Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τομέας Πληροφορικής και Τεχνολογίας Υπολογιστών

Νεφο-υπολογιστικές αρχιτεκτονικές για τη διαχείριση βιο-ιατρικών δεδομένων

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Χρήστος Ο. Ανδρίκος

Αθήνα, Μάιος 2022

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Τομέας Πληροφορικής και Τεχνολογίας Υπολογιστών

Νεφο-υπολογιστικές αρχιτεκτονικές για τη διαχείριση βιο-ιατρικών δεδομένων

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Χρήστος Ο. Ανδρίκος

Συμβουλευτική Επιτροπή: Παναγιώτης Τσανάκας
Νεκτάριος Κοζύρης
Ηλίας Μαγκλογιάννης

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 17η Μαίου 2022.

....
Τσανάκας Παναγιώτης   Κοζύρης Νεκτάριος   Μαγλογιάννης Ηλίας
Καθηγητής   Καθηγητής   Καθηγητής
Εθνικό Μετσόβιο Πολυτεχνείο   Εθνικό Μετσόβιο Πολυτεχνείο   Πανεπιστήμιο Πειραιώς

....
Κουτσούρης Δημήτριος-Διονύσιος   Παυλάτος Χρήστος   Ματσόπουλος Γιώργος
Καθηγητής   Επίκουρος Καθηγητής   Καθηγητής
Εθνικό Μετσόβιο Πολυτεχνείο   Σχολή Ικάρων   Εθνικό Μετσόβιο Πολυτεχνείο

Σούντρης Δημήτριος
Καθηγητής
Εθνικό Μετσόβιο Πολυτεχνείο

Αθήνα, Μάιος 2022

. . . .

Χρήστος Ο. Ανδρίκος

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών

Εθνικό Μετσόβιο Πολυτεχνείο (2022)

Σε εκείνους που δουλέψαν μαζί μου, και σε αυτούς που θα ακολουθήσουν ...

Εν συντομία

Η παρούσα διατριβή υπερασπίζεται μια νέα διαδικτυακή πλατφόρμα που προωθεί προηγμένες υπηρεσίες τηλε-διαβουλεύσεων σε πραγματικό χρόνο για την απεικόνιση ιτρικών δεδομένων. Οι αρχές των ετερογενών Συστημάτων Διαχείρισης Ροής Εργασίας (WFMSs) και οι τεχνολογίες αιχμής, όπως το αρχιτεκτονικό σχέδιο μικροϋπηρεσιών, η δικτύωση peer-to-peer και η ιδέα της Εφαρμογής Ενιαίας Σελίδας (SPA) συνδυάζονται για τη δημιουργία μιας επεκτάσιμης πλατφόρμας για να καταστήσει τη συνεργασία μεταξύ γεωγραφικά κατανεμημένων επαγγελματιών υγείας εφικτή αλλά και αποδοτική. Οι δυνατότητες επικοινωνίας σε πραγματικό χρόνο βασίζονται στο πρωτόκολλο webRTC που επιτρέπει την άμεση επικοινωνία μεταξύ των πελατών. Η εργασία συζητά τις εννοιολογικές και τεχνικές λεπτομέρειες του συστήματος, δίνοντας έμφαση στα καινοτόμα στοιχεία του. Εμβαθύνει σε έννοιες αρχιτεκτονικής και ορθών πρακτικών σχεδιασμού πλήρως κατανεμημένων υπηρεσιών. Αξιοποιώντας τις αρχές των εφαρμογών δώδεκα-συντελεστών καταφέρνει να προσφέρει μιά λύση η οποία:

- Αξιοποιεί το δηλωτικό μοντέλο περιγραφής πόρων για να επιτύχει αυτοματοποίηση

- Διατηρεί μια ξεκάθαρη σχέση με το υποκείμενο λειτουργικό σύστημα προκειμένου να φέρει υψηλή μεταφερσιμότητα

- Παρέχει δυνατότητες συνεχούς ανάπτυξης

- Φέρει χαρακτηριστικά υψηλής κλιμακωσιμότητας

Επιπρόσθετα. ενσωματώνει προηγμένους αλγορίθμους νευρωνικής μάθησης για την επεξεργασία και τρισδιάστατη αναπράσταση δεδομένων ιατρικής εικόνας και RNA ακολουθιών προτείνοντας παράλληλα μεθοδολογίες που ανοίγουν το δρόμο στην καθιέρωση νέων πρακτικών στον ευρύτερο κλάδο της βιο-ιατρικής επιστήμης.

Abstract

The present dissertation defends a novel web-based platform promoting real-time advanced tele-consultation services on medical imaging. Principles of heterogeneous Workflow Management Systems (WFMSs) and state-of-the-art technologies such as the microservices architectural (MSA) approach, peer-to-peer networking and the Single-Page Application (SPA) concept are combined to build a scalable and extensible platform to aid collaboration among geographically distributed healthcare professionals. The real-time communication capabilities are based on the webRTC protocol to enable direct communication among clients. Implanting the twelve-factor application principles, we end up with a solid system that:

- Uses declarative format for setup automation

- Features a clean contract with the underlying operating system to offer maximum portability between execution environments

- Minimize divergence between deployments to empower continuous deployment (CD) and enhanced agility

- Scales out/in without changes to tooling, architecture or development practices.

The system features increased extensibility. The domain-driven-design we imply lets us explore multiple biomedical applications. In this sense, we introduce two novel algorithms. The first one concentrates on the three-dimensional reconstruction of the urinary bladder. The second one focuses on predicting patterns in RNA sequences. Both algorithms are integrated into the original system to push biomedical engineering frontiers further.

The work discusses the conceptual and technical details of the system, emphasizing on its innovative elements.

Contents

List of Figures

List of Tables

Preface

At the time being most of us are working from home while Slack CEO complains that Microsoft is trying to kill Slack business. And this is due to the COVID-19 outbreak. It is remarkable how quickly we have adapted. Most of us used to commute to our work every single day, to travel abroad to participate in conferences or to teach in congested amphitheatres and classes. A couple of months ago we used our pc screens mostly for private work. Web calls used to be significantly fewer than the exchanged emails. But all these seem like an ancient era shifting the meaning of BC to Before the Coronavirus crisis. That shift may rival great workplace transformations of the 19th and 20th centuries. Dominant tech companies announced that the work from home policy is going to be widely adopted; Twitter has already said that all its employees will be allied to work from home permanently while Facebook expects half its staff to do so within a decade.

The teleworking transition has been more abrupt than the one that occurred by the factories back in the 1800s. At those times steam power was snaked through a single building to move belts and pulleys around. People had to be part of that system in-person to contribute with their own work. Later on, electricity enabled machines to have their own power supply, but factories required almost half a century to be redesigned to take advantage of that new power source. Those days humanity used to adopt the technology trends. Nowadays, technology is disrupting humanity.

Apparently, teleworking became part of our presence due to specific preconditions. First things first, the latest technological advances have matured enough. Broadband connections, along with cloud services and high-end portable devices, empower videoconferencing along with decent file-sharing services. That is considered a necessary but not sufficient precondition. The real transition advocate is the fact that advanced economies revolve around services instead of manufacturing. The fourth industrial revolution is said to be the AI one. But for me, it is the democratization of services. And it is actually already happening. The pandemic is the condition that triggered that instant change from an office-based to a home-based working force. Being severe, it obeyed the humanity to stay at home and to mainly work from home. We were expecting that our time at home may be affected, but this was not the case. At least at the degree, we did expect.

The exciting part of that is that pandemic made remote work seem not only reasonable but highly acceptable. The suspicion that 'remotees' (remote employees) are bunking off has been replaced with the perception that the ones insisting being at office sound self-important.

Of course, things are missing for a completely remote-working environment. The impulse of in-person meetings cannot easily be substituted by any video conferencing engine; creativity is presumably harder to foster. But is it? In our case, I am not sure that these general rules do apply.

The scope of my thesis is to provide a sufficient plan for healthcare transition to the new teleworking era. I personally aspire to exhibit the way of disrupting the well known centralized healthcare systems transitioning them into a cloud of distributed knowledge and automation via an efficient and performant digital collaboration environment.

Yet although hospitals and clinics will not disappear (I also hope not to), it is hard to imagine the reason why health care practitioners and patients have to get crowded together into cost-aggressive and health-risking structures such as large hospitals. I am not sure why geographical barriers have to deteriorate the quality of the offered healthcare services or the reason why automated diagnosis have not to be part of the modern healthcare providing routine.

Many industries have already their "Wizard of Oz" moment: the old school HQ is shown to be an old man behind the curtain. Centralized offices may vanish. Why do we have to keep a centralized way of running the healthcare industry? Pandemic taught us a hard and literally expensive lesson: the modern healthcare system is already obsolete due to lack of scalability. What we have actually figured out is that Covid-19 would not feature a high mortality rate if we were able to afford large scale hospitalization. Large and centralized hospital units, however, are not able to provide such agility. That's why Hermes has been built...

CHAPTER **1**

Introduction

Nowadays, significant market forces are dramatically reshaping the landscape of any industry, even those that have historically been deliberate and methodical in its adaptation to change. The effect of standard business disruption, caused majorly by digital technologies, is forming a new global service mesh network, a result that I call "service democratization". In this setting, healthcare companies are challenged to adopt more thoughtful product development and supply chain lifecycle management practices to not only survive but grow and defend market share. The shift in the United States to value-based care (VBC) model, in which healthcare cost reimbursements are being driven by the efficacy of treatment, along with the entrance of non-traditional players to the marketplace (i.e. digital native competitors eager to test-drive their expertise in the healthcare domain) are the most prominent indicators of the upcoming disruption.

In healthcare, the end-goal has always been to protect the patient and help improve patient outcomes. This has, historically speaking, anchored therapy as the primary driver of ongoing product iterations, with patient individuality taking a back seat. To the best of my knowledge, the last time that practice has been strongly revisited was back in 1910 when William J. Mayo (one of the Mayo brothers and founders of Mayo Clinic) in his commencement address to Rush Medical College declared:

"It has become necessary to develop medicine as a cooperative science; the clinician, the specialist, and the laboratory workers uniting for the good of the patient, each assisting in the elucidation of the problem at hand, and each dependent upon the other for support."

1.1 Contribution & impact

In a particular thesis, we revisit that concept once again. we argue that medicine should be updated to a cooperative but broader science in the manner of considering the participation of clinicians, medical experts and laboratory experts along with data scientists and algorithms in a scalable fashion that eliminates any geospatial or time-related barriers. Thus, we introduce a novel web-based platform offering:

1. real-time collaboration or standalone sessions

2. coherent PACS integration

3. medical imaging out of the box

4. integrates novel artificial intelligence algorithms to incorporate

- computer aided diagnosis
- wet lab data analysis

to unleash the holistic approach of transdisciplinary healthcare provisioning. The platform highlights a unique architectural design based on state-of-the-art web and cloud technologies. It utilizes principles of three of the core aspects of modern computer science (Figure 1.1). It eliminates the overall server-side footprint to promote scalability and enhance user experience; Hellenic Radiological Society evaluated the work quite positively and corroborated the impact of its application to the modern healthcare system.
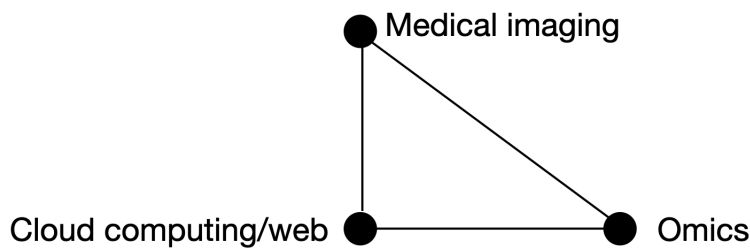
Figure 1.1: Thesis center of focus

Today's physicians acknowledge that the more they involve other disciplines in their practice, the more likely their patients will receive advanced healthcare services. Personalized healthcare services do require teamwork in a multidisciplinary or even transdisciplinary way. When healthcare providers from multiple disciplines collaborate and share ideas from the beginning to create a complete health care plan that covers all necessary diagnosis and treatment for the patient, the end result should be a treatment plan as unique as the patient himself. Nowadays this idea may be extended further to include inter-domain collaboration. Computer scientists, biomedical engineers, wet lab experts can join their forces to tackle to tackle the toughest challenges and shape the modern landscape of healthcare provisioning. Besides teamwork, automation onward with decision support tools, may reveal new points of view that correlate to any specific patient case or are driven by pure large scale statistics. Plenty of AI algorithms do outperform into medicine tasks, such as medical image annotation or even more complex challenges. Still, they rarely find their way to the front line of medicine while the adoption of a transdisciplinary based treatment planning is also cumbersome. Specific technical barriers are affecting both cases. Poor infrastructure, the geographic distribution of medical personnel, scientists or data, schedule conflicts or even the lack of a common communication channel prevent the full exploitation of such a healthcare upgrade. Small or regional hospitals usually are short of equipment or specialized healthcare personnel discouraging the transdisciplinary approach. In contrast, crowded metropolitan hospitals suffer from heavily scheduled staff, failing to promote productive transdisciplinary teams (TTs).

In this context, we introduce a novel platform that eliminates all these barriers. By adopting modern web technologies and cloud computing methodologies to provide powerful collaboration

sessions to the geographically distributed and over-scheduled healthcare workers. It offers an advanced and extensible image analysis toolkit to meet the needs of modern medical diagnosis.

The second contribution of our work lies in the extension of this toolkit to incorporate a novel algorithm for the visualization of the urinary bladder anatomy in the 3D space. The proposed methodology utilizes information depicted on ultrasound and photoacoustic medical imaging modalities to extract the inner and the outer wall of the urinary baldder respectively. Given the fact that the acquisition of the two image modalities is a based on non invasive techniques and that the overall 3D space mapping methodology achieves a remarkable accuracy ($err <$ $0.05mm$) the value of this work is uncontested.

A detailed survey conducted on a total of 368 radiology professionals exhibited that 65% of them were using some kind of teleradiology [RB13]. The main usages were in-house image distribution (71%) and on-call readings from home (44%). The key benefits are improved collaboration with other radiologists (46%) and efficient distribution of workload (38%). Outsourcing is performed by 35% of the participants, while among them, 68% noted that they use commercial services. The significant advantages of outsourcing are the availability of second opinions (82%) and additional capacity for on-call services (71%). The drawbacks hold insufficient integration of patient history and priors (69%), and limited communication with clinicians (68%). The majority expressed a positive opinion regarding the future of teleradiology (80%) predicting growing importance (46%). Opportunities ought to be found in emergency reading services, flexible support of small practices and in collaborative platforms.

Lately, collaboration in the electronic health (eHealth) environment is focused on the co-operation of healthcare professionals belonging to the same organization via the Computer-Supported Cooperative Work (CSCW) paradigm as presented in [DJ10] and Share4Health [sha]. The majority of the existing infrastructures have been deployed as in-house, centralized systems featuring point-to-point communication between medical servers and clients. However, in practice, collaboration is proved to be an inherently challenging task imposing a variety of additional requirements, such as latency, flexibility, scalability, tolerance and robustness which are usually not met by the common centralized architectures. Typical screen-sharing or accessing a machine remotely means that only one person is in control while the other is only observing. Instant messaging, email or other tools are great for primitive collaboration but collaborating on real medical data often requires more than an image or a single file to share the necessary context. Additionally, the validity of analysis can be affected by hardware capabilities, while trust is also of high importance, as privacy and security issues are essential in medical data handling. Thus solid data protection policies are required to cover strict rules like EU's impending GDPR regulation (https://www.eugdpr.org).

The success of any collaboration technology supporting any distributed medical team relies on the understanding of the human-computer interaction entailed in the technology and a socio-technical approach in the design process [RRTN06] [Ack00]. Any MDT requires an appropriate electronic collaboration infrastructure tailored to standardized work environments, so as to improve efficiency and productivity by tackling any barriers. In this context, we present a platform that provides remote situation awareness and shared artefacts exploring a variety of settings. By utilizing video teleconferencing techniques alongside a rich graphical user interface (GUI), the communication among different specialists is being enhanced to eliminate any MDTM limitations: the meetings are easily integrated into the work schedule of specialists,

while small or district hospitals need no special facilities for teleconferencing to support remote MDTMs. The proposed platform may also be utilized in teaching medical students and junior doctors [Gao13] or in supporting research groups to jointly solve related issues such as image-based co-registration [WEC+13] or co-segmentation [HFW+11] [SNH+09] [Må06]. It features an extensible design able to integrate inter and intra domain applications.

In this setting the third point of contribution of our works lies in the Introduction of a novel approach for wet lab results analysis. In particular in our work in our recent work in [AMK+22] we discuss the analysis of the ribonucleic acid (RNA) using syntactic pttern recognition techniques. Obtaining valuable clues for noncoding RNA subsequences remains a significant challenge, acknowledging that most of the human genome transcribes into noncoding RNA parts related to unknown biological operations. Capturing these clues relies on accurate "base pairing" prediction, also known as "RNA secondary structure prediction." As COVID-19 is considered a severe global threat, the single-stranded SARS-CoV-2 virus reveals the importance of establishing an efficient RNA analysis toolkit. This work aimed to contribute to that by introducing a novel system committed to predicting RNA secondary structure patterns (i.e., RNA's pseudoknots) that leverage syntactic pattern-recognition strategies. Having focused on the pseudoknot predictions, we formalized the secondary structure prediction of the RNA to be primarily a parsing and, secondly, an optimization problem. The proposed methodology addresses the problem of predicting pseudoknots of the first order (H-type). We introduce a context-free grammar (CFG) that affords enough expression power to recognize potential pseudoknot pattern. In addition, an alternative methodology of detecting possible pseudoknots is also implemented as well, using a brute-force algorithm. Any input sequence may highlight multiple potential folding patterns requiring a strict methodology to determine the single biologically realistic one. We conscripted a novel heuristic over the widely accepted notion of free-energy minimization to tackle such ambiguity in a performant way by utilizing each pattern's context to unveil the most prominent pseudoknot pattern. The overall process features polynomial-time complexity, while its parallel implementation enhances the end performance, as proportional to the deployed hardware. The proposed methodology does succeed in predicting the core stems of any RNA pseudoknot of the test dataset by performing a 76.4% recall ratio. The methodology achieved a F1-score equal to 0.774 and MCC equal 0.543 in discovering all the stems of an RNA sequence, outperforming the particular task. Measurements were taken using a dataset of 262 RNA sequences establishing a performance speed of 1.31, 3.45, and 7.76 compared to three well-known platforms. The implementation source code is publicly available under knotify github repo and an instance has been deployed to extend the introduced platform functionality in inter domain fashion.

Since the thesis discusses the three pillars presented above (Contribution & impact) the overall structure of the text come into the following parts: Chapter 2 provides a thorough survey of literature related to the higher collaboration concept in the domain of medicine. It also discusses the value of multidisciplinary approach to modern healthcare landscape. Chapter 3 discusses the principles of the design of the introduced platform along with explicit implementation details. Backend details of the system along with client side architecture and deployment strategies fill most of the lines of chapter 3. The chapter concludes with a solid evaluation of the introduced platform platform. Next Chapter 4 presents our novel approach for the 3D reconstruction of the urinary bladder. We present the fusion of two medical imaging modalities that are acquired in

a non-invasive fashion along with the basic principles that drove our methodology. The chapter concludes presenting the related results and a detailed evaluation of the suggested approach. Chapter 5 presents a novel strategy for RNA sequence exploration that our platform incorporates. The chapter highlights the value of the particular challenge along with a thorough survey of the related literature. It presents the introduced approach and concludes with a thorough comparison of our approach against the top performers that bibliography suggests. Finally, the thesis concludes the text chapter 6 presenting the most significant closing remarks along with future work directives. Reader is welcome to check our appendix for any work in progress that highlights the future extension of the platform we suggest.

CHAPTER 2

Related Work

Computer science always flirted with medicine in a mix and match way. Biomedical imaging has far been one of the hottest domain aspiring to combine computer science and medicine to one specialized hyper scientific field. AI algorithms aspire to automate biomedical imaging tasks such as detection of abnormalities at Computer Tomography (CT) images, image annotation and captioning. Database and system researchers have spent millions of hours to come up with universal solutions to manage and analyse medical data, while statisticians hustle to reveal correlations that hide beneath the constantly growing (bio)medical data silos.

That is how we ended up with PACS, a technology which provides economical storage and convenient access to images from multiple modalities (source machine types).

Collaboration services in medical imaging have a long history of systems and tools. Researchers have produced various research papers, along with a plethora of real-world products and standards. At the beginning of 1980s, American College of Radiology (ACR) developed Digital Imaging and Communication in Medicine (DICOM), a standard for handling, storing, printing and transmitting information in medical imaging [MEM02]. Since its introduction, DICOM has been expanded to support new applications [BHPVS97] while multiple studies explored its potential use in innovative use cases [GKK$^+$02][GPS05][CFB$^+$11].

Due to the proliferation of the web, DICOM committees recognized the necessity of a web medical standard. They specified WADO (Web Access to DICOM Objects) service [KL06], so that system interaction could also take place over the web in a standardized way, allowing interoperability and proper information management within PACS (Picture Archiving and Communication Systems).

Apparently, such a coherent standard was able to facilitate the development of various biomedical imaging applications and to inspire the early "Internet of Things" (IoT) for medical devices. On the one hand, in hospitals and clinics equipped with PACSs, any images or reports are transmitted digitally via them, eliminating the need for someone to file, retrieve or transport film jackets manually. On the other hand, healthcare professionals started to use specialized software packages (DICOM viewers) such as HTML5 Zero Footprint DICOM viewer (the latest medical imaging product by LEADTOOLS) [htm], Synapse 3D [sun] by Fujifilm, OSIRIX [RSR04] by the University of Geneva, Radiant or ImageJ [BPYM05] etc.

A brief inspection of the long list of DICOM viewers [HPD16][HPKD15] shows that rich functionality and suitable interfaces can be found in the commercial as well as in the open-source

domain of DICOM viewers software. However, each tool has its strengths and weaknesses; an all-rounder solution does not exist.

In the introductory chapter, the value of collaboration among healthcare professionals has been confirmed. Furthermore, a survey of the literature reveals two outstanding initiatives adopting early Internet technologies to support this claim, INTERMED and TeleMed. INTERMED was a pioneering project designed to enhance communication among patients and multidisciplinary health providers. INTERMED proved to be a reliable method for classifying patients' care needs, primarily when used by experienced raters scoring by patient interview. It can be considered a useful tool in assessing patients' care needs as well as the level of needed adjustment between general and mental health service delivery [HLS$^+$99][SdJH$^+$99]. TeleMed was a promising visualized, action-oriented decision-support tool, providing collaborative sessions and access to primary patient data to physicians [KF97]. However, the insufficient hardware, the poor connectivity infrastructure, combined with the lack of sophisticated software solutions, made either of those ventures almost inviable. The work in [LFC00] presents the development of a web-based collaborative system for medical image analysis and diagnosis aiming to equalize healthcare delivery services among areas of higher and lower population density in Australia. The proposed platform was implemented in Java as a set of applets and Common Gateway Interface (CGI) scripts. In [MDK06], a collaborative working environment for physicians is introduced, enabling a peer-to-peer exchange of electronic health records based on the HL7 standard [DAB$^+$01].

In this context, the thesis treats technological issues such as video, audio and message communication, workspace management, distributed medical data management and exchange. At the same time, it emphasizes the security challenges because of the sensitive and private nature of the medical information. The novel architecture of the proposed system taking advantage of a minimalistic microservices approach aspires to provide a facelift on developing, deploying and maintaining large scale cooperative systems. The rich domain-related functionality does tackle the multiple aspects of modern transdisciplinary approach in medicine; RNA sequencing is just an example of the provided integration freedom that the system supports.

The authors of [LLH09] depict a universal method to integrate computer-aided detection and diagnosis (CAD) results with PACSs to assist physicians in the decision making process. A medical image atlas is described in [MY14]. The researchers developed a platform allowing registered users to upload, visualize, process and collaboratively comment medical images by using Web 2.0 Personalized Learning Environment tools for social learning and knowledge construction and sharing. Similarly, the work in [MBA13] exhibits the way that a social network can be provided to physicians and patients, to foster cooperation and to overcome the problem of unavailability of doctors on site any time; patients can submit their medical images to be diagnosed and commented on by several experts instantly. Another collaboration network intended to gather patients' medical images and physicians' annotations expressing their medical reviews and advice, is proposed in [BAAH16]. Authors in [MV08] present a potential solution on integrating grid computing and DICOM to feature advance medical imaging capabilities and security while researchers in [SDO10] introduce a virtual laboratory for medical image analysis; the platform is based on Dutch grid and adopts a service-oriented architecture to decoupling the user-friendly clients running on the user's workstation from the complexity of the grid application. Finally, P2Care, motivated by real-world social groups, organizes healthcare professionals, providers and patients in groups according to shared characteristics and interests [ABL15] [ABL13].

Last but not least, three related patents have been filed so far to demonstrate the value of collaborative medical environments further. The first one refers to a system that aggregates medical images via an uploading application; the system can further transmit the acquired images to the plurality of client applications to be displayed [VG11]. The second one presents a simple browser-based medical imaging system that requires no installation of application software or any browser plug-in and functions in the same way as traditional full-blown medical imaging PACS viewer fat clients [Xie12a] [HFW$^+$11]. At the same time, the third patent depicts the idea of a web-based medical collaboration platform [Ahm17] between two browser-based clients. It is noteworthy that the patents have been filed after our initial paper publication related to the specific work.

Despite the long list, the surveyed applications found in literature do not meet the requirements of a universal approach that enhances user experience. Medical image data exchange is a bandwidth-intensive task, due to the size of imaging data as well as the demand for high QoS (Quality of Service) and minimum response times. Thus server-based architectures that assume subsequent server operations to produce intermediate image processing tasks to be rendered by a client app may introduce severe bottlenecks that lead to reduced user experience. Moreover, the preponderance of the existing systems are device-dependent or highly coupled with the DICOM standard while they do not support the collaborative creation of workflows over shared medical imaging data. Consequently, our contribution relies on the complete integration of both concepts —collaboration services and biomedical imaging— into a web-based real-time platform tackling all those issues and limitations, modern transdisciplinary medicine suffers from—the collaboration among healthcare professionals is going to upshifted to the new digital era.

Moreover, an extra leap is made which is not directly related to medical imaging but lies into the broader biomedical data visualization and analysis field; a novel RNA folding algorithm has been proposed as part of the current thesis. That is to demonstrate the functionality wise scalability of the platform.

The work has been progressively communicated to the biomedical imaging and data analysis community; initially, we proposed an intelligent collaboration system fusing the principles of heterogeneous WFMSs, AI, existing electronic medical infrastructure (PACS, EMR) and standards (HL7, WADO) [AMB$^+$14] [ARTM15a]. Next, the work focused on the conceptual and technical details of design and implementation process, delving deeper into the overall system architecture and the way it facilitates the provision of easy-to-use, transparent and interoperable online MDTM capabilities and in the last revision the RNA analysis toolkit is going to be integrated to the platform.

The presented platform leverages state-of-the-art web technologies (WebRTC, HTML5, CSS3) to provide rich and fluent communication (audio/video streams) and real-time interaction mechanisms combined with medical image processing functionalities. It features the concept of distributed state sharing among clients to eliminate any server-side footprint. The novel architecture of the proposed infrastructure favours all state-of-the-art technological advances (i.e.domain driven development paradigm, micro-services implementation, dockerization) that now seems to start getting in the core of software production pipeline. Practices and tactics that have been proposed are aligned with products and scientific directions that have been introduced by tech giants far later than our initial designs.

The motivation of the specific work can be summarized into (a) the benefits for public health that stem from interconnecting MDT members efficiently and transparently, (b) the advantages of peer-to-peer networking and its suitability for developing real-time collaboration services as it features robustness, scalability, dynamically adapted system architectures and low-cost, (c) the maturity of the web technologies offering advanced programming capabilities to facilitate the deployment of image processing algorithms and (d) the lack of academic research as well as commercial solutions in the specific domain. The integration of a novel RNA folding exploration algorithm should be considered a bonus remark.

The main innovative features of the proposed platform are the exploitation of client-side computing (fat client) paradigm as well as the efficient cloud utilization; a Single Page Application (SPA) based on the newly introduced WebRTC standard, adopting the most recent web technologies, enables users to communicate, share and work on medical images simultaneously [KGS+13]. Everything is orchestrated by a highly scalable ecosystem of fine-grained backend microservices. In this way, clinicians can collaborate in real-time following intelligent pathways using a medical image processing toolbox on any device without any installation concerns. Besides, the proposed system is easily extensible to offer additional functionality, and it does not suffer any severe burden of maintenance. That is why an RNA analysis toolkit has been built on top of that with ease!

CHAPTER 3

System Architecture

In a high level, the architecture of the proposed system is illustrated in figure 3.1. It involves a federation of modules to provide (1) real-time tele-consultation, (2) data management, (3) computer assisted diagnosis services to the clinical personnel. The system users are allowed to use medical collaboration services (i.e. real-time data sharing and video conferencing services, dedicated image processing algorithms, medical data analysis tools) along with instant messaging and file sharing on the browser upon registering to the system. Registration may take place in a standalone, OATH2 or SAML based fashion [Nes19]. Besides, it tracks personal information about the registered user and about potential roles along with auditing information and access credentials. The latter may be revoked on the fly. Users can also upload medical data to a dedicated PACS system into a sharable or private-only mode (Figure 3.1). In this regard, User Access Control (UAC) module focuses on user authentication and authorization, either by interacting with some private data structures or by negotiating with some third party Authentication and Authorization Infrastructure (AAI). Users are able to enter the platform by registering their credentials explicitly (typical sign-in schema) or implicitly (single sign-on schema). Next, User Management (UM) enumerates all registered user profiles securely as well as manages any social activity of theirs (i.e. collaborative circle, networking, friend requests). It is also responsible for the user role assumption; every user of the system affords a set of dynamically assigned credentials (role) that restrict the resources he/she is able to consume [1] (i.e. medical data, tools, users and other). Content Management (CM) defines the means and the policies behind content delivery. It facilitates medical data management and sharing and auditing [2] in offline mode, based on the assumed user roles and authorization policies. The whole system features strict security requirements due to highly sensitive medical data manipulation and exchange. Thus, it leverages encryption protocols in all communication stages. Futhermore, it implements a novel distributed secure authorization schema of certified communication 3.4.2 The following paragraphs outline the architecture of the system in detail. The chapter focuses on the design and the implementation details while justifying all the decision made via thorough literature examination. Our approach to designing and deploying cloud services aspires to provide a standardized way to perform into the modern cloud-enabled ecosystem.

---

[1] The word resource is more like a wildcard in computer science. It is context defined, and its meaning alters based on the domain referring to.

[2] The process of logging access to sensitive information
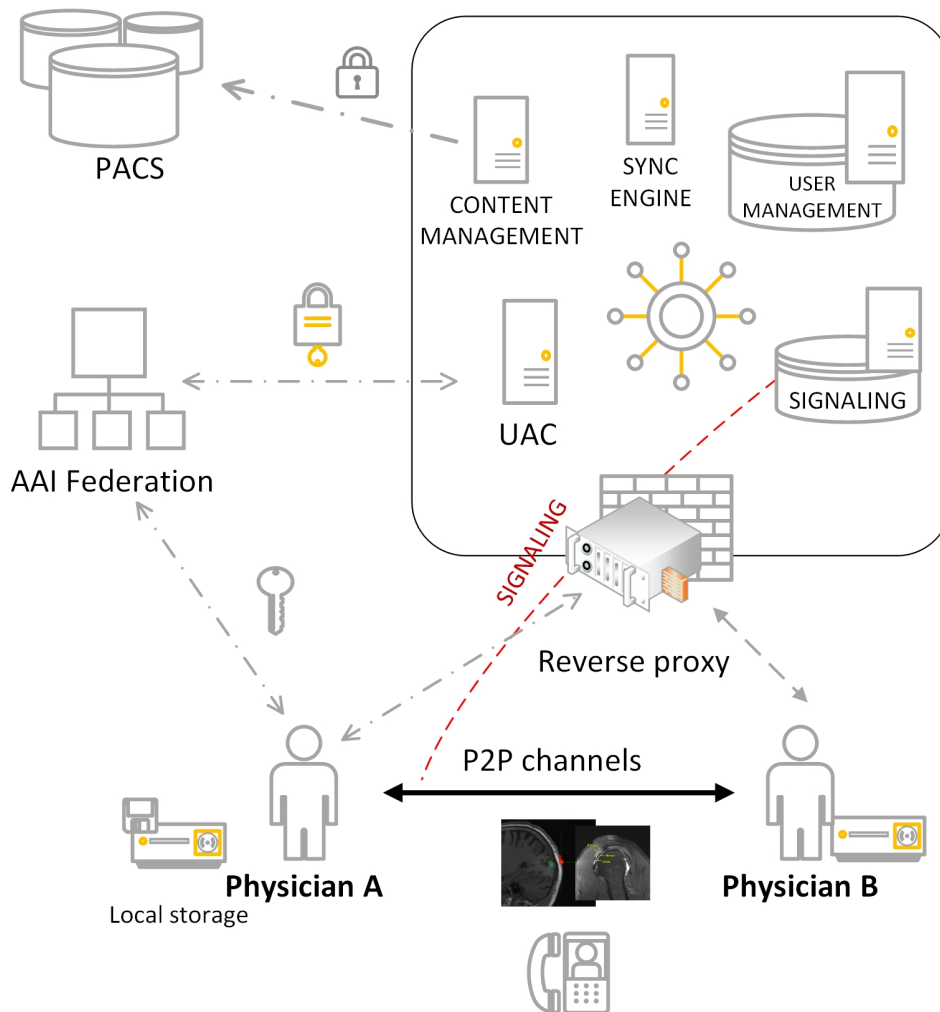
Figure 3.1: Overall system architecture

3.1 Background information

Before delving into the "subatomic" level of the micro-scale architecture of the introduced system, we need to explore the microservices design approach. In general, microservices represent an evolution of Service-Oriented Architectures (SOA) used to build (a) **flexible,** (b) **independently deployable** software systems.

In a microservice architecture (MSA), services are processes that communicate with each other over a network in order to fulfil a goal. Each microservice-based system consists of a federation of autonomous, self-sufficient processes. In this terms, any process may be replaced by a completely new one featuring the same communication interface. The deployment of any service (process) completely decouples the deployment of any other one, while the whole system deployment maintains the **independent** aspect (any deployment should outcome the same results that any other one ought). (Un)fortunately, there is no single definition when it comes to

the MSA approach. Despite it is considered being more an art than science, there is at least some consensus that evolved in the industry over time is highlighted in the next six ideas:

- Services in a microservice architecture (MSA) are often processes that communicate over a network to fulfil a goal using technology-agnostic protocols such as HTTP. [Fow18][New15a][Wol16]

- Services in a microservice architecture are independently deployable. [PZA$^+$17][Che18b][Che18a]

- Services are organized around business capabilities and operations.[NMMA16] Domain-driven development is a concept that may take advantage of the microservices paradigm. In this setting, any microservice should be responsible for the processes of at most one business domain.

- Services can be implemented using different programming languages, databases, hardware and software environment, depending on what fits best. (polyglot effect).[Che18b]

- Services are small in size, messaging-enabled, bounded by contexts, autonomously developed, independently deployable, decentralized and built and released with automated processes.[PZA$^+$17]

- Services do communicate over "dump" pipes through "smart" endpoints. In this context every microservice is responsible for sharing information with each peers considering any underlying communication channels to be unreliable.

3.1.1 From monoliths to the democratization of services

To start explaining the microservice style and the reason why it is adopted in this work, a comparison to (a) the typical monolithic along with (b) the standard SOA approach should be enlightening.

Any **monolithic application** is built as a concrete, single-tiered software application which is self-contained, and independent from other computing applications. That design pattern is based on the conception that *the application is responsible not just for a particular task but for every step needed to complete a particular function*. Designing monolithic systems does not necessarily involve a lack of code reusability and flexible design patterns. Object-based development is generally considered to be system development at a low level of granularity [Sny93] that plenty of modern monolithic systems feature. The evolution of object-based development is the *component-based* one, according to which a number of objects are packaged together as a solid *component*. The key benefit of component based design is the *increased code reusability* as consequence of working at a relatively higher level of granularity.

Next, in the Service Oriented Architecture (SOA), *a service encapsulates an arbitrary number of components into a single interface to provide a discrete business function.* Services work in an even higher level of granularity than the components with the latter considered being a service if it is wrapped in a service layer [AZDR10]. A service is consumed through late binding at runtime [AZDR10] to increase the decoupling of codebase versioning (i.e. any service version may be instantiated and deployed transparently to all other services of the system). In the

matter of distributed systems, enterprise firewalls incommode any inter-component communication. Thus, service layers do enable industry-standard (widely accepted protocols) to promote interoperability.

On the other hand, Microservices epitomises the decentralisation of the software design. Isolating business functionalities while decoupling data representation aims at enhancing the autonomy and replace-ability of the individual microservice(s) [Fow18]. This, in turn, *enhances decentralised governance of the microservices*; each microservice encapsulates a subset of fine-grained business functionality. By featuring a clearly defined interface, a network of microservices is established to facilitate any complex business workflow. In this setting, a series of microservices are chained together or coexist, one next to another, to perform a high-level business function. To enable these characteristics, each microservice has its own data model and a class model preserving its own perspective regarding the business world they are built to operate on.

(a) SOA based design   (b) MSA based design
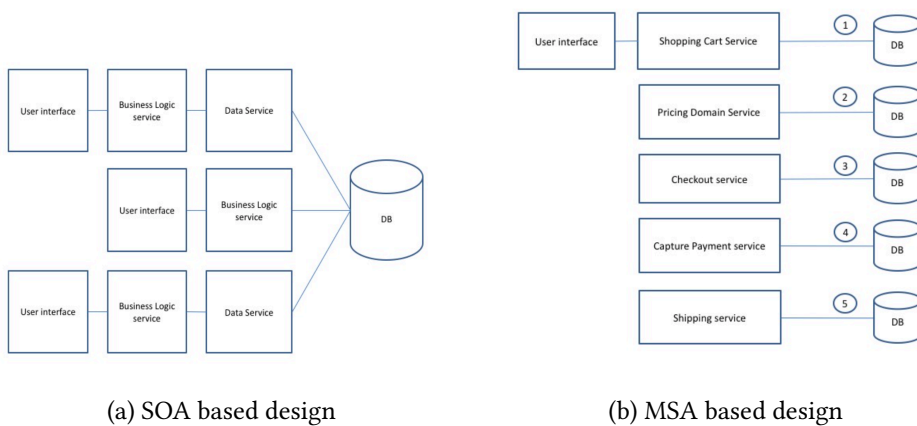
Figure 3.2: SOA vs MSA comparison

Figure 3.2 depicts the design difference between the approaches of SOA and Microservices respectively. Considering the "bounded context" concept of the Domain-Driven Development (DDD) literature [Eva04] any microservice may be defined as a "loosely coupled service in a bounded context". The concept of "bounded context" captures the key properties of a microservice architecture: the focus upon business capabilities, rather than program code decomposition and reuse (key incentives of monoliths and SOA). Such a perspective supports the capture and modelling of requirements in complex multi-agency domains such as the delivery of community healthcare [BHH03], [HPB05], applications for the Internet of Things [IAH+15] or in our case the development of a web platform to assist real-time collaboration of medical experts over the web via jointly biomedical data analysis.

3.1.2   Services bounded to contexts

Domain Driven Design (DDD) [Eva04] focuses on modeling real-world domains into a neat yet lean software designs. DDD is based on interesting ideas such as the use of universal language, repository abstraction and others. The most interesting and probably the fundamental idea of DDD, though, is the concept of the bounded context. In DDD domains define the sub-problem space, while bounded contexts are the distinct sub-solutions; it is considered to be a divide and

conquer methodology applied in the space of higher abstractions such as business logic, services, clients or any other generic in regards of implementation entity.

The idea is that any given domain consists of multiple bounded contexts, and residing within each are things that do not need to be communicated outside as well as things that are shared externally with other bounded contexts. Each bounded context has an explicit interface, where it decides what models to share with other contexts. Moreover the co-located entities (i.e. entities residing into the same domain model) have a coherent perspective of the world laying out of the particular context, a fact usually not true considering the inter-context coherence. Despite that, all system modules should be eventually be consistent in terms of outcomes regardless the context they are bounded by. For example in the case of two modules of distinct models performing transactions to the same databse table, the transactions should end to a consistent database state (i.e. database integration pattern).

Another definition of bounded contexts is: "a specific responsibility enforced by explicit boundaries"; any information request from a bounded context, or functionality request within a bounded context, involves communication with its explicit boundary using models. The author of [Eva04] uses the analogy of cells, where "cells can exist because of their membranes". Similarly, bounded context defines what is in or out of the scope of a particular service. It also determines what can move inwards or out of it.

Multiple models are in play on any large project. Considering that every model applies in a single well defined and strictly bounded context eliminates multiplexing models, software bugs, unreliable, and difficult to understand code routines. Team members communicate over sound channels while non-ambiguous terminology vanishes.

Let us consider the operation of the introduced system and some "cells" of it. Our domain is the whole business in which we are operating. It covers everything from user registration to the medical digital rooms, from text messaging to advanced syncing algorithms. Apparently, we may not model all of that in our software, but that is nonetheless the domain in which we are operating. Let's think about parts of that domain (cells) that look like the bounded contexts that DDD refers to. UAC, for instance, is a traditional warehousing service that manages user objects. Registration, login, AAI verification, user profiles and other. At the core of UAC resides User which shapes the perspective of the entire domain from within it. Sync Engine is a hive of activity—managing synchronization requests by clients, while Content Management distributes any content to the clients. Table 3.1 overviews the conceptually autonomous components (cells) of the system.

Any attempt to define the bounded contexts, in which our software models will reside, will probably start with a number of coarse-grained bounded contexts. These bounded contexts may in turn contain further bounded contexts. Considering that we introduce a system providing real time tele-consultation services,one of its core bounded contexts is the Synchronization capabilities. These, can be further decompose into capabilities associated with call singling, those being associated with the workspace syncing (e.g. joint annotations) and those associated with the asynchronous operations (file uploads and other).

Considering the boundaries of our microservices, first, we thought in terms of the broader, coarser-grained contexts, and then we subdivided along with these nested contexts while looking for the benefits of splitting out these closures. The subdivision of each context into a set of distinct ones would increase the number of the top ones while reducing the overall structural depth. In

any iteration, the nesting depth never exceeded the second degree. In this way, we managed to simplify the system design while avoiding any over fragmentation.
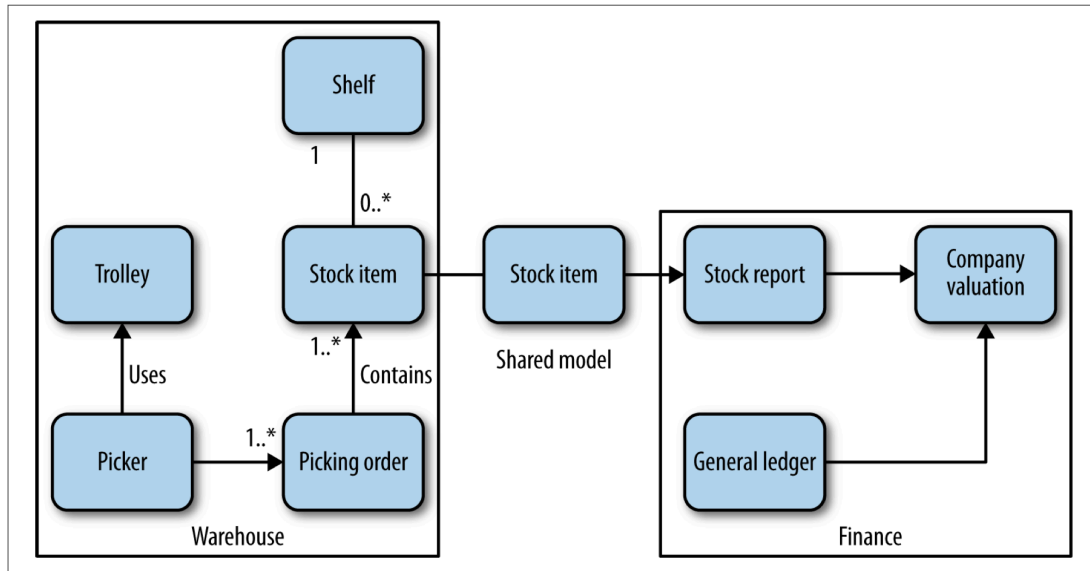
Figure 3.3: Sharing models across domains – an illustrative example (origin: web)

Lose coupling & High Cohension

The concept of bounded contexts is essential for MSA designs to empower loose coupling among microservices. "Loosely coupled" services mean that any change to one service does not involve any change to another for the system to conserve its seamless operation. That is an important aspect, and a core idea of the MSA approach, that we kept in our mind while designing the introduced system: to avoid any shortages originating from "tight coupled" designs. Being able to make changes to any arbitrary set of services and to deploy them at once with no need of changing any other already deployed service enhances system agility in terms of scalability and extensibility. Any cloud-hosted system affording such features is considered antifragile and time-insensitive. Any future update add on, or issue may be provided in a tight scope; operators should not suffer any lack of the broad domain knowledge. The system would be able to scale out or in, based on the needs of the business logic on top of the typical bottleneck thresholding technique. Loosely coupling services though, is not an easy way to go even while following the MSA paradigm and strictly obeying the concept of the bounded context. A common pitfall is to select an integration style that tightly binds one service to another, causing changes inside the service to require a change to consumers.

A loosely coupled service should afford a limited knowledge about the services with which it cooperates. That indicates we presumably want to deteriorate the inter-service communication, because, beyond the apparent performance degradation, chatty communication often implies tight coupled services.

Usually, that kind of flows are shaped by the structure of the teams that implement the related system. Planning and grooming sessions ³are established to define the API provisioning of each team and the way that the rest of the teams are going to consume them. In our case, though, the primary technical guide to that design decision was to chunk up our architecture to simplify testing, in regard to modules consuming modules that belongs to another context, To explain that further let us consider two bounded contexts, the User Management (UM) and the gate-keeping (user access) one. Each context should be bound to a particular service to foster the included domain business logic. The first one should handle socializing operations along with authorization requests for the data access given that the later one (user access) has already managed the authentication of the client successfully (i.e. registration, login, oauth, SAML login). In case of a non-MSA design, any test related to the UM modules may depend on a "spy-mock"⁴ of the a subset of the modules laying into the user access domain. This results to all failing scenarios related to user access to be part of the tests of the UM modules as well. That cannot scale when developing systems of multiple contexts, neither in terms of implementation nor in terms of maintainability. By separating the concerns of the two, we provide two completely decouples services, to unit-test both of them in the subdomain they foster the only prerequisite is just mocking out the communication end-points of service A when testing service B and vice versa. While any communication testing is performed through a lean integration or A to B testing;

(a) Unittesting implicitly invokes collaborators services, sociable testing

(b) Unittesting explicitly mocks collaborators services, solitary testing

(c) Unittesting services mocks endpoint reactions, highly isolated testing
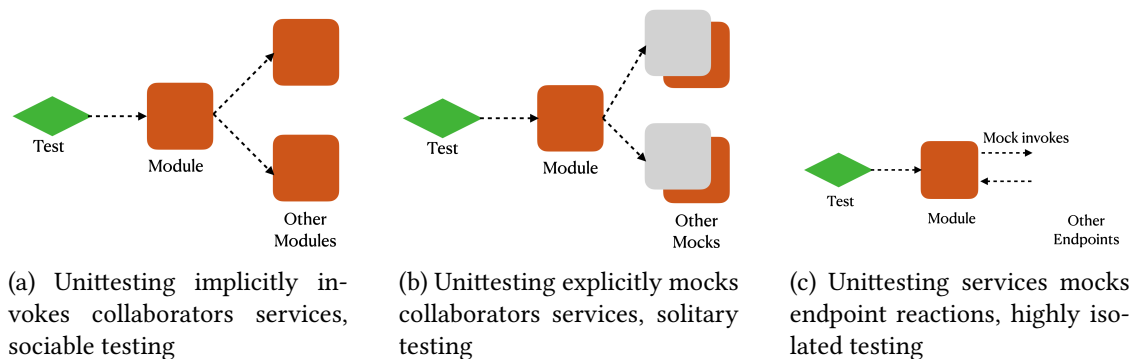
Figure 3.4: Design composite testing operations

What we did simply thought is to find a way to provide a robust testing coverage while writing the most straightforward test cases we could. By tackling the test simplification problem as an optimization one, we ended up with an increased unit of isolation when considering larger-scoped tests of entire contexts. Consider, for example, end-to-end testing against a module that belongs to context A. In that case(figure 3.4a), all collaborating modules might be stubbed out by just mocking their interface (figure 3.4b). In our case though, we have uplifted that paradigm to stub out only the communication endpoints (figure 3.4c).

---

³In the typical agile software development lifecycle planning, grooming sessions are in-person or virtual meeting focused on defining the scope of a subsystem implementation details, the high-level designs and the requirements of a subsystem,

⁴A spy-mock or just spy is a special software enginnering pattern that is made to check the way that the tested code utilizes some particular object which is he one that is mocked.

Fine-grain scaling & polyglot persistence

Figure 3.5 summarises the key takeaway benefits of the microservice paradigm in an intuitive yet illustrative visualisation, fine-grained scaling and polyglot persistence.

(a) Polyglot persistence

(b) Coarse vs fine-grained scaling, monoliths vs microservices scaling
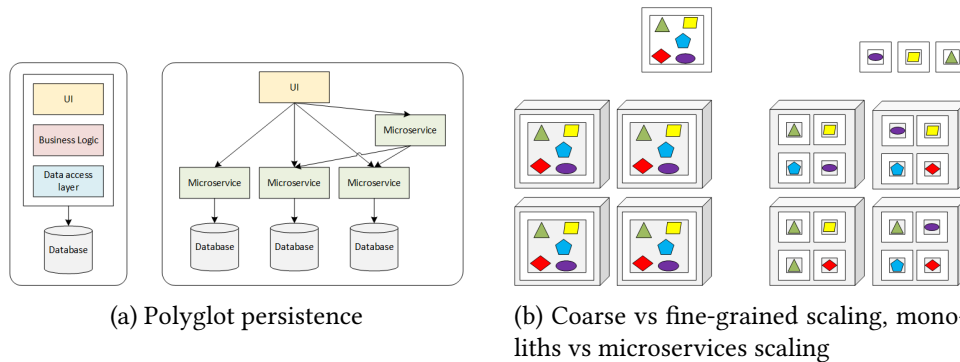
Figure 3.5: Microservarization take aways

On top of decentralizing decisions about conceptual models, microservices also **decentralise data storage** (Figure 3.2b). At the most abstract level, it means that the conceptual model of the world will differ between systems or even subsystems of the same system. Microservices prefer letting each service manage its own database —either different instances of the same database technology or entirely different database systems, an approach called **polyglot persistence** (see Figure 3.5a).

Besides, just-in-time scalability is facilitated [DLL+17], since there is no more need for scaling the entire application but only the services that require more resources. Microservices can scale-out effectively in contrast to the monolithic ones that can scale only up. Let us consider a typical web application built in a monolithic way. In that case, scaling the monolith out involves running multiple instances behind a load-balancer

Now consider our monolithic app affording two kinds of code routines, memory-intensive and a CPU intensive ones. We still can scale the app out by increasing the number of running instances, though it is evident that **the resource allocation is not going to be even close to the optimal!** That usually affects the efficiency of any cloud infrastructure seriously and ending up in blown running and maintenance costs. By following the microservices paradigm, though, the monolith of our example can break apart into two segments, the memory intensive and the CPU intensive one, accordingly. These segments may not even have the same world perspective (polyglot persistence), but they would be able to scale-out efficiently by allocating the required infrastructure resources in a decoupled way.

3.1.3   White vs black box design

Splitting the monolith's components into services or designing an MSA system from scratch, eliminates any programming language, software engineering paradigm or database system obligation. Examples of large-scale applications which adopt microservices include Netflix [HAB17], Amazon [Wag], Uber [Rei] and Etsy [Har]. Netflix for example receives around one billion streaming requests every day [Hus16], thereafter routing each request through an API to multiple

back-end microservices [HAB17]. MSA aspires to enhance the flexibility of large scale distributed applications to make them better cope with the operation, maintenance and evolution uncertainties. Such flexibility promises improvement to the maintenance costs and quality of service (QoS) provisioning to the end users. The developer can tackle any sub-challenge utilizing the tools and technologies that fit best — MSA is the adoption of divide and conquer principle in the field of software design. Such an heterogeneity contributes further to the overall system performance but not always at zero cost. Remote calls are more expensive than in-process calls, and thus remote APIs need to be more coarse-grained, which is often awkward to use. If someone needs to change the allocation of responsibilities between components, such shifts of behaviour are harder to operate when crossing process boundaries. Thus **determining the optimal microservice boundaries (i.e. microservice granularity) is among the vital design decisions that influence the Quality of Service (QoS) of the application at runtime**. In this perspective, this work adopts an architecture-centric approach to tackle the following decision problem (D):

> "*D: Given a set of x required functionalities, what is the optimal microservice landscape to provide them considering no infrastructure constraints (i.e. any resource type is available for optimal usage)?*"

There are two good examples of schools trying to tackle that issue, the black and the white box one. The black box approach is an agile operation of continuous development and monitoring of an already deployed platform. Predefined bottleneck tracks of a live platform highlight the need for more fine-grain design. Thus the platform is revisited and a new version is deployed to be monitored to a perpetual extent. That approach, however, often suffers two major drawbacks. The first pitfall is that during the development of a microservice-based system the reasoning about microservice granularity takes place at the code level instead of the design level. That ends up with systems featuring high complexity while missing heterogeneity advantages or decentralised governance. Although intuitively the main trigger for MSA adoption is "people finding they have a monolith that's too big to modify and deploy" [Fow15], unfortunately, to the best of our knowledge there is no sufficient problem formulation of "too big" or "too small" notions. There is even no technique for creating an MSA based system from scratch. That is an open design problem where only good practices do contribute so far.

Furthermore, the definition of potential bottleneck points requires domain-related knowledge or personal sense of contributing individuals. In that case, any architect may miss the big picture of what is to be built or even be misled by hidden correlations to wrong and costly conclusions. In this setting, the black-box approach may suit best in cases of migrating a monolithic architecture into a microservices one rather than building a microservices-based system from the ground up.

On the opposite side, the white-box figures that approach the concerns of a microservice are separated into several layers (Figure 3.6).The hexagonal architecture consists out of the domain model, application services and adapters with ports. Each side of the hexagon stands for a particular port, although, in practice, there could be more than six distinct ports. These ports are the way through which the microservice upstreams or downstreams information to the clients or to other microservices. To employ a microservice, any client has to route its requests through one of the exposed ports and its corresponding adapter. The adapters work like an anti-corruption layer; the core domain model that the microservice is built around would never be affected by

the exposed functionality. According to the dependency inversion principle (DIP) [Mar02] high-level modules should not depend on low-level modules. Both should depend on abstractions (e.g. interfaces, adapters). Abstractions should not depend on details, but the other way around; details (concrete implementations) should depend on abstractions. In this context, the domain model, which represents the business logic, is acknowledged to be the cornerstone of the particular design approach. It is independent of the surrounding application, services and adapters. The layer dependencies are applied from the outside to inside according to DIP (Figure 3.6).
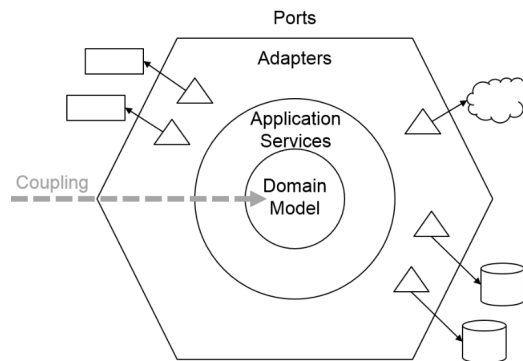
Figure 3.6: The hexagonal architecture pattern

3.2   Microserverization, determining the microservice boundaries

In our case, that we need to build a system from scratch, we are going to follow the white-box approach (see section 3.1.3) of modeling the back-end substructure to guaranty a sufficient granularity level in terms of distributed governance and scalability.

Software architecture can be divided into macro- and micro-architecture; it can further be divided into organization, system and building block level. The organization and system levels belong to macro-architecture whereas the building block level can be assigned either to the macro or micro-architecture perspectives, depending on what is required for the concrete architecture [VACK11]. Despite their names, microservice architecture and the domain model describe the macro-architecture of a system. A microservice is a functional or technical software building block that requires a platform to run on. Neither DDD nor microservices limit the underlying platform. When using DDD, microservices are structured according to the organizational units using bounded contexts originating from the domain model [New15b][Ver13]. The domain objects within a bounded context specify the core architecture of a microservice.

Domain-driven design requires a layered architecture to separate the domain from other concerns [Eva04]. A typical approach to that is a five-layered architecture, consisting of the user interface, application, domain, and infrastructure layers. Figure 3.7b presents the distribution of these layers among the software building blocks of microservice-based applications Traditional design approaches promote the separation of the concerns into a vertical fashion – tiers (Figure 3.7a). *Data-tier* groups different data sources such as relational databases, XML databases, Microsoft (MS) excel spreadsheets, object databases and others. *Integration-tier* manages the

(a) Multitier design approach (origin: web)

(b) Software building blocks layered taxonomy (MSA)
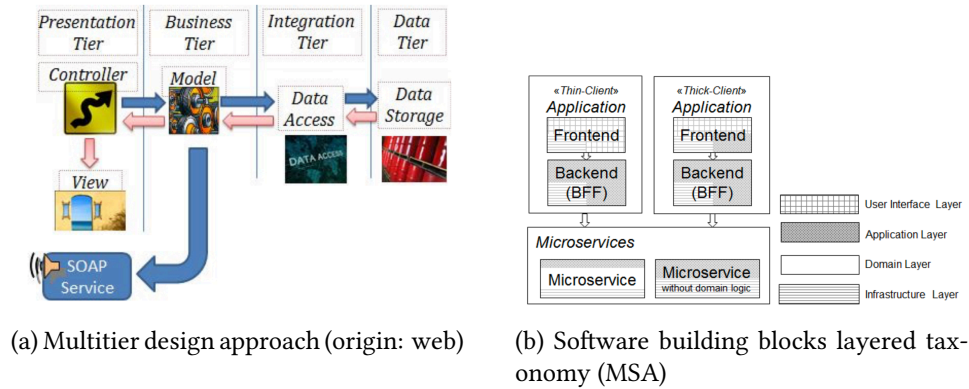
Figure 3.7: Divide and conquer design

connections to data sources (connection classes), and the execution of queries against each connection. It is decoupled from the rest of the tiers. No business logic should be considered part of this tier. Program code has to be restricted to data access operations and only. Next, *business-tier* consists of all classes that carry out business logic. Business logic could be bespoke for the application, for instance, special requirements or dealing with legacy applications and may include standard libraries from a third party if available. This layer lacks any database utility. *Presentation-tier* layer encapsulates all of the presentation applications. It can be extended to include web forms, mobile applications, desktop applications etc. Finally, the optional *web-services* tier is responsible for any required business layer classes that has to be exposed directly as a as web service. In that respect a web service becomes another type of presentation tier but without a user interface and merely a data service and/or business processing service.

On another abstraction level, microservice-based applications can be divided into applications and microservices. The application consists of a **frontend**, which is either *thick* or *thin* (meaning that it may contain application logic or not), and its backend, which provides the application logic. The backend uses the microservices to access the domain layer or general infrastructure functionality. Each microservice has an application layer on top. The application layer translates requests into either the domain or infrastructure layers. Infrastructure logic may be part of each software building block. In our approach, we applied the layer distribution following Miller's approach [MT15] according ro which a layered architecture, higher layers can communicate with lower layers. Figure 3.8 depicts the layered architecture's communication paths applied to the above-mentioned software building blocks [MT15]. The frontend should not directly call the microservices; we emphasize this by using dashed arrows.

3.2.1 The onion architecture

The layering from Figure 3.7b shows that the layering introduced by DDD is also applied to the microservice architecture. Going into the implementation of a microservice, the sequence of the layers changes. The underlying structure of the microservices is defined through the hexagonal architecture pattern introduced by the authors in [Coc07]. The work in [HGS+17] extends that paradigm to model precisely the fundamentals of a microservice. The onion paradigm (Figure 3.9) is set to model the increased communication among software building blocks while maintaining
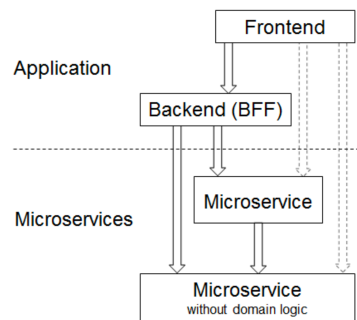
Figure 3.8: Communication among MSA building blocks

minimizing any adoptions. On the outer layer of the onion architecture, we find the infrastructure as well as the exposed interface as fine-grained building blocks. The infrastructure part of the layer provides technical functionality for the operation of the service such as database access. The exposed user interfaces focuses on the provisioning of the underlying business domain that can be used by clients for interaction. Similarly to the original hexagonal architecture. the adatper pattern is once again the gatekeeper of the domain model at the core of teh service to provide a "corruptiontight" layer. . Despite that the bibliography suggests that the hexagonal architecture and the onion architecture are the same [Ver13], the onion architecture adds, or focuses on, a new layer, the domain services one. Domain services represent behavior that cannot be mapped directly to a single domain object [Eva04][vernon2013implementing]. For instance, this is the case if the behavior is spread over multiple domain objects to form a business workflow. This way, the domain services are tightly coupled with the domain objects; together they build the whole domain. The work in [HGS$^+$17] highlights the interface definition of the repository on this layer. Nevertheless, that approach mixes the domain-specific layer with technical aspects. That is the essence of the "glue layer" of the onion approach (Figure 3.9), to perform as a link layer between the application and the domain. That layer consists of all the known repositories or factories originating from the DDD while the domain models represent the core of the onion architecture.

The layered architecture is applied to the whole application and divides it into horizontally-divided layers. Meanwhile, the onion architecture is applied to the microservice building blocks from Figure 3.7b and divides them in a vertical manner. Going back to DDD, a microservice is defined through a single bounded context in the customer's domain [New15b]. Considering the microservice architecture as an onion makes it more suitable with the concept of bounded contexts. Every microservice has to work autonomously, which is intended with the layering layout of the onion architecture.

3.2.2   Microserverization, the battle plan

Next, we are introducing the workflow (battle-plan) involved in building microservice-based the system that we introduce in our work. The goal of the battle plan is to provide a group of operations to develop a finite set of applications within similar domains – e.g. an application that offers information on points of interest, an application navigating from and to points of
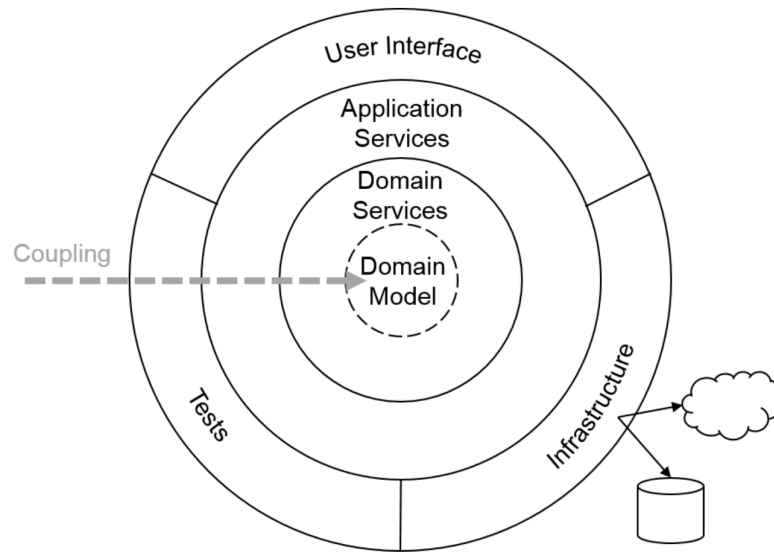
Figure 3.9: The onion architectural approach

interest and an application that enables the management of points of interest. Our workflow aligns with the typical software development activity flow [BD09]. However, we extend that framework to include any artifacts that are associated with the client application, the deployment and maintenance presented in later chapters.
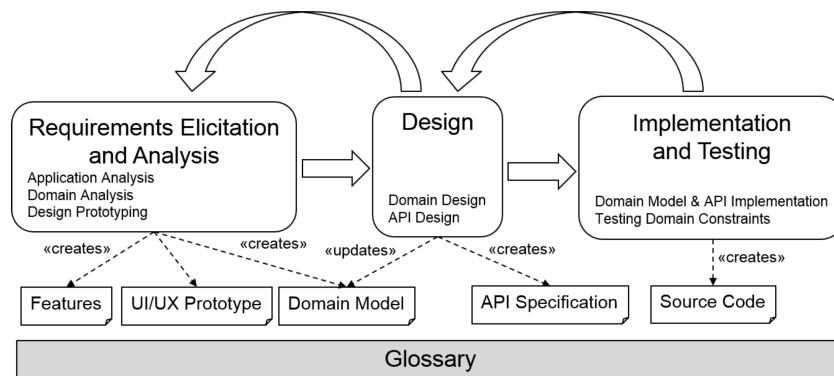
Figure 3.10: Development evolution cycle

Requirements Elicination: The requirements elicitation and analysis consist of three steps that connect via a feedback loop: first, the application requirements are stated in features with and from users by following the Behaviour-Driven Development (BDD) approach; next, the information model (a primitive domain model) – the input of domain experts (in our case medical

personnel) is essential during that step; last but not least, a prototype is designed to be asserted by the related stakeholders or even a subset of the domain experts. All steps are closely associated (feedback loop) and thus combined into a single activity; reviewing prototypes increases the knowledge of the domain that engineers afford and may trigger the change of terms or workflows. Discovering the information model reflects on eliciting requirements operation to sharpen the goal of the application itself.

Design: The design step implies the collection of activities incorporated in designing the domain and the APIs of the microservices. The initially shaped information model is refined based on the UI/UX design and further inquiries with the domain experts (i.e. medical personnel in our case). In this sense, by applying DDD patterns, the information model is going to be transformed into a fully-fledged domain model. According to Conway's Law [6] the system consists of distinct subsystems that can be realized by individual teams or in distinct time frames using bounded contexts. Domain design results in a domain model that must be bound to the implementation artefacts. In this context, domain design aligns with the system design activity. Any microservice manages access to the domain model by transposing information of the application layer to the domain layer. The UI/UX, along with the formated domain model, shape the exposed APIs; the user interface layer and the application layer should eventually consume domain model capabilities though the microservice gateways. This activity can be assigned to the object design activity discussed in [13].

Implementation: The next step is to implement and test the microservices designed during the first step. The web APIs describe the microservices' entry points and in conjuction with any encapsulated business logic constitute the constraints of the defined domain model such as multiplicities and directed associations. Any DDD attempt suffers the ubiquitous language; therefore, we introduced a glossary to capture the domain terms. Each term of the ubiquitous language is listed and described by a few sentences. We see the glossary as a cross-sectional artefact. It is created and updated in each activity of Figure 3.10.

The work in [5] clarifies that any "deep model" requires "explorations and experimentation". There should be no single best-fitting solution. To mitigate that issue, we connected all the steps of the battle-plan with the corresponding feedback loops. By iterating over that flow for a couple of times, we gained insights into the domain across the whole software development activities. The hard part here is to be open-minded and ready to revisit decisions made and thus update any artefacts produced by the previous iteration. In this setting, we were able to form an agile technique to optimize the development by extracting hidden knowledge and incrementally mitigating misunderstandings.

3.3   Battle plan application

3.3.1   Elicintating user requirements

The goal of any human-designed system is to satisfy the needs of a portion of people (users), into a particular context (domain), through the seamless operation of it. In this scope, the first step of our battle plan is to understand the needs of the potential system users. To eliminate any
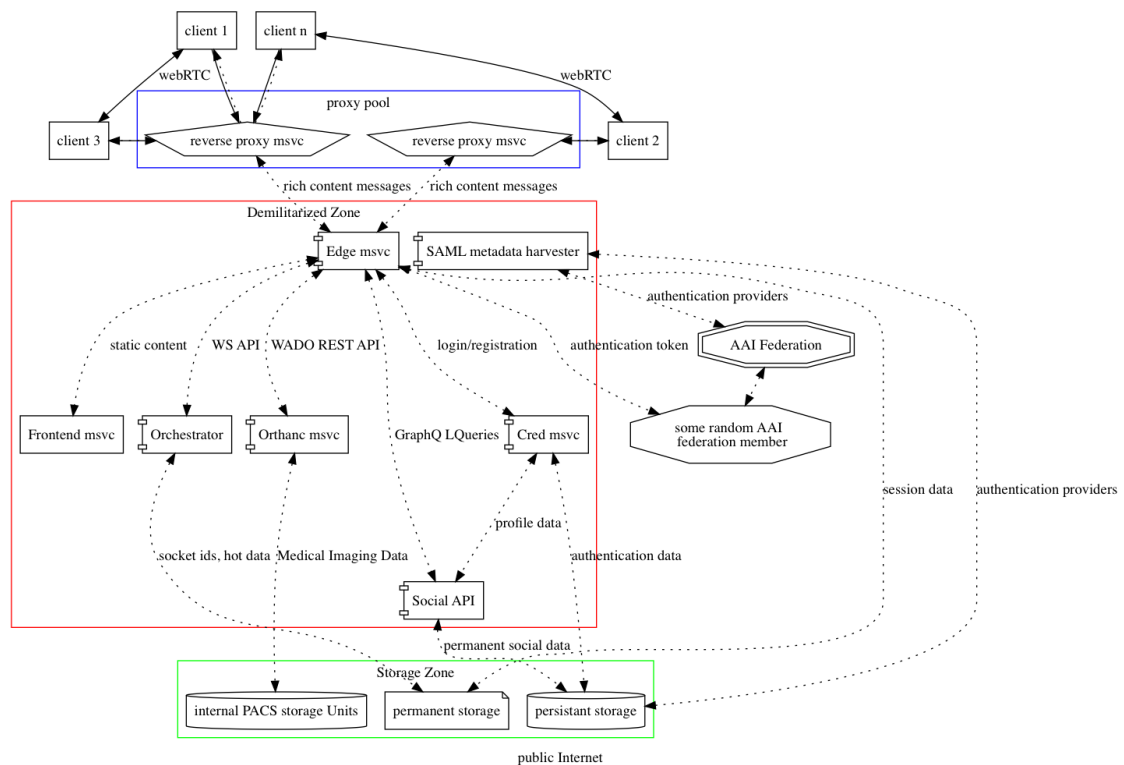
Figure 3.11: System architecture in details

uncertainties, we follow the agile methodology of the iterative cycles of continues inspection and adaptation; (1) we analyze the application of the system, (2) explore the related domain while (3) designing a prototype. All three pursuits are highly correlated and impact each other; terms from the application features should get aligned to the domain model and implemented to the prototype. The other way around, any inspection of the prototype may provide novel insights that may alter the application features and/or the domain model itself.

3.4 The architectural landscape

Figure 3.11 depicts the plain system architecture following the MSA paradigm. The backend features a **network of cooperating autonomous services** (see also Table 3.1). The whole network resides into a Demilitarized Zone (DMZ) —all services communicate through a private network and they are not directly accessible over public Internet— in a two-tier fashion architecture. The first layer consists of all the services that constitute the provided API, while the second layer comprises "rich" database management systems, which autonomous services providing database and storing functionalities through REST APIs.

EDGE SVC is the front gate for all the requests to the backend of the introduced system. It is built to provide dynamic routing, resilience and security. Featuring a rich and divergent API, it interacts with the corresponding services to fulfill any incoming request. Moreover, it assigns a unique ID to every autonomous incoming request or to any group of requests that constitute a

more complex operation. This unique identification of requests is vital for monitoring due to the distributed fashion of MSA. Last but not least, the core EDGE SVC manages cookies, sessions as well as interacts with third-party AAI (Authentication and Authorization Infrastructure) to offer a flexible and robust AAA (Authentication, Authorization and Accounting) scheme.

FRONT-END SVC delivers content to the user such as the client application as well as any supplementary static assets (i.e. HTML pages, CSS files, etc.) while **CRED SVC** is responsible for the management of the registered users's credentials. Registration, notification e-mails or any further update of registration data is been provided via that service.

SOCIAL API features the social networking infrastructure based on GraphQL. GraphQL is a novel query language introduced by Facebook for managing data efficiently in a transparent fashion. The service implements a runtime for executing queries and mutations by using a well-defined type system for data. GraphQL isn't tied to any specific database or storage engine, queries and mutations accompanied by use regulations are defined and can be applied to any database system.

WEBSOCKETS API is the core of every real-time feature of the provided platform. The service hosts a WebSockets connection coming from each online user. In this manner, the backend can asynchronously interact with any client, in order to facilitate status discovery, notifications, instant messaging and signaling, they are vital for driving remote conferencing sessions. The service depends on a Redis key value store to store any temporal data before archiving them to the corresponding persistent data stores through SOCIAL API.

Metadata Harvester is dedicated to aggregate all the necessary SAML (Security Assertion Markup Language) metadata [Lew09] coming from the third-party AAI federation. In this way the system is aware of the members of the federation and thus users are able to be authenticated via an XML-based SSO (Single Sign-On) scheme.

MONGODB-Store hosts both registration and social databases while a **PostgreSQL DB** powering an Orthanc instance [JBD$^+$], a free and open-source DICOM server whose source code is available to the benefit of hospitals and researchers, and they are combined to mimic a fully functional PACS system. Decentralizing responsibility for data across microservices leads to more flexible and performant system but it introduces implications for managing updates.

3.4.1 The API specifications, GraphQL

The management and processing of complex data structures along with the administration of data-flows are two of the cornerstones of the design of any data-intensive system. Especially "micoreserverized" systems require high-level of interoperability to allow the collaboration and independence of both internal and external to the overall system components. While typical REST APIs is the standard solution to achieve interoperability, sometimes they lack the required flexibility and performance. Facebook INC. introduced GraphQL a couple of years ago, to trigger the revolution of GraphQL APIs. GraphQL is a flexible, fast and stable protocol for data management operations. That makes it an exciting approach for data-intensive and complex data-driven systems [VICBGPn17].

GraphQL stands for Graph Query Language. Despite there is no formal definition of its semantics, the literature suggests that it affords a very low complexity for evaluation. More

specifically, the combined complexity of the primary decision problems is in NL (Nondeterministic Logarithmic Space); consequently, they can be solved in polynomial time and are highly parallelizable [HP17].

In this context we use GraphQL to blur and abstract away the bounded contexts and domain models. Any domain model related operation should be part of the microservice −> ..... graphql is there to translate any client request to the domain field.

GraphQL over REST

REST is the widely-used acronym for REpresentational State Transfer that was first introduced back in 2000. It is based on a set of guiding principles to offer a lean and transparent way of client-server communication (or service integration). First, it separates the user interface concerns from the ones of data storage, to increase UI portability across multiple platforms and advance scalability by simplifying the server components. On top of that, the layered system design vertical isolation. Requests are fulfilled by chained components one operating on top of the other, and each component is not aware of anything outside the intermediate layer it operates in.

The Stateless aspect of REST obeys that any request to the server must contain all of the information necessary to interpret the request. With the client maintaining any session data, the server lacks any request context. Data within a response to a request is implicitly or explicitly labelled as cacheable or non-cacheable. Any cached response is there for the client to use on any fothcoming equivalent requests. REST suggests the application of "generality" to the component interface, a software engineering principle describing software that is not limited to one particular case. In this setting, the overall system architecture is simplified to provide transparency in its interactions. The ultimate goal, though, is uniformity; REST signifies a uniform interface which, unfortunately, implies multiple architectural constraints – i.e. identification of resources [5]; manipulation of resources through representations; self-descriptive messages; and, hypermedia as the engine of application state.

Although its broad adoption, REST does suffer some considerable drawbacks that may raise severe limitations to any modern system that manages a wide range of information models (resources). In particular:

- **Query Complexity**: REST requires multiple and complex HTTP requests to fetch multiple resources

- **Overfetching**: Overfetching is characterized by returning more data than required by an application.

- **Under-fetching** and n+1 request problem: Under-fetching indicates that a particular endpoint does not give sufficient information. This results in making an additional request by a client application to a server. This is referred to as n+1 request problem.

- **API versioning**: An API creates a contract between two systems for information exchange and hence it should be stable and consistent. However, the business goals of any

---

[5]Key abstraction of REST: any piece of autonomously defined information such as a document, a photo, a calendar and other

organizations change, so the APIs must be adaptable for modifications according to their behavior. This is handled by API versioning. An empirical study [12] addresses the perennial issue of REST API versioning and how evolution of such API affects the clients.

These are the issues GraphQL was build to tackle. GraphQL operates as a thin service abstraction layer [UKT+19] to provide a single API endpoint for resource acquisition, creation or modification. GraphQL APIs feature (1) strongly typed schema (2) introspection [6], (3) rapid product development (4) rich open-source ecosystem (5) composite API [7], (6) faster request-response cycles than the typical REST implementations and (7) hierarchal structure – i.e. GraphQL queries are hierarchical, and structurally reflect the data it returns [TODO: ref a figure here]. It is noteworthy that the introspection of the GraphQL endpoints facilitates the comprehension of the interface with ease by the developers while machine-readable representation empowers the dynamic and loose coupling between the server and clients. In this setting GraphQL leverages the creation of flexible maintainable, interoperable and secure APIs [11], aspects that should be considered essential in any medical data management system.

3.4.2   Inter-service communication & security

Concerning the building of communication structures between different processes, the microservice community favors the approach of **smart endpoints and dumb pipes**. Applications built from microservices aim to be as decoupled and as cohesive as possible —they own their own domain logic and act more as filters in the classical Unix sense— receiving a request, applying logic as appropriate and producing a response. These are usually choreographed using simple REST protocols rather than complex protocols, such as WS-Choreography or BPEL or orchestration by some central tool. The two protocols used most commonly are HTTP request-response with resource APIs and lightweight messaging [SQV+14].

The second approach in common use is messaging over a lightweight message bus. The infrastructure chosen is typically dumb (dumb as in acts as a message router only) —simple implementations such as RabbitMQ or ZeroMQ do not do much more than provide a reliable asynchronous fabric– the smarts still live in the endpoints that are producing and consuming messages, i.e. the services. In a monolith, the components are executing in-process and communication between them is via either method invocation or function call. The biggest issue in changing a monolith into microservices lies in changing the communication pattern. A naive conversion from in-memory method calls to RPC leads to chatty communications which do not perform well. Instead you need to replace the fine-grained communication with a more coarse-grained approach.

In the proposed system, inter-service communication is based on the REST approach which offers many benefits, such as development simplicity, rapid prototyping, flexibility in scaling and independence from implementation language and underline platform. Nevertheless, as someone might expect, REST/HTTP is not a silver bullet; there are scenarios that it simply does not fit in. Such a case is the intrinsic asynchronicity regarding instances of the same service: two users are connected to two different instances of the WebSockets service and one of them is trying to send a message to the other one. In that case, every instance is not aware of the users connected

---

[6]clients can query against fields types or available queries themselves

[7]schma federation allows merging multiple GraphQL APIS [HP17]

to the others, so there is not an obvious way to relay the messages through some REST API call. The solution to this problem is supplied by the use of the aforementioned message brokers/buses that implement well-known and useful patterns for these purposes, like Pub-Sub. According to Publish/Subscribe messaging paradigm, senders (publishers) are not programmed to send their messages to specific receivers (subscribers). Rather, published messages are characterized into channels, without knowledge of what (if any) subscribers there may be. Subscribers express interest in one or more channels, and only receive messages that are of interest, without knowledge of what (if any) publishers there are. This decoupling of publishers and subscribers can allow for greater scalability and a more dynamic network topology. In our case, Redis provides such kind of functionality.

Secure endpoints, certified requests

Most cloud-based systems are trying to provide the extra level of security by invoking private endpoints, strong passwords and multiple levels of encryption. The introduced system, however, facelifts the concept of the "demilitarized zone" (DMZ); it follows an entirely different approach to boost safety, it promotes trust! DMZ is not a centralized "private" set of resources but a distributed cluster of *veryfied* ones.

In the field of computer security, a demilitarized zone is a "low level" isolation practice. Mostly it is implemented as a subnetwork, either physical or logical, that contains and exposes the external-facing services of an organization to the hostile world of the world wide web. A DMZ can have a completely different IP addressing scheme from the internal network, or it can reside on the same network broken up logically through the use of VLANs. Originally that practice meant to add a layer of security to an organization's local area network (LAN); an external network node could only access what is exposed in the DMZ, while the rest of the organization's network is isolated behind network firewalls. By having this design feature utilized in a network, an organization can make selected services available for public access such as hypertext transfer protocol (HTTP-web servers) and file transfer protocol (FTP) or even domain name system (DNS) while still protecting the computer assets on the internal land area network (LAN) [WWF13].

Cloud computing extends that practice beyond that perspective; almost every cloud-based application bears a DMZ to protect any sensiteive piece of information. Usually, whenever a product goes worldwide, multiple DMZs are set to host distributed storage or other sensitive services. Even though that modern cloud providers promote managed operations [8] to abstract away low-level details, DMZs are complex networking setups that may lead to the extensibility degradation of the deployed system or introduce certain boundaries such as cloud provider opacity.

Nowadays, any service that is being provided to the end user publicly is usually set in a DMZ. The servers that directly interact with any databases should not be publicly accessible, Any compromise of such a server may lead to data leaks. In our case medical data, the intellectual property of specialists and private messaging are considered to be private and sensitive data. In that sense we should and have guarded all "sensitive" services inside teh boundaries of a DNZ, We updated that design though to abstract the DMZ concept up into the application layer of hhe

---

[8]Managed operations: DevOps automation. A set of services to automate lower deployment details such as networking and storage

standard OSI specification, by utilizing the idea of certified distributed requests, In this setting the overall system design is not bounded by a centralized DNZ pattern anymore.

Certificates and authorities   Typically the everyday user browsing looks for two things that ensures a safe and trusted website:

- a padlock icon in the left-hand area of the address bar on the browser

- **https://** preceding the website address

By clicking the padlock, the user may access the website certificate that he has visited to check that it is valid or not. In particular, a modern web browser considers a certificate to be valid if a trusted Certificate Authority (CA) issues it. A CA is an entity that is trusted by everybody [9], and it is responsible for three major tasks:

- *Issues certificates*: Every user and every computing resource such as a web site or an SSH server must have an identity and a way to prove its validity. A certificate is an electronic version of a "driver's license", it is just a file which contains the information about the identity of its owner. A certificate authority issues such certificates.

- *Confirms the identity of the certificate owner*: A certificate authority guarantees that the identity information embedded in a certificate is valid.

- *Provides proof that the certificate is valid*: A certificate authority can also guarantee that a certificate is valid. A certificate can become invalid due to expiration, revokation, tempering. A certificate authority can provide proof that a given certificate is valid. What this boils down to is "we, the certificate authority, vouch for the owner of the certificate and ensure that they are whom they say they are because we have validated the information they claim".

Figure 3.12 illustrates the way a CA issues any certificate. First, the requester makes a private key and public key pair and submits an "application" called a certificate signing request (CSR) to a trusted certificate authority. **The CSR has all the information about the requester that will be displayed on the resulting certificate.** Next, the CA should verify whether the information on the CSR is true or not. In the first case, the CA issues and **signs** a certificate using its private key to deliver that back to the requester ready to use. In any other case, the certificate request is aborted. The verification of the submitted information invokes multiple steps such as challenges to the requester to prove that the domain being requested is controller by the requester, the requestor must sign a nonce (an arbitrary) number with its private key pair to prove control of the key pair. Once verification succeeds the requestor is may request, renew or revoke certificates for the registered domain.

---

[9]CAs to be trusted are registered in a public list

Figure 3.12: Issuing Certificates (origin: web)

JSON web tokens JSON Web Token (JWT) is an open standard (RFC 7519) that defines a compact and self-contained way for securely transmitting information between parties as a JSON object. This information can be verified and trusted because it is digitally signed. JWTs can be signed using a secret (with the HMAC algorithm) or a public/private key pair using RSA or ECDSA.

JWTs can also be encrypted to provide secrecy between parties, but this is not always a mandatory need. The most exciting part of a JWT is the signature one. Signed tokens can verify the integrity of the claims contained within it, while encrypted tokens hide those claims from other parties. When tokens are signed using public/private key pairs, the signature also certifies that only the party holding the private key is the one that signed it.

JWT was introduced to initially assist the authorization of any single user interacting with a federation of services. Once the user is logged in a JWT will accompany any later request of his allowing the user to access routes, services, and resources that are permitted with that token. Single Sign-On is a feature that widely uses JWT nowadays, because of its small overhead and its ability to be easily used across different domains.

Another application of JWT is the trusted information exchange. JSON Web Tokens are a good way of securely transmitting information between parties. Because JWTs can be signed, using public/private key pairs and thus sender's identity could be verified with ease. Additionally, considering that the signature is calculated using both the header and the payload apart of any private keys, tampering can be once again detected with ease.

$$xxx.yyyyy.zzzzz \rightarrow \begin{cases} x : Header \\ y : Payload \\ z : Signature \end{cases}$$

Figure 3.13: JWT distilled

In its compact form, JSON Web Tokens consist of three parts separated by dots (.), the header, the payload and the signature like explained in 3.13. Getting a closer look at the JWT specifications, the header typically consists of two parts: the type of the token, which is JWT, and the signing algorithm being used, such as HMAC SHA256 or RSA. The token payload contains the claims, i.e, statements about an entity (typically, the user) and additional data, encoded in Bse64Url format. There are three types of claims:

- Registered claims: These are a set of predefined claims which are not mandatory but recommended, to provide a set of useful, interoperable claims. Some of them are: iss (issuer), exp (expiration time), sub (subject), aud (audience), and others.

- Public claims: These can be defined at will by those using JWTs. But to avoid collisions they should be defined in the IANA JSON Web Token Registry or be defined as a URI that contains a collision resistant namespace.

- *Private claims*: These are the custom claims created to share information between parties that agree on using them and are neither registered or public claims.

```
{
    "sub": "00302107724394",
    "name": "Stan the Man",
    "admin": true
}
```

Figure 3.14: JWT Paylod example

Figure 3.15 depicts the way that the JWT signature is created in the case of HMAC SHA256 algorithm. It a actually is the combination of the encoded header, the encoded payload and a secret. All the three of them are combined into a signature that derives from the algorithm specified into the header section.

The signature is used to verify the message wasn't changed along the way, and, in the case of tokens signed with a private key, it can also verify that the sender of the JWT is who claims it is. The output is three Base64-URL strings separated by dots (Figure 3.13) that can be easily passed in HTML and HTTP environments, while being more compact when compared to XML-based standards such as SAML.

In authentication, when the user successfully logs in using their credentials, a JSON Web Token will be returned. Since tokens are credentials, great care must be taken to prevent security

```
HMACSHA256(
    base64UrlEncode(header) + "." +
    base64UrlEncode(payload),
    secret)
```

Figure 3.15: JWT signature

issues. In general, tokens should be invalidated as soon as possible. Whenever the user wants to access a protected route or resource, the user agent should send the JWT, typically in the Authorization header using the Bearer schema.

3.4.3 Logging and Monitoring

Software solutions based on microservices architectural pattern involve increased inter-service communication, which should not be taken always for granted. As a result, algorithms and well-known workflows are altered to groups of asynchronous processes. Therefore block diagrams should be substituted by sequential diagrams embracing any potential communication failure.

In order to efficiently monitor the operation of the platform we ended up with the Elasticsearch, Logstash and Kibana (ELK) stack [BBG+15][AAB+17].

- Elasticsearch is a search server based on Apache Lucene; it provides a distributed, multi-tenant capable *full-text* search engine with a RESTful web interface and schema-free JSON documents.

- Logstash is an open source, server-side data processing pipeline that ingests data from a multitude of sources simultaneously, transforms it and then sends it to some "stash" (in our case Elasticsearch).

- Kibana is a rich visualization tool consuming the Elasticsearch API. In this fashion, every node of the constituted Kubernetes cluster has been configured to monitor any deployed containers as well as a bunch of individual process that are considered vital for the operation of the cluster.

Table 3.1: Architecture components and their short description

Component	Short description
EDGE SVC	Reverse proxying, dynamic routing, resilience and basic security
CRED SVC	User registration, authorization handling, token issuance
FRONTEND SVC	UI views and SPA serving
SOCIAL API	User and contact management
WEBSOCKETS API	Message passing, WebRTC signaling, real-time features support
Metadata Harvester	GRNET's AAI metadata aggregator
MONGODB-Store	Various instances as backing stores of the rest microservices
Orthanc	Open-source DICOM server with REST APIs [JBD+]
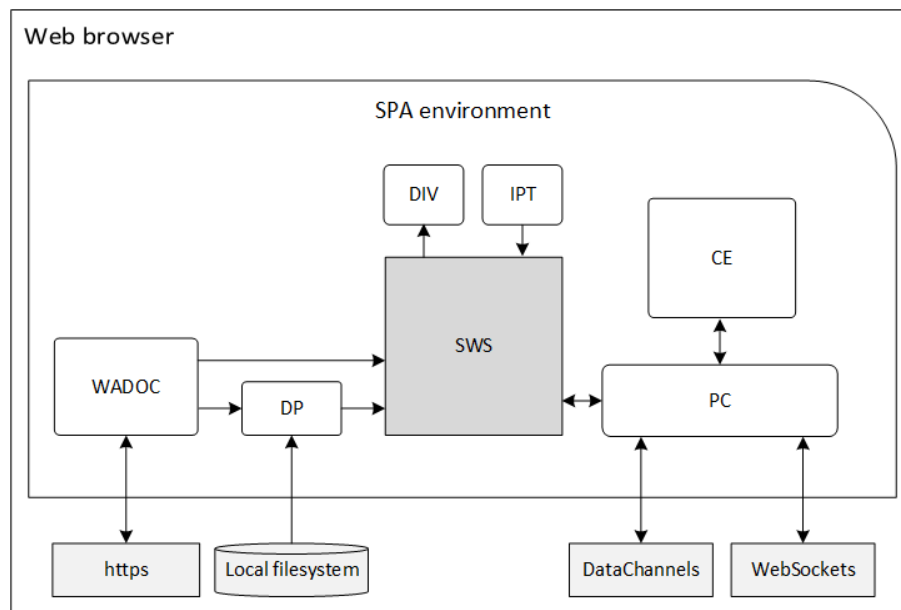
Figure 3.16: The internal architecture of SPA.
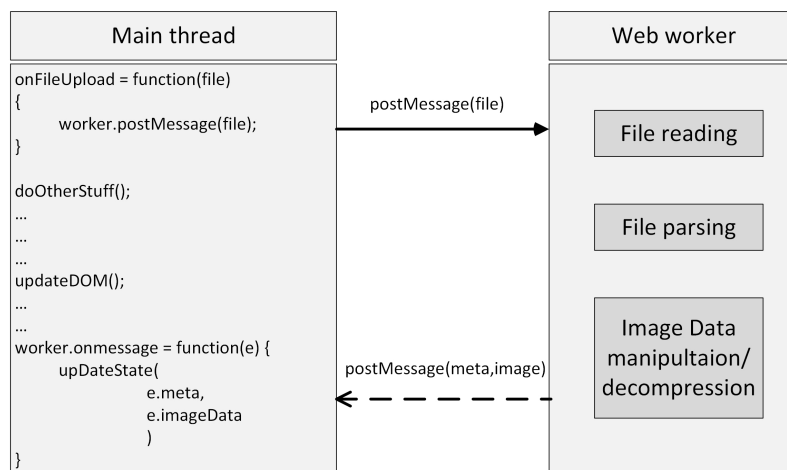
Figure 3.17: Parsing DICOM files concurrently.

3.5   The workspace, a Fat client

Single page application, SPA

The SPA provides a sound DICOM viewer. It supports reading and displaying a wide-range of single-frame and multi-frame DICOM files, encoded according to most common transfer syntaxes. Interaction with PACS via WADO (Web Access DICOM Objects) as well as local storage devices in conjunction with smart series integration (synthesis from distinct DICOM files) and browsing is also supported. Basic image processing (e.g. zooming and panning, rotation and
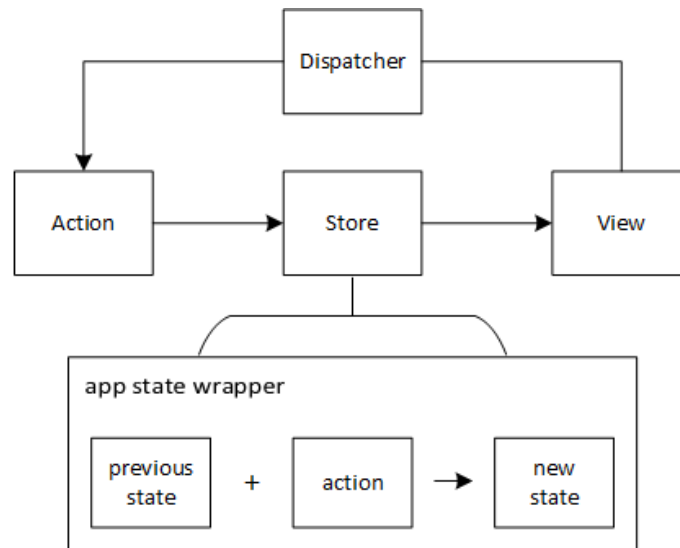
Figure 3.18: SPA state updates.

flips, brightness and contrast adjustment, preset window level setting for Computed Tomography and Magnetic Resonance modalities, etc.) filtering and ROI (Region Of Interest) selection are supported to enhance image viewing.

However, the **basic innovation** of the SPA itself is the real-time collaboration feature based on multi-channel video conferencing, workspace and file sharing, chat rooms and shared operations and annotations. While typical implementations of similar collaboration environments involve centralized architectures, we introduce a pure peer-to-peer interconnection scheme to eliminate any intermediate party overheads and handle scalability issues. The communication mechanism relies absolutely on the WebRTC protocol, which is a free, open source project that provides browsers with secure (encryption embedded) RTC (Real-Time Communication) capability. Implementation details are provided in the following paragraphs.

3.5.1   Client Application & Implementation Details

The introduced software consists of seven autonomous entities: WADO Connector (WADOC), DICOM Parser (DP), DICOM Image Viewer (DIV), Image Processing Toolkit (IPT), File Sharing Engine (FSE), Conferencing Engine (CE) and PeerConnector, as depicted in Figure 3.16. The whole SPA is implemented in HTML5, CSS3 and JavaScript. The User Interface (UI) has been built in a reactive manner [CC] exploiting the declarative approach and one-way data flow of **ReactJS**, an efficient, and flexible JavaScript framework for building user interfaces. By enforcing a purely functional programming style, the classic MVC (Model-View-Controller) paradigm is upgraded to support explicit modeling of complex time-dependent relationships in a high-level declarative way.

Views are the declarative React components which are responsible for the UI layout that is being presented to end users through composition, while **actions** are plain JS objects with a type property that indicates the type of action being performed. **Dispatcher** is a special component/function that produces the actions as instructed by the registered event handlers of views.

Finally, the **store** is essentially a wrapper around the whole application state, implemented by a single JS object; it ensures the latter's **immutability** by deterministically calculating and saving the new state based on the previous one and a single dispatched action. The SPA consists of views that are organized in a hierarchical way. High-order views get notified whenever a new state is produced; they consequently propagate parts of that state to their child views which update themselves accordingly in order to reflect the changes.

This structure allows us to **reason easily about our SPA** in a way that is reminiscent of functional reactive programming, or more specifically data-flow programming or flow-based programming, where data flows through the application in a single direction —there are no two-way bindings. The different parts of the application remain **highly decoupled**. The two-way data bindings originating from the heavy use of **Backbone.js** and its core dependency, jQuery, often resulted in cascading and subtle updates that were difficult to reason about. Now, updates due to user interaction can only change data within a single round and thus the system as a whole becomes more predictable.

Shared Workspace State (SWS) is the most important component of the SPA. It comprises only a part of the whole application state but it deserves special reference as it needs to be kept synchronized across client instances for all the duration of a remote collaboration session. In other words, as long as a call is being carried out, SWS essentially becomes shared state with eventual consistency semantics. SWS holds all the information from the loaded DICOM files —images and metadata, as well as all the user applied filters and annotations, in order for the connected clients to be able to render the same output every moment.
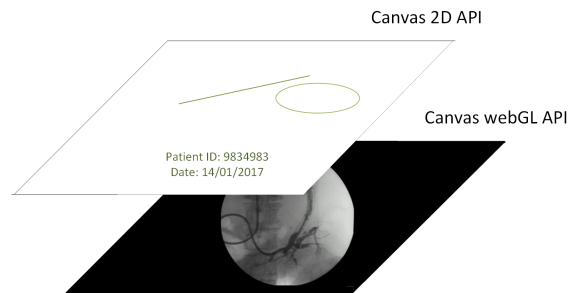
Figure 3.19: Architecture layout in practice.

WADO Connector (WADOC) is the logical part of the application which is responsible for requesting and receiving DICOM persistent objects residing in remote PACS through HTTPS, based on DICOM UIDs. It relies on the WADO standard, which is part of the DICOM specification. Data can be retrieved either in a presentation-ready form as specified by the requestor (i.e. JPEG or GIF) or in a native DICOM format.

DP (DICOM Parser) is the JavaScript module that handles the parsing of DICOM files and the extraction of the image data and its associated metadata. It is designed to utilize modern multi-core architectures efficiently and reduce memory footprint. Typically, in common server-based architectures DICOM parsing would have taken place server-side. Nevertheless, we hand off that parsing process to the client to reduce network traffic and improve application responsiveness. Therefore, the client-side component (DP), written in JavaScript, is becoming larger and more compute-intensive, increasing the demand for high performance JavaScript execution.

Client-side computation helped to improve the performance of JavaScript engines in the web browsers. Furthermore, considering the widespread deployment of multicores in modern computing systems; exploiting parallelism in these applications is a promising approach to meet their performance requirement [MHSM11]. However, JavaScript has traditionally been treated as a sequential language with no support for multithreading, limiting its potential to make use of the extra computing power in multicore systems. Therefore, the SPA utilizes Web Workers, a Web API that provides a simple means for web content to run scripts in background threads. The worker thread can perform tasks without affecting the user interface. Once created, a worker can send messages to the JavaScript code that created it by posting messages to an event handler specified by that code (and vice versa). In this manner, DP resides in a web worker, main thread posts files to be parsed and DP responds with a similar post message containing all the extracted data (Figure 3.17); obviously, the performance of the main thread is not affected and hence SPA remains always responsive.

DICOM Image Viewer (DIV) is responsible for displaying medical images, while Image Processing Toolbox (IPT) provides key functionality for image manipulation and processing; the combination of these two modules produces a fully-fledged web-based DICOM viewer which is the core unit of the introduced SPA. User experience has been highly prioritized while designing and implementing the specific modules. On one hand HTML5 comes with the canvas element which supports drawing graphics via scripting in JavaScript and while utilizing GPU for rasterization. However, in our case, canvas paradigm suffers some serious performance degradation; any image processing operation involves (1) transferring all pixel data of an image from GPU memory to RAM, (2) executing some JavaScript routine to transform a proportion of that pixel data, and finally, (3) summoning pixel data back to GPU memory. In addition, rendering the content of DICOM files utilizes the canvas-specific APIs, a process that requires special handling of the underline pixel data. That, in many cases, should lead to user experience degradation. For instance, loading DICOM series of multiple instances involve excessive CPU computations resulting in unresponsiveness. To tackle these serious performance issues, a hybrid architecture has been adopted (see Figure 3.19). By stacking up two canvas elements, we achieved the performance of pure WebGL without sacrificing the utilization of HTML5 canvas element for tasks that WebGL lacks sufficient programming capabilities, such as text rendering, etc. The bottom layer supports medical image rendering and post-processing by delegating image processing computations directly to the GPU using GLSL routines, while the upper one facilitates text features and annotation capabilities. Figure 3.20 depicts the performance increment as the decrease of a complete loading cycle (UI latency). The latter is defined (in the scope of the paper) as the overall time required from the selection of a DICOM file by the user until the completion of loading all of its content into the SPA. DICOM content loading requires excessive computation compared with any other visual update. Consequently, it can be considered the bottleneck of each UI update operation and this is the reason why it is defined as the maximum UI latency. Note that time has been asserted in millisecond level accuracy by specialized low-level monitoring routines. Obviously UI latency decreases drastically by the adoption of the aforementioned hybrid architecture.

PeerConnector (PC) is a library developed to harness the power of the latest Web APIs offered by most major web browsers, like WebSockets and WebRTC-related (i.e. RTCPeerConnection, RTCDataChannel). All these relative new standards offer alternatives for the traditional

client-server communication or add new functionalities formerly met only in custom standalone solutions, like peer-to-peer communication. More specifically, PC utilizes WebSockets for client availability detection, chat messages relaying and call signaling --since WebRTC as a standard does not cover this area. The media negotiations and direct session establishment among clients are achieved through the use of the RTCPeerConnection API, whereas the peer-to-peer data transfers (for DICOM files and JSON messages representing user actions) are offered by the RTCDataChannel API. PC manages to abstract away all the quirks of this variety of technologies that are implemented differently among browsers and exposes a simple and intuitive API to the higher-level component providing the social functionality of the SPA, i.e. the Conference Engine (CE). The latter manages all the views enabling user communication in real time, including the contact panel, chat panel, video elements, user and call notifications. The developed SPA may run on a variety of different devices, following the Bring Your Own Device (BYOD) trend.

3.5.2 The system in practice

The developed platform is called HERMES and it is publicly available. In a previous approach [ARTM15a], we described a state representation and implementation based on dynamic data structures conforming to DICOM hierarchy that resided in client memory.

Figure 3.21a depicts a JSON containing instructions of how to render a particular series. This is an indicative example of what is used to transfer back and forth between clients in an active session. In the described implementation, the size of the transferred data is reduced drastically (Figure 3.21b). This is achieved through the radical redesign of the client local state representation based on immutability rules. In this context, every possible user action over the application data, is assigned with a unique name (string of characters) and then it is describe (in code) how that action, when fired, produces a new state. The action representation itself, being a plain JavaScript object, contains the least amount of information required to capture the special meaning of that action and, due to the collaborative nature of the application, it is the only thing that has to be transferred (see Figure 3.22). At this point, it should be mentioned that the advantages in
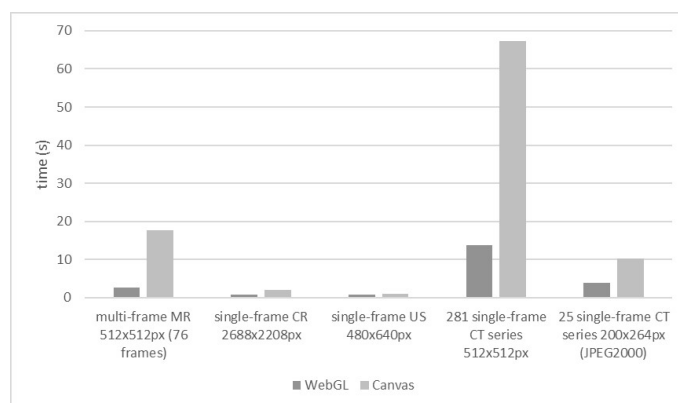
Figure 3.20: Hybrid vs Canvas-based architecture performance. Diagram presents the overall load time of a DICOM file.

```
{
    "peerID": "ho234jf34jfsdfl_43kf",
    "seriesInstanceUID": "1.3.12.2.1107.5.99.2.5562.4.0.575080411334689",
    "left": 458,
    "top": 317,
    "angle": 270,
    "scale": 1,
    "flipX": false,
    "flipY": false,
    "windowCenter": 40,
    "windowWidth": 400,
    "type": "SeriesRendererModel",
    "negativeFilter": false
}
```

(a) A JSON example of what was transmitted in our previous work.

```
{
    "peerID": "ho234jf34jfsdfl_43kf",
    "seriesInstanceUID": "1.3.12.2.1107.5.99.2.5562.4.0.575080411334689",
    "type": "SERIES_MOVE",
    "left": 458
}
```

(b) A JSON example of what is transmitted in our current implementation.

Figure 3.21: The decrease of transmitted bytes increases performance and reduces network bandwidth utilization.

addition to the reduced bandwidth utilization extend to reduced CPU utilization, since less data to be serialized/deserialized.

3.5.3 User interface & visualization
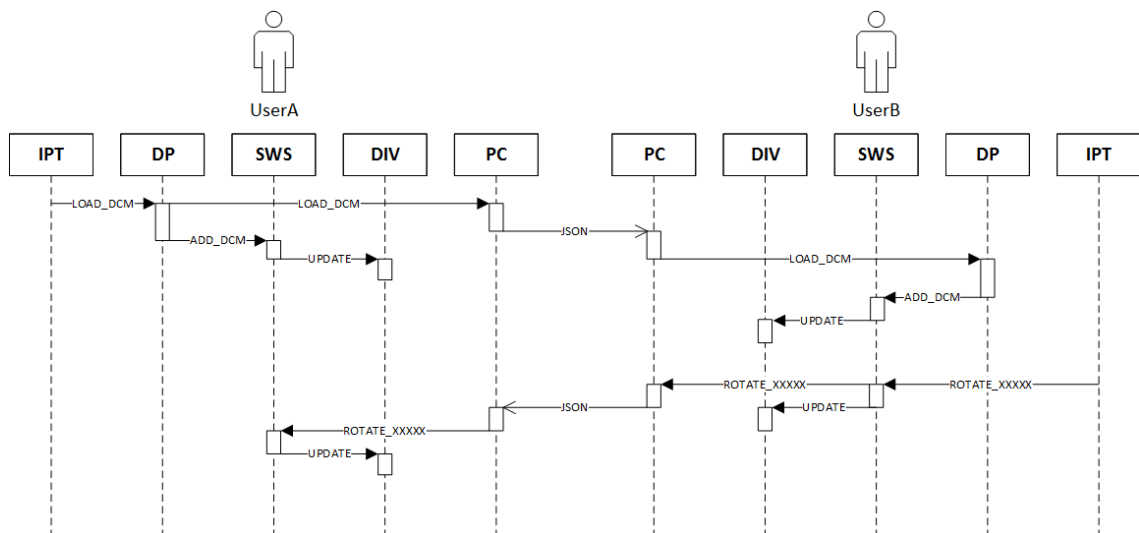
Figure 3.22: Interaction of client modules.

A significant objective of this work is the provision of visualization tools that are intuitive and can be easily used by the physicians [KZH$^+$14]. Browsing across multiple DICOM sources,
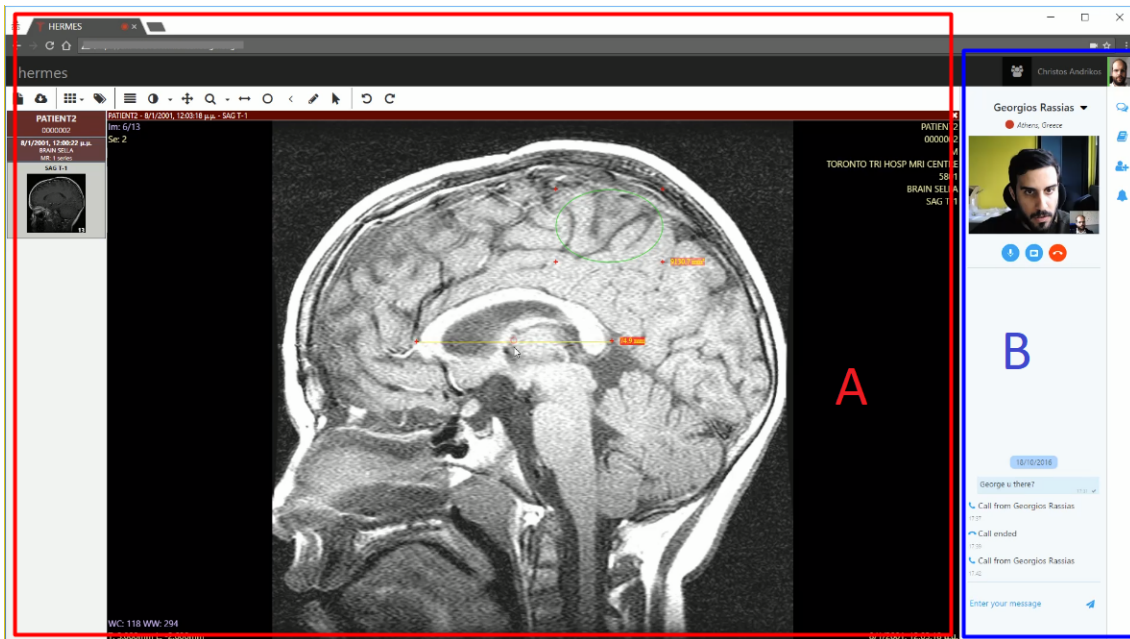
Figure 3.23: SPA screenshot with highlighted subpanels.

aggregating and displaying the evidence required in making informed decisions and reusing the retrieved results to address challenges in different contexts, are aspects of the utmost importance.

To facilitate the intuitive exploration of the information available in any DICOM source, while enhancing the collaboration process among MDTM members, we present a User Interface (UI) comprising various components enabling real-time image processing. The resulting workspace conforms to the basic design guidelines dictated by major medical imaging software solutions (i.e. RadiAnt, OsiriX).

Figure 3.23 exhibits an indicative screenshot of the introduced SPA demonstrating the aforementioned assets. Figure 3.23-A depicts the **medical imaging panel**, the core of our SPA, since it hosts all the key functionality required for medical image viewing. A toolbar facilitates the access of DICOM files or directories and enables user interaction with the loaded images. At the
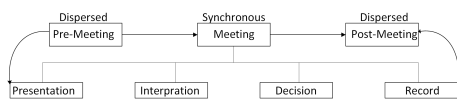
(a) MDTM conceptual model.
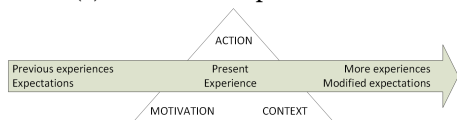
(b) User experience conceptual model.
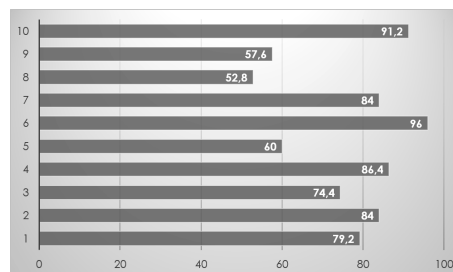
(c) SUS evaluation.

Figure 3.24: User experience evaluation by ten random physicians.

left-side column, there are previews of all the opened series organized according to patient and study unique identifiers. A user can navigate to the desired series and display the corresponding images, by clicking the appropriate thumbnail view, while the actual rendering takes place in the adjacent interactive viewport.

Figure 3.23-B presents the **social panel** where all social widgets reside to support remote MDTM and enhance collaboration among healthcare professionals. It is divided into two sub-panels. The first one dynamically renders information depending on the user selection in the adjacent second one. So, it might contain all user contacts highlighted according to their availability status, the grouped history of user past interactions with others, screen for searching and adding new people, or a window for direct interaction with a specific remote user. The latter combines the video elements that output the local and remote video streams, when available, and the chat history.

3.5.4   Design for failure, deployment automation

Each one of the presented services is "containerized" —an OS-level lightweight visualization technique for deploying and running distributed applications without launching an entire VM for each one. Thus, multiple isolated systems, called containers (i.e. **Docker Containers** (`https://www.docker.com/what-docker`)), can be run on a single control host relieved from the burden of hardware virtualization, and access a single kernel that is shared with the host itself. Consequently, hosts treat the provided services as isolated internal processes resulting to a more efficient scheme regarding to memory, CPU and storage. Any service call could fail due to unavailability of the supplier; the client has to respond to this as gracefully as possible to reduce failure impact to the user experience.

The proposed system deals with that issue through replication of services and smart endpoint deployment over SDN (Software Defined networks). Each container is replicated and every replica participates in a private virtual network that spreads out on the entire cluster of hosts. Replicas of the same service reside behind smart endpoints facilitating ad-hoc discovery of replicated services in order for the latter to be reachable in a transparent way. Thus the proposed architecture ensures high availability and network partitioning tolerance. The lifecycle and replication factor of each microservice is managed by a newly introduced master-slave orchestration system by Google, Kubernetes (K8s). In kubernetes deployment plans and networking rules are defined in a plain declarative way while a description of the state of the entire cluster is backed to an ETCD key-value store to provide failure recovery. As containers can be created and terminated much faster than hypervisor-based instances, our deployment approach is much more agile than the typical hypervisor-based ones.

3.5.5   Use case evaluation and User experience

In this section we present the establishment of a collaboration session. Consider two potential users: physician A and physician B as illustrated in Figure 3.1. Upon login, physician A acquires an authentication token; by triggering UAC to request a digital signature from a third trusted party. User management service returns to physician A the manifest of all online participants matching to his contact list. After physician B logins, the same service updates asynchronously the contact list of physician A, while the latter may access the supported PACS service of his

choice to retrieve and process any medical data. A conference session between the two physicians takes place after the initiator creates a new room. Initial negotiations and further signaling are propagated to participants through the WebSockets API, to result to symmetric WebRTC data channels set up.

Figure 3.24 depicts the graphical representation of user experience conceptual model. User actions are based on previous experiences and expectations to form novel experiences while re-adjusting future expectations. As a consequence, enhancing user experience equates reinforcing the amplifying character of that model. Guided by this concept we applied a novel design resulting to an agile and efficient architecture; any further API extension, forming a distinct microservice, will not affect the provided functionality neither the performance of the platform and thus user experience will not decline. Furthermore, fine-graining the backend infrastructure while maintaining low overall complexity (i.e. microservices interact each other as less as possible) eliminates any design bottlenecks.

The provided architecture in conjunction with the SPA ergonomic design ensures that (i) increased workloads can be easily handled by dynamically spawned service instances, (ii) further functionality extension could not affect the performance and responsiveness of the platform.

Our claim about user satisfaction is not untenable; rather it is backed by the results of a SUS questionnaire [B$^+$96] that has been filled by ten random physicians, after they engaged in a collaboration session with one of the members of the development team. They were asked to evaluate certain aspects of the platform related with the user-friendliness, the ease of learning, the provided feature integration and observed inconsistencies, if any. The results, depicted in Figure 3.24c, are highly positive and motivational, especially if someone takes into account the alpha release that physicians were given access to —well above 68 that is considered as the baseline and pretty close to 80, above which people are expected to recommend it to their friends or colleagues.

Table 3.2: Related work, brief feature comparison. **Ref** column holds the number of the corresponding citation.

Ref	audio-visual communication	real time workspace sharing	end-to-end encryption	P2P connections	image processing toolkit	Integration with PACS	3D reconstruction	Session Recording	BYOD - installation free	Social Networking aspects	AI
[RSR04]					✓		✓				
[BPYM05]					✓		✓				
[htm]					✓	✓	✓		✓		
[sun]					✓	✓	✓		✓		
[LFC00]		✓			✓	✓	✓				
[MDK06]	✓	✓			✓						
[MY14]	✓	✓			✓			✓		✓	
[VG11]	✓	✓	✓	✓	✓	✓	✓	✓	✓		
[Xie12b]	✓	✓	✓		✓	✓	✓	✓	✓	✓	
HERMES	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓

CHAPTER 4

Urine Bladder 3D reconstruction with fusion of non-invasively acquired medical imaging modalities

The urinary bladder is an elastic reservoir that is responsible for the low pressure storage of urine. A continent, pain free and healthy bladder is crucial for the preservation of a good quality of life. Numerous conditions disturb the anatomy and physiology of the urinary bladder, leading to insufficient and restricted evacuation of urine [VRP+12].

Urinary bladder imaging is essential to monitor disorders such as urinary incontinence vesicoureteral reflux, cystitis, glomerulation, bladder cancer [DHT+04][KS03]. The severity of these diseases varies from causing minor discomfort to being fatal. Urinary bladder cancer is a common disease strongly related to morbidity worldwide. It ranks the fourth and twelfth cause of death in men and women, respectively [VRP+12]. Moreover, it presents a high recurrence rate which dictates a long-term surveillance plan after the initial treatment. According to specialists, early detection of bladder cancer holds the key to its successful treatment. The cancer begins on the wall of the bladder and eventually spreads, which is fatal to the patient. Therefore, imaging the urinary tract (ureter, bladder and urethra) is very important for early diagnosis [KBR+17]. Considering the value of such a successful early diagnosis, we present a **novel approach assisting the detection bladder cancer to its early stage**. It is about a Computer-Aided Diagnosis (CAD) workflow to highlight any "suspicious" area around the bladder, leveraging inter-modality knowledge through classic computer vision approaches combined with the most recent Neural Network advances. Healthcare experts require access to realistic views of anatomical lesions to fully understand the spatial anatomic relationship of the 2D image frames and conclude any successful diagnosis. The literature proposes multiple innovative visualizations methodologies of the bladder silhouette alongside any potential tumour artefacts [SIKW+20][HRX+12]. Still, the majority of these methods lack the multimodality approach, which deteriorates their overall sensitivity [FD96].

In this setting, the introduced platform integrates a hybrid proof of Concept POC) methodology for an accurate urine bladder 3D reconstruction leveraging both **ultrasound** and **photoacoustic** imaging modalities. In a high level approach, the ultrasound images (slices of the same study-series in DICOM) are combined with the photoacoustic ones to feed a hybrid[1] **convolutional autoencoder** which segments accurately the area area between the inner and the outer bladder wall on each DICOM slice. Next, by combining the extracted segments of each slice, an

---

[1]it also takes advantage of classic computer vision algorithms

accurate 3D model of the bladder is constructed to provide a visual artefact of the thickness of the bladder's wall. Given that the thickness of the bladder's wall is a well established biomarker for the early cancer diagnosis of urinary bladder, the significance of such visualization artifacts is self explanatory. Furthermore, by including the particular methodology into the imaging toolkit of the the introduced platform [ARTM15b], we pave the way the next generation of collaboration where software engineers and domain experts (clinicians) may work in a perpetual feedback loop to improve healthcare services. We argue that The chances of developing new models and CAD strategies are proportionate to the experts using the introduced platform and the data availability per se. The novel methodology we discuss may be extended to leverage further imaging modalities such as Computed Tomography (CT), Position Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) [ZAL15][HH13] so as to be applied to various analogous cases of medical image analysis.

4.1   Background

The majority of computer vision advances and applications on medical imaging data focuses on outperforming doctors' eyes. But this is not due to a race between human and algorithms rather due to the need of automating tasks that can enhance healthcare provisioning. For instance organ segmentation is a prerequisite for CAD systems that may assist a medical diagnosis. In fact, the segmentation algorithm is the most important and basic for any image analysis task. Especially in the case of medical imaging it arguably enhances the level of disease prediction and therapy. Authors in [PXSY18] use the Ensemble of Polyharmonic Extreme Learning Machine (EPELM) and positive feedback of perception to detect salient objects, which is totally data-driven without any prior knowledge and labeled samples compared with the existed algorithms. A positive feedback module based on EPELM focuses on fixation area for the purpose of intensifying objects, inhibiting noises, and promoting saturation in perception. Experiments on several standard image databases show that the novel algorithm outperforms the conventional saliency detection algorithms and also segments nucleated cells successfully in different imaging conditions.

Recently, random walkers (RW) have attracted a growing interest to process segmentation of medical images. However, classical RW method needs a long computation time and a high memory usage because of the construction of corresponding large-scale graph to solve the resulting sparse linear system. The work in [DZL^{+}17] incorporate the prior (shape and intensity) knowledge in the optimization of sparse linear system − a perspective we also follow to our core methodology. Integrating the Bayes model into the RW sparse system, the organ is automatically segmented for the adjacent slice, which is called RWBayes algorithm in the article. Compared with the conventional RW and the state-of-the-art interactive segmentation methods, their method can significantly improve the segmentation accuracy and could be extended to segment other organs in the future.

The practice of modern medicine relies heavily on synthesis of information and data from multiple sources; from initial diagnosis to treatment planning and from delivery to monitoring the patient post-therapy, these data drive the decisions about how the patient is treated and help assess the progress and efficacy of therapy. While X-ray CT remains the primary imaging modality for most aspects of treatment planning and delivery, the use of data from other modalities such as MRI and MR spectroscopy (MRS) and positron/single photon emission tomography

(PET/SPECT) is becoming increasingly prevalent and valuable, especially when taking advantage of highly conformal treatment techniques such as intensity-modulated radiotherapy [1–3]. These additional imaging studies provide complementary information to help elucidate the condition of the patient before, during and after treatment. The work in [HPS+20] offers a deep dive into medical data fusion techniques.

The standard test for the diagnosis of bladder cancer is cystoscopy and biopsy. Flexible office cystoscopy is as reliable for the diagnosis of bladder cancer as rigid endoscopy [GLW+01]. Currently, white light cystoscopy (WLC) remains the standard of care for diagnosing a bladder tumour and enables the urologist to map and resect all visible tumours. Although WLC has excellent sensitivity and specificity for diagnosing large papillary tumours, it is less reliable for diagnosing small papillary tumours and carcinoma in situ (CIS). In particular, standard white light cystoscopy misses up to 20% of bladder tumours. Authors in [SJC+19] suggest a novel methodology that enhances classic light cystoscopy with deep learning techniques to tackle that discrepancy. Another advanced imaging process is the porphyrin-induced fluorescence cystoscopy. It benefits from photoactive porphyrins, which tend to collect in neoplastic tissue and emit red fluorescence under blue-wavelength light. Although HAL and 5-aminolevulinic acid (5-ALA) have been assessed in clinical studies, only HAL is FDA approved, with indications for bladder cancer surveillance for rigid and flexible cystoscopy [PWK+20]. Most of these methodologies are either invasive or try to locate tumours explicitly and potentially suffer a high false-positive rate. For this reason, they cannot be considered a de facto approach for an initial clinical case assessment.

An alternative to the aforementioned approaches is the inspection of accurate and recent 3D models of the urinary bladder of the patient. Any tumor should be expressed in some sort of bladder wall thickness abnormality that could be easily located by a domain expert. The work in [MJMT13], is based on the information from multi-view planes, to propose a level set algorithm for reconstructing the 3D shape of the bladder using the cross-sectional boundaries of Computed Tomography (CT) frames. The authors of [JLSH+19] present an immunohistochemical staining method that delineates neurovascular structures in the mouse urinary bladder by using immunohistochemical staining with three-dimensional reconstruction to evaluate pathologic changes in a variety of urinary bladder diseases. They managed to reconstruct three-dimensional images of rich anatomic information obtained from thick-cut sections.

This work suggests a novel approach for semantic segmentation (i.e. clustering parts of images together which belong to the same object class) of the urinary bladder wall to be the intermediate step of an accurate 3D bladder model. By integrating this methodology into the toolkit of the work in [ARTM15b], we admire providing accurate visualization artifacts to assist experts in locating potential tumours. Our approach's novelty lies in the fusion of multimodal imaging data usually acquired through non-invasive medical imaging techniques that simplify the overall diagnosis process.

4.2 Methodology

A normal urinary bladder has a thin, smooth muscular wall mainly composed of the detrusor muscle and is the temporary storage reservoir for urine. Thus we argue that any 3D reconstruction methodology should focus on extracting sufficient information related to the bladder wall.
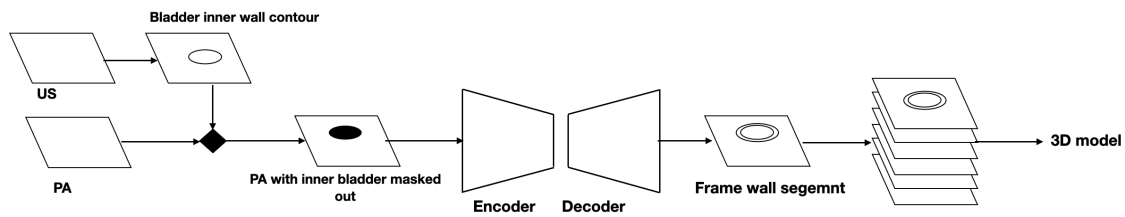
Figure 4.1: Bladder 3D reconstruction, an overview of the methodology

A common procedure for 3D reconstruction is to use the cross-sectional regions of the bladder segmented on the images: a rough profile of the bladder can be sketched through stacking the sequential boundaries from the same spatial view; then, the 3D surface model can be reconstructed using, for example, the marching cubes algorithm[LC87]. Therefore, the reliability of a 3D model depends on two critical factors:

- the accuracy of the segmented regions

- the spatial interpolation between neighboring slices

Regarding the first factor, effective segmentation algorithms [DLB$^+$10], [MJMT11] have been proposed based on the appearances of the bladder. Although some aspects of these algorithms still need to be improved, in most cases correct results can be obtained; hence, segmentation accuracy is not the focus in this paper rather proposing a novel method of hybrid inter-modality segmentation. **Our focus is improving the life of the patient by working with a dataset acquired in a non-invasive way.** Regarding the second accuracy prior, the spatial interpolation is used to recover the missing information at the slice gap. The slice gap is a parameter determined at the imaging phase: the closer to zero this value is, the more information has been acquired. If the gaps between neighboring images are tight enough, a realistic 3D surface model can be obtained. High quality multiplanar reconstruction can be rendered precisely through isotropic or near-isotropic scanning provided in the CT urography (CTU) and MR urography [VDMCML$^+$08], [SLAJ09], [CCK$^+$02] but this is out of the scope of our work.

Our focus targets the utilization of imaging data acquired through two non-invasive techniques (viz. Ultrasound and Photoacoustic imaging) to provide an accurate 3D model of the urinary bladder. Shortly, the two image modalities are combined into a series of co-registered frame. Next, each resultant frame feeds a neural network that extracts the bladder wall segment (i.e. the area between the inner and outer bladder wall). Finally, the extracted segment frames are combined to form the final 3D reconstructed model of the bladder (figure 4.1).

4.2.1 Data input, dataset

The proposed software receives from user as input data both ultrasound and photoacoustic image sequences in Digital Imaging and Communications in Medicine (DICOM) format. Figure 4.2 presents the image modalities that our approach relies upon.

On the one end, considering the epidermal ultrasound is an imaging modality that provides essential artefacts for assessing the urinary infection and the overall bladder condition. It is also a prestigious tool of the clinician's arsenal since it displays 2D images of the under-skin anatomical

structure of tissue and organs in real-time to enhance clinical examination. Ultrasound images feature a 92.7% sensitivity in depicting lesions greater than 5 mm. However, an early diagnosis requires locating tumours sizing less than 5 mm, which challenges the 26.5% sensitivity of ultrasound images to this task. Thus, during any regular cancer assessment, the ultrasound findings are considered vague and insufficient, leading the patient to additional cystoscopy examinations to clarify potential bladder tumours [HH13] further.

On the other end, Photoacoustic imaging (PAI) it is a molecular imaging tool that has emerged as a noninvasive and nonionizing in vivo biomedical imaging modality with relatively deep tissue imaging ability. Briefly, tissue is irradiated with laser pulses and absorbs the light in variable degrees based on its optical properties, causing the tissue particles to heat and thus increase their kinetic energy. The overall result may be enhanced further by the infusion of nanoparticles designed for increased photoacoustic contrast. That produces acoustic waves that may be detected and visualized due to the thermoelastic effect. Relative to other optical technologies, such as fluorescence imaging, PAI has a 100% relative sensitivity to optical absorption and can provide an image in a region up to several centimetres deep in biological tissue [FD96]. Combining the spectral selectivity of molecular excitation with the high resolution of ultrasound images is capable of depicting tumour depth in bladder cancer.

This work has been conducted on a dataset of in-vivo clinical trials of animal models. In particular Two female Fischer rats and one female Albino mouse are the three subject animal models. The rats where periodically monitored for 9 weeks and have either been exposed to N-butyl-N (BBN) nitrosamine to develop a slow growth of tumour in their urinary bladder area over a prolonged period [VNCLO12], or tumour cells were injected directly into the bladder area to accelerate the development of carcinoma (orthotopic). The same ultrasound and photoacoustic imaging were applied to all three subjects to extract the required dataset. The size of the acquired data is in total 700 images. The dataset has been heavily augmented by a set of non-elastic image transformations.

Figure 4.2 depicts a typical instance (a tuple of frames) of the input data of the suggested pipeline. A typical Ultrasound (US) image is depicted in 4.2a, while 4.2b presents the photoacoustic one. The two imaging modalities come from different sensors, and feature the same spatial resolution so as to be easily co-registered. The rationale of co-registration is to ensure that the two images being spatially aligned should afford overlapping ROIs.

4.2.2 Inner bladder wall extraction, a classic computer vision approach

Figure 4.2a depicts an ultrasound image. The internal space of the bladder is depicted as a distinguishing dark concrete area. Since the available dataset is restricted in size, **the inner contour of any US image is chosen to be approached according to the active contour method**, [CCCD93] along the ultrasound images. Active contour is a pure (in the manner of not affording any domain semantics) segmentation method that uses energy forces and constraints to separate the pixels of interest from a picture for further processing and analysis. Contours are the boundaries that define the region of interest in an image. Any contour is the interpolation of finite set of cardinal points that have been interpolated in a linear, splined, or polynomial way, according to the curve complexity that is to approached. Active contours is a family of optimization techniques for image segmentation. Similarly to every optimization problem, all of them do employ an energy function which is related to the curvature of the image curve to be approached, and

a set of forces to interact with that energy function. They do also define the notion of external energy to be the sum of all the external forces. External energy is specifically used to control the location of the contour onto the image, while internal energy is used to govern deformable changes. In this sense, the active contour methodology refers to a family of iterative algorithms that lookup for some equilibrium regarding the internal and external energy models. Active contour algorithms vary mainly to the energy model definition (and in particular in the internal energy one) and consequently to the convergence rate, the accuracy and the stability of the estimated equilibrium.

A typical active contours algorithm (model) is the Geometric Active Contours (GAC) which is mostly used in medical image computing and in particular on image-based segmentation. (GAC) is a contour model that adjusts the smooth curve established in the Euclidean plan by moving the curve's points perpendicular to each other and at planar steps proportional to the curvature of image's region (Figure 4.3).

4.2.3 Inferring the outer bladder contour, The Unet approach

The quintessence of any visual task especially in biomedical image processing is the automatic annotation of a Region of Interest (ROI). Given any image and a domain of concepts any "highly intelligent" operation should come up with mapping of ROIs to concept:

$$\mathcal{I} : \{(b, x, y), \dots\}, \, where \, x, y \, \in \mathbb{R} \, and \, d \, imaging data \tag{4.1}$$

$$\mathcal{P}(\mathcal{I}) : the \, power \, set \, of \, I \tag{4.2}$$

$$\mathcal{C} : a \, set \, of \, concepts \tag{4.3}$$

$$G(I) : I \mapsto (k, c) \, where \, k \, \in \, \mathcal{P}(\mathcal{I}) \, and \, c \, \in \, \mathcal{C} \tag{4.4}$$

But concepts are fuzzy. The same stands for human interpretation of reality thus the definition 4.4 should be updated to include the notion of probability as it follows:

$$G(I) : Pr((k, c)|I) \, where \, k \, \in \, \mathcal{P}(\mathcal{I}) \, and \, c \, \in \, \mathcal{C} \tag{4.5}$$
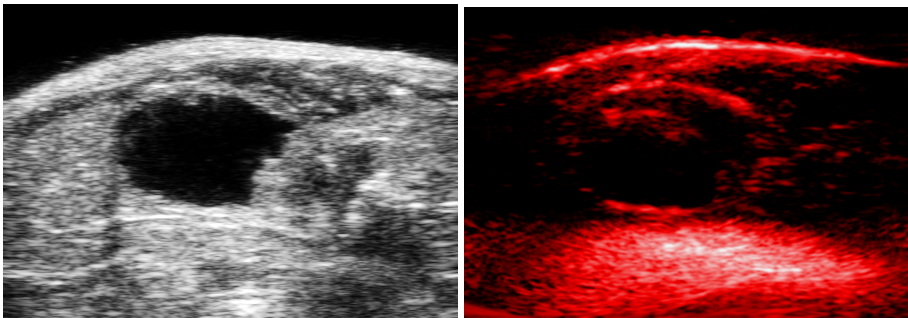
(a) An arbitrary Urine bladder Ultrasound frame

(b) An arbitrary Urine bladder Photoacoustic frame

Figure 4.2: Instances of the two input imaging modalities
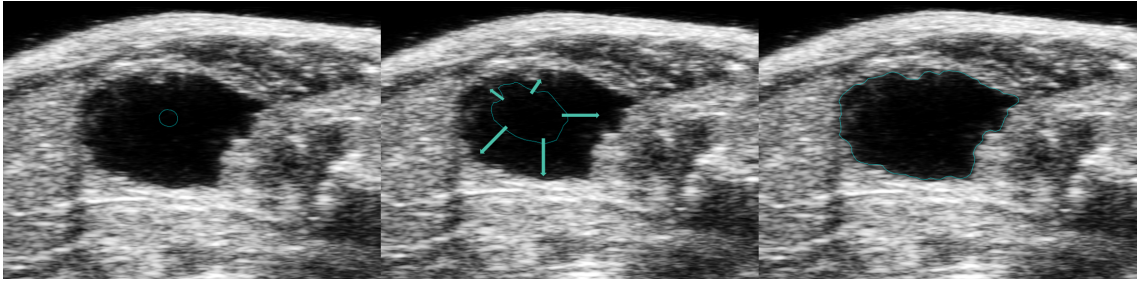
Figure 4.3: Steps A to C of active contour algorithm

What 4.5 defines is the probability space of pixel to concept correlation. Exploring such a space is usually a difficult task that requires heuristics trading the number of concepts against the quality of ROIs against the computing power necessary for some particular task step.

The most advanced family of heuristics is the convolutional neural networks (CNN)The typical use of CNNs is on classification tasks, where the output to an image is a single class label (i.e a concept).

According to the bibliography U-Net (Figure 4.4) is one of the most successful deep Neural Network topologies that do not require a vast amount of training data when it comes to semantic image segmentation. And that this the main reason why U-net fits perfectly to such a segmentation task: **the lack of serious datasets in the manner of quantity**.

The network architecture is illustrated in Figure 4.4. It consists of a encoding (left side) and a decoded path (right side). The information flows though the encoding to path to be encoded to a lower dimensionality space and next it moves through a decoder to be decoded once again back to the original dimensionality one. The encoding is the typical contracting path that any convolution neural network applies to its input. It consists of the repeated application of two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. At each downsampling step we double the number of feature channels. The decoding path, is an expansive one that consists of steps of upsampling of the feature map followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the encoding path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In total the network has 32 convolutional layers. Figure 4.5 depicts the remarkable learning of the network. It seems to be a perfect fit to tackle the particular challenge.

Data input, preoprocessing   According to Figure 4.1, U-net consumes photoacoustic images to extract the outer wall of urinary bladder. The inference of the outer part of the bladder is quite a challenge even for high skilled experts. In US images there is not sufficient information to accurately recall the outer contour of the bladder. That is the reason why our methodology invokes PA imaging modalities as well. Figure 4.6a highlights how the vascularization of bladder is able to highlight the outer bladder contour when checking a PA image. This is the appropriate way for the experts to annotate outer bladder contour. Though this may be extra tricky if there
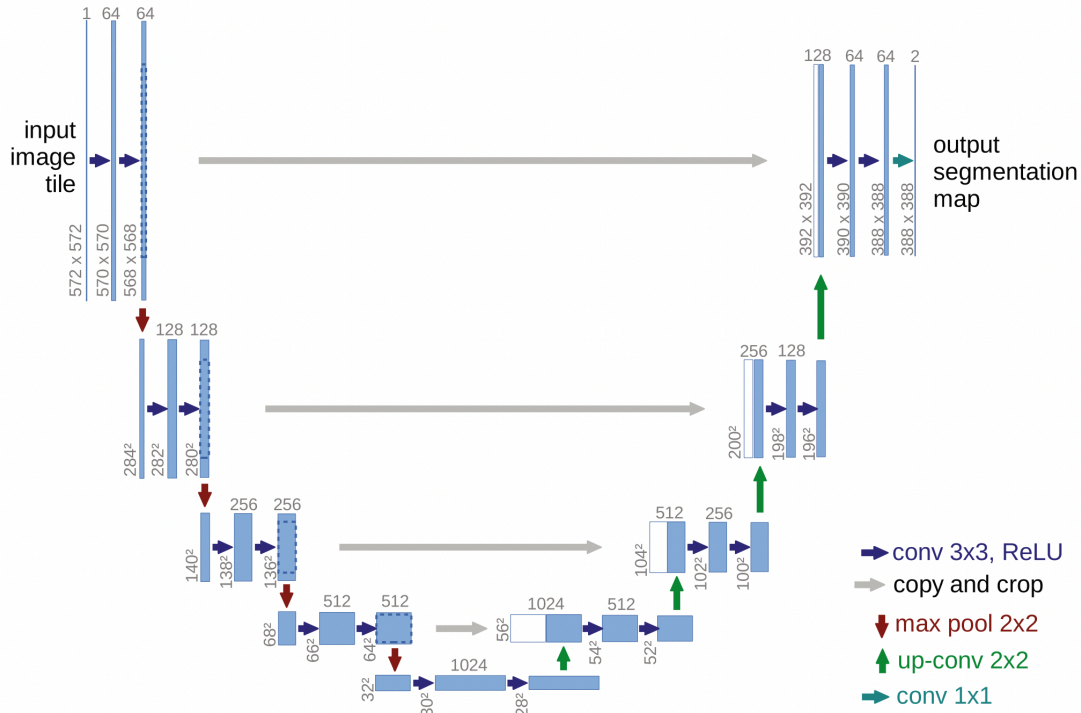
Figure 4.4: Typical U-net architecture

is no inner contour highlighting while they do work on PA images. For this reason bladder internals was masked out before the feeding the U-net structure. According to figure 4.1, the extracted bladder internal ROI registered directly onto the related PA image to be masked out.

Furthermore, due to the confined dataset, excessive data augmentation has been performed. By applying elastic deformations to the available photoacoustic images (viz. rotations, zoom, crops, random noise addition, shear, etc.), the training process has been seriously enhanced. The network learns invariance to such deformations, without the need to see these transformations in the original dataset. This is particularly important in biomedical segmentation, since deformation used to be the most common variation in medical imaging and realistic deformations can be simulated efficiently to contribute further to learning invariance as highlighted in [BSH17] in the scope of unsupervised feature learning.

4.2.4 Reconstructing to 3D space

Figure 4.7 depicts the final 3D reconstructed bladder model according to the following steps:

- for each (PA_i, US_i)

  - Extract the inner bladder wall according to active contour algorithm in the US_i image frame
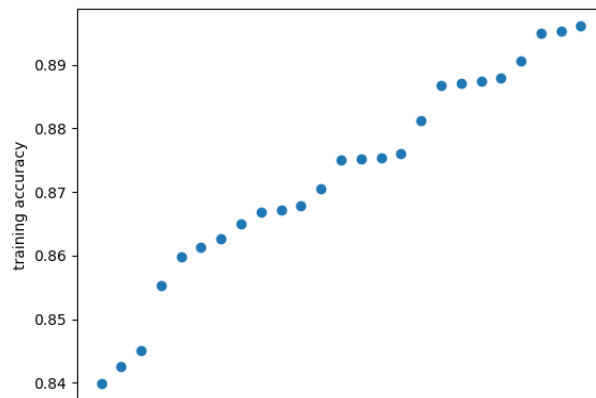
Figure 4.5: Unet training accuracy for the first 25 epochs
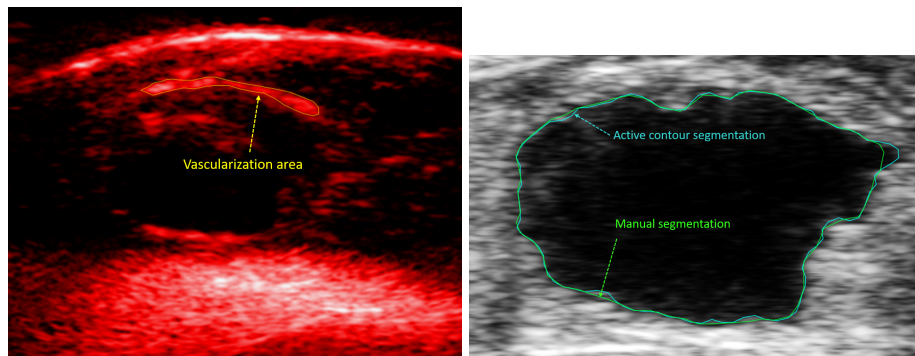
(a) Vascularization: The distribution of vessels in a photoacoustic image that drives the inference for the location of the bladder outer wall

(b) Extracted bladder wall inner contour

- – Mask out the interior of urinary bladder in the PA_i image frame according to inner contour extracted in the previous step
- – Infer the exterior by feeding the masked PAI to the Unet model

- Convert all points of the two contours to polar coefficients relatively to the center of gravity of the interior and exterior walls respectively

- Sort the points in a clockwise fashion

- Enhance points of current frame to feature the same cardinality with the previous one according to the cubic spline interpolation

- Perform an image triangle based reconstruction step for the current and previous frame

At this point it is worth noting an optimization we introduced to the active contour algorithm that significantly reduces the overall computation cycles. Given that the input to the inner wall is inferred into an iterative notion for a set of consecutive imaging frames, we utilize the concept of vertical spatial locality. Considering that, we approximate the inner wall of the k_i frame utilizing information by the next adjacent k_{i-1} one. In particular we select the already extracted to be the initial contour of the active contours algorithm while assessing k_i frame. In this way, we significantly reduced the execution time of the inner borders segmentation algorithm. Thus, the active contour algorithm needs quite less iterations to find the inner borders of the bladder because it is trying to detect them close to the segmentation result of the previous frame.

4.3   Evaluation

The authors in [TH15] present the 20 widely accepted metrics related to 3D reconstruction operations. In this work we employ the related 2D metrics following four metrics to communicate a clear performance retrospection. This is due to the fact that both the any inferred or ground truth value belongs to the 2D space.

- Dice Coefficient: Overlap metric. It is the most used in metric in medical image segmentation. Used to directly compare automatic and ground truth segmented areas.

- Sensitivity measures the portion of positive pixels in the ground truth that are also identified as positive by the automatic segmentation.

- Specificity measures the portion of negative pixels (background) in the ground truth segmentation that are also identified as negative by the segmentation being evaluated

- Hausdorff Distance Hausdorff distance, the Average distance between the finite set of n samples (GS) of the ground truth and the subset of n samples (IS) of the inference
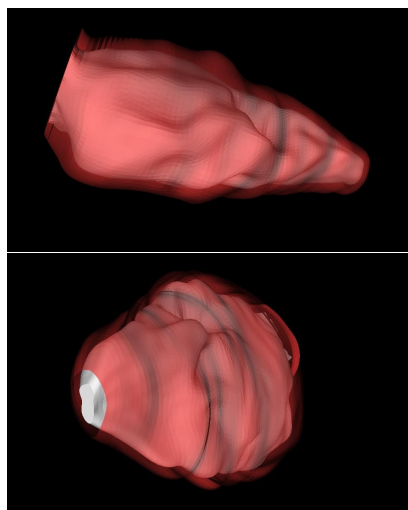
Figure 4.7: 3D reconstructed urinary bladder

Inner Wall According to the results depicted at table 4.1, our methodology regarding the segmentation of the interior bladder wall performs a remarkable accuracy. The Hausdorff distance between the extracted contour and the one suggested by the domain experts remains less than 0.3 millimeters across all the three of subjects. The testing dataset is a randomly generated horizontal set of 60 arbitrary [2] ultrasound frames per subject (rat).

	Dice Coefficient	Sensitivity	Specificity	Hausdorff Distance (mm)
subject A	0.99 < 0.02	0.99 < 0.02	0.99 < 0.01	0.04 < 0.06
subject B	0.92 < 0.04	0.85 < 0.07	0.99 < 0.01	0.20 < 0.05
subject C	0.98 < 0.03	0.98 < 0.03	0.99 < 0.01	0.05 < 0.09

Table 4.1: Mean evaluation metrics of the extraction of the inner wall of the bladder for 60 arbitrary frames per subject ± the standard deviation of the error

Outer Wall Table 4.2 evaluates the overall performance of the suggested hybrid pipeline for the extraction of the outer wall of the urine bladder. Once again the testing dataset is randomly generated. It consists of 60 arbitrary photoacoustic frames of the urinary bladder per subject. The remarkable Hausdorff Distance of 0.3 millimeters highlights that the suggested methodology should be considered successful.

	Dice Similarity	Sensitivity	Specificity	Hausdorff Distance (mm)
subject A	0.93 < 0.04	0.90 < 0.04	0.99 < 0.01	0.30 < 0.40
subject B	0.95 < 0.01	0.98 < 0.01	0.98 < 0.01	0.15 < 0.02
subject C	0.92 < 0.04	0.92 < 0.01	0.99 < 0.01	0.23 < 0.17

Table 4.2: Mean evaluation metrics of the extraction of the outer wall of the bladder for 60 arbitrary frames per subject ± the standard deviation of the error

Validation of masking hypohtesis One of the major contributions of this work relies on the hypothesis of the masked Unet input. We argue that masking out the pixels of the Photoacoustic image input that represent the internal area of tbe urine bladder, may boost the segmentation performance. To test out that hypothesis we trained two Unet instaces. The classic approach is trained to provide the outer wall of the bladder given a photoacoustic image while the hybrid one refers to the Unet instance trained to operate on masked photoacoustic frame inputs. Figures 4.8, 4.9, 4.10 and 4.11 present the metrics for arbitrary testing frames for all subject. Hybrid metrics (blue) not only clearly (1) outperform classic ones (orange), but also they (2) keep a more solid distribution. The later suggests that that the low level classification of pixels is consistent when it comes to the hybrid implementation.

---

[2]due to the geometry of bladder, the top and bottom two frames have been excluded to avoid any validation discrepancies
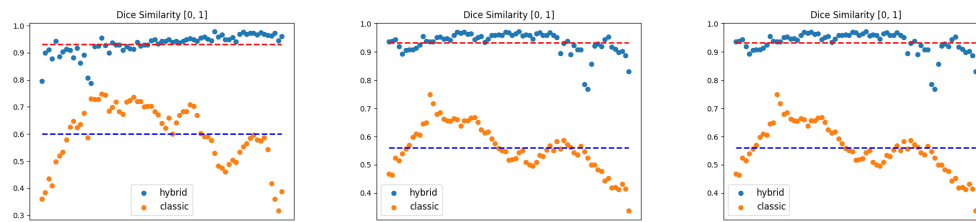
Figure 4.8: Dice Similarity, hybrid vs classic approach for for subjects A, B, C respectively
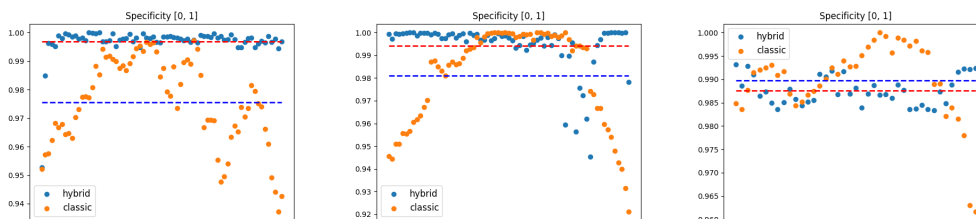
Figure 4.9: Specificity, hybrid vs classic approach for for subjects A, B, C respectively

4.4   Discussion

In this section we discuss a novel methodology to 3D reconstruct the urinary bladder of a patient, leveraging only non invasive medical imaging data. The novelty lies on leveraging inter-modality knowledge through classic computer vision approaches combined with the most recent Neural Network advances to tackle an important medical challenge: to assist **early urinary bladder cancer**. We argue that by extracting an accurate enough 3D representation of the urinary bladder we may pave the way for the development of new CAD operations or further research opportunities. By integrating the particular methodology into a concrete platform described in [ARTM15b], we expose healthcare practitioners to new strategies for healthcare provisioning and patients to closer to personalized treatments. The chances of developing new models and CAD strategies are proportionate to the experts using the introduced platform and the data availability per se. The novel methodology we discuss may be extended to leverage further imaging modalities such as Computed Tomography (CT), Position Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) [ZAL15][HH13] so as to be applied to various analogous cases of medical image analysis.

This should not be considered a black box integration. Due to the hybrid canvas implementation the polygon mesh that approaches the actual image contour belong to the higher visualization layer (Figure 3.19), thus all points are event-able; user can manually operate at each on of one of them. Thus the supported "manual" functionalities are:

- correct the border points by shifting them on images

- cut a part of the borders and replace it with new manually interpolated points
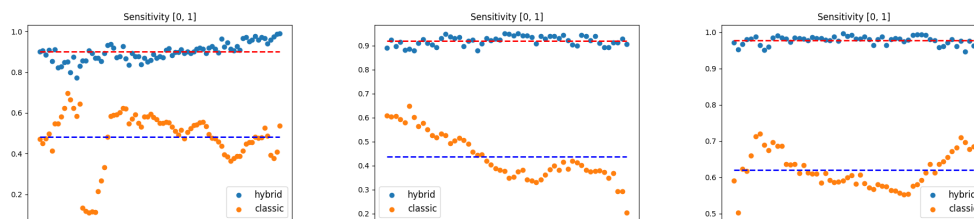
Figure 4.10: Sensitivity, hybrid vs classic approach for for subjects A, B, C respectively

- perform fully manual segmentation of the bladder borders from scratch.

Another interesting insight is related with the cycle of the medical data annotation and exploration. In retrospect, the development of the current or any similar functionality requires close communication between doctors and computer scientists. Doctors should annotate data and share knowledge with the computer scientists. The latter should share artifacts with the doctors to be reviewed. Hermes seems to be the common place for such a consistent collaboration loop. In this work, for instance we as computer experts had to understand the "vascularization" concept (Figure 4.6a), and doctors should be able to evaluate the extraction of the inner contour of the urinary bladder at the first step of the pipeline (Figure 4.6b).
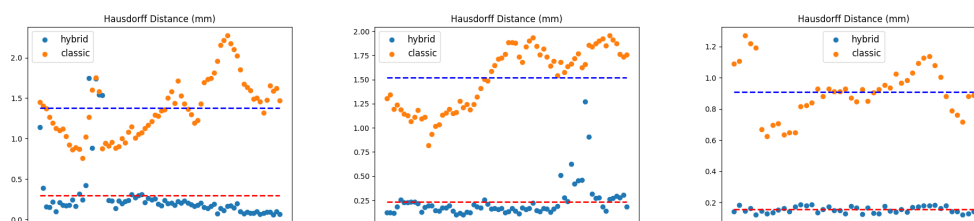
Figure 4.11: Sensitivity, hybrid vs classic approach for for subjects A, B, C respectively

CHAPTER 5

The RNA psuedoknot use case

The use of RNA-sequencing (RNA-seq) in molecular biology research and clinical settings has increased significantly over the past decade. Despite its widespread adoption, there is a lack of simple and interactive tools to analyze and explore RNA-seq data. Many established tools require programming or Unix/Bash knowledge to analyze and visualize results. This requirement presents a significant barrier for many researchers to efficiently analyze and present RNA-seq data.

Once the wet lab components of an RNA-seq experiment are completed, the data must be analyzed computationally. To date, a multitude of tools are available to researchers depending on the experimental question (e.g. the discovery of novel transcripts or determining gene expression changes) [CMT+16, CSDL17].

Regardless of the analysis tool selected, the vast majority of currently available tools require knowledge of programming (C/C++, Perl, Python, R) or shell scripting (Unix/Bash shell). DE-Seq2, one of the most popular analytical software packages for DGE, is written in R and requires an understanding of this language to manipulate data and visualize results [LHA14]. The requirement for users to navigate one or more computational languages in order to analyze RNA-seq data presents a substantial barrier for many researchers who are adept with respect to the wet lab components of RNA-seq but unfamiliar with the computational aspects.

This fact contradicts the rise of user-friendly LC/NC tools that admittedly increases the overall effectiveness of healthcare practitioners and researchers. Having said that, let us introduce a novel RNA analysis workflow that can be included in the arsenal of the introduced platform.

5.1 Introduction

The RNA molecule, being the intermediate representation of the information flowing from DNA to proteins, holds a crucial role in many biological processes. Recent studies reveal that its detailed structural analysis is of utmost importance. The RNA structure is often depicted by a 2-D representation of the base-pairing nucleotide (A-U, C-G, and G-U pairs) known as the "secondary structure," which conduces to the more-complex construction 3-D representation, termed the tertiary structure. The substantial role of RNA in the expression of proteins, along with its contribution to the gene expression regulation, catalysis, and site recognition, requires the

enlightenment of its territorial structure, which is associated with considerable biological functionalities.

More specifically, non-coding RNAs (ncRNAs) are functional RNA molecules transcribed from DNA but not translated into proteins. The latter must not be misinterpreted as not enclosing important information or contributing to any biological operation. Current evidence implies that ncRNAs transcribe most of the genomes of mammals and other complex bions, to contradict the widespread assumption that proteins transcribe most genetic information. Their purpose is to fulfill diverse catalytic and structural functions, along with regulating gene expressions at the transcriptional and post-transcriptional level.

The literature contains a considerable number of publications discussing the prediction of RNA secondary structure. The employed methodologies are mainly based on dynamic programming algorithms, thermodynamic models, stochastic methods, and syntactic pattern-recognition techniques. A thorough analysis of the related work is presented in Section 5.3. In this study, a methodology of predicting H-type RNA pseudoknots, a familiar yet complex structure of the RNA secondary representation, was granted. Initially, the RNA raw string was parsed via a context-free grammar parser for all trees that include a pseudoknot to be produced. Next, all trees are traversed to identify additional base pairs around the pseudoknot. Finally, the optimal tree was selected, maximizing the number of base pairs, while minimizing the free energy of the pseudoknot. For the first task, i.e., detecting possible pseudoknots, an alternative methodology was implemented as well using a brute force algorithm. The proposed methodology succeeded in predicting the core stems of any RNA pseudoknot of the test dataset by performing a 76.4% recall ratio. It achieved a F1-score equal to 0.774 and a MCC equal 0.543. A dataset [dat] of 262 RNA sequences was used, proving a performance speed of 1.31, 3.45, and 7.75 compared to three well-known platforms [JWMW18, SKH^{+}11].

The article is organized at the following sections. In Section 5.2, all required definitions and preliminaries are given. Section 5.3 presents any related publications thoroughly. In Section 5.4, the proposed methodology is analyzed, and an illustrative example is presented. Section 5.5 focuses on implementation details, while, experimental results for several RNA structures are discussed in Section 5.6. In the end, Section 5.7 concludes the presented methodology and implementation and describes future modifications and improvements.

5.2 Theoretical Background

Non-coding RNAs (ncRNAs) are functional RNA molecules that are transcribed from DNA but not translated into proteins. This is not to say that ncRNAs do not provide knowledge or serve a purpose. Although it has been generally assumed that proteins transcribe most genetic information, recent evidence suggests that ncRNAs transcribe most of the genomes of mammals and other complex bions. Their purpose is to fulfill diverse catalytic and structural functions and regulate gene expressions at the transcriptional and post-transcriptional level.

5.2.1 RNA

RNA is the cornerstone of various biological processes; it carries genetic information that is encoded into four distinct nitrogenous bases, i.e., adenine (A), cytosine (C), guanine (G), and

uracil (U). As a messenger [CS14], it acts as a regulator for gene expression [WB08] or even as a catalyst [DC02] to complex biological operations. Recent studies reveal its contribution to functions not directly related to protein conformation [OM11, GKQ16].

Contrary to DNA, RNA features a single-stranded molecule resulting in a solid inclination of its bases to interact in two principal ways, either perpendicular to their planes (stacking) or hydrogen-bonded within the base planes (pairing). Those interactions form base pairs via hydrogen bonds between the corresponding nucleotide, i.e., the standard set of RNA base-pairs (AU, GC) known as Watson–Crick base-pairs [WC03] and the regularly appearing GU wobble-base pairs. On top of that, RNA bases also form bonds between their Hoogsteen- or CH-edge and their sugar-edge to allow "pairing" between three bases at once, known as base-triplets. In this setting, RNA molecules usually come as single strands, but they tend to fold themselves into what is known as their tertiary structure, which determines the spatial stability of its single-stranded helix. Simple linear strings of RNA form various complex three-dimensional structures due to the exact hydrogen-bonding mechanism stabilizing the well-known DNA double helix. Helices, also known as stems, are formed in the case of RNA in an intra-molecular way.

According to the literature, there is a strong correlation between the structure of a RNA molecule and its function [Edd02, Zuk00, RML13] to the extent of determination. Accurate RNA secondary-structure predictions highlight the RNA's functionality. For instance, according to work in [HSS+17], RNA secondary structure determines vital responsibility in the central nervous system that may be relevant to the aetiology of neurological disorders. For quite some time, researchers have been working on identifying and accurately measuring the structural elements of RNA to monitor their structure. To identify the RNA structure, chemical mapping has been proposed in works such as [KTS15]; however, these methods are not generalizable and often have biases in their reactivities. A compelling method widely used for measuring RNA structure in vitro is SHAPE [SCF+12] and its transformation icSHAPE [CFS17] for in vivo measuring. Other methods such as X-ray crystallography [Shi14] and nuclear magnetic resonance (NMR) are time-consuming and cost-inefficient. For all the experimental methods, the accuracy is restricted by the length of the RNA, i.e., the prediction accuracies are higher for shorter RNAs. Therefore, in RNAs of more considerable lengths, secondary-structure prediction provides a significant key to deciphering their potential functions.

The Pseudoknot Pattern

The pseudoknot pattern is one of the most-common RNA-folding motifs. It was first identified in the *Turnip Yellow Mosaic virus* [RVPP+82], consisting of two helical segments that are bound by at least two single-stranded sections or loops. Pseudoknots are met into various folding motifs. In general, there are four basic types [KHSQ16] of pseudoknots that have been distinguished (i) the H-type, (ii) the K-type, (iii) the L-type, and (iv) the M-type as shown in Figure 5.1.
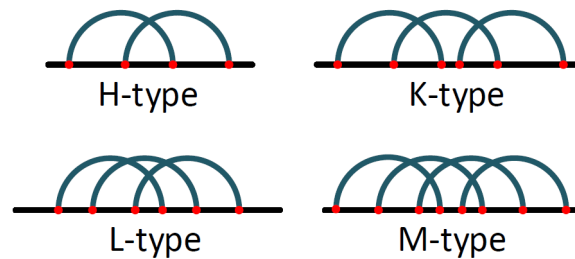
Figure 5.1: The four basic types of pseudoknots.

The most prevalent is the H-type [SB05], which is the one this work mainly focused on. In this pseudoknot pattern, the single-stranded RNA sequence folds into an "S" fashion to form two tangent loops, each one of arbitrary length. The two base pairs can be piled on top of each other to develop an essentially continuous helix. During the pseudoknot folding formation, the single-stranded loop parts consist instantaneously of bases, which urge to form hydrogen bonds with any adjacent "free" neighbor to contribute to the overall arrangement of the RNA molecule.

Even though a pseudoknot is a typical pattern, it is the springboard for remarkable yet robust RNA structures. Being a structurally diverse group varying in length and loops, and stems (the hydrogen-bonded base-pairs), pseudoknots are related with a plethora of biological operations such as holding the catalytic role of various ribozymes [RBOC96, KZD$^+$04], self-splicing introns [ASK$^+$04], and telomerase [TBF05]. Pseudoknots are even contributing, sometimes to the extent of definition, to the alternation of the gene expression of many viruses [STJ95, NRK$^+$02, MVV$^+$01].

5.2.2   Syntactic Pattern Recognition

The underlying model of the proposed predicting methodology of pseudoknots in RNA structures is that of syntactic pattern recognition. In syntactic pattern recognition, a language is defined as a set of syntax rules, which may construct a string belonging to that language [HU69]. The set of syntax rules is part of a grammar and determines the way accurate strings of symbols, which are components of the defined language, may be produced [HU69]. All grammars belong into four specific classes defined by Noam Chomsky [Cho56], which is acknowledged as the Chomsky hierarchy. Context-free grammars (CFG) are one of those four categories and are widely used for the implementation of programming languages and human-language recognition [Sip06].

Context Free Grammars

The formal definition of a CFG [ALSU06] is a group of four sets (quadruple), i.e., $CFG = \langle NT, T, R, S \rangle$. S ($S \in NT$) is the start symbol, which is also called the root of the grammar. T is a set that includes all terminal symbols, while NT includes all non-terminal symbols. All syntactic rules are contained in set R. The syntax rules follow the formalism $C \rightarrow \delta$, where $C \in NT$ and $\delta \in (T \cup NT)^*$, defining that symbol C, which is a non-terminal symbol, may be altered by δ.

Latin capital characters are used to represent the non-terminal symbols, while Latin lowercase characters are used to represent terminal symbols. Greek lowercase characters represent

strings of terminals and non-terminal symbols.

Parsing is the process that makes use of the syntax rules of a grammar, in order to produce a string, meaning it validates a string's syntactic exactitude. When a methodology solely resolves if a string of symbols may be produced by a grammar, then it forms a recognizer. In case a methodology, apart from recognizing a string, constructs a parse tree as well, then it is called parser. The root of a parse tree is the root symbol of the grammar; each node of the parse tree is formed by a syntax rule; and the leaves of the parse tree are terminal symbols forming the recognized string. A grammar is called ambiguous when a string can have more than one leftmost derivation or parse tree.

Primitive Pattern Selection

During the design process of an implementation based on syntactic pattern recognition, it is really vital to select the appropriate primitive patterns. In case of a RNA sequence that consists of the four key bases of adenine, guanine, cytosine, and uracil, the most-common case is to consider RNA as a string of symbols *a*, *g*, *c*, and *u* such as *auacggc* or *cugcaucccgcauauacg*. Consequently, the vocabulary of a grammar aiming to recognize strings representing RNA should contain only four terminal symbols T = (*a*, *g*, *c*, and *u*).

CFG Parsers

Due to the high expressive capability of CFG grammars, numerous parsing algorithms have been proposed for them. Two well-known CFG parsing algorithms are those proposed by Cocke, Younger, and Kasami (CYK) [You67] and by Earley [Ear70]. Based on the two algorithms mentioned above, several worth-mentioning extensions [GHR80, Ruz78, GXM⁺14] and parallel versions [PDK⁺09, CF84] exist in the literature as well.

Earley and CYK are algorithms of comparable complexity [GHR80], as both of them have adopted a similar dynamic programming method. Earley's parser was selected for the proposed implementation due to his efficiency and ability to handle ambiguous grammars.

Earley's Parsing Algorithm

The parsing algorithm for CFG grammars presented by Earley in 1970 constructs the parse tree using a top-down methodology. Earley's algorithm locates the dot symbol "$\bullet$" $\notin (N \cup NT)$, in each rule producing dotted rules. The existence of a dot in a rule indicates that the part of the rule left of the dot has been recognized, while the part of the rule right of the dot has not yet been recognized. In case a dot reaches the last position of a rule that has the root symbol at its left side, then the input string is considered recognized. This algorithm defines and applies operations, which are named Scanner, Predictor, and Completer. The input string $\alpha = a_1 a_2 a_3 \ldots a_n$ is traversed from a_1 to a_n. As each input symbol is scanned, a data set is constructed, representing the state of the recognition procedure at this place in the scan. Consequently, the algorithm builds $n + 1$ data sets of states. A state is simply a set of three integers $\{SR, p, F\}$. SR indicates the number of the rule; p is the position of symbol "$\bullet$"; and F is the enumeration of the set where the dotted rule was initially generated. A state in an Earley's data set is of form $i : {}_F Y \rightarrow \alpha \bullet Z \gamma$, meaning syntactic rule $Y \rightarrow \alpha Z \gamma$ having symbol "$\bullet$" at the pth position ($|\alpha| = p$), ~~was~~ initially generated at the Fth data set and is located in data set S_i. As the reading of the input symbols is moving

on, new data sets of dotted rules are generated. The three operations are sequentially applied to each dotted rule of all sets. The presence of a completed dotted rule having a root symbol at the left side of the rule in the last data set denotes the recognition of the input string.

The implementation of the proposed method is based on a Yet Another Early Parser (YAEP) parser [yae], which is one of the most efficient Early's parser implementations capable of parsing ambiguous grammars as well.

The Earley's parser algorithm is presented in Algorithm 1. In the main function EARLEY_PARSER, an array of sets containing states is initialized according to input sting length (INITIALIZE(input_string)) and adds a state having "$\bullet$" to the left side of start symbol S at the set with enumeration 0 (ADD_TO_SET((Start $\to$ $\bullet$ S, 0), Sets[0])). Then, a double-nested loop is executed. The nested loop examines each state in each set, and a set may expand during this loop as the three operations are adding states to sets. In each state, it is examined whether the right of "$\bullet$" is a nonterminal symbol, a terminal symbol, or the state is completed ("$\bullet$" is at the end of the rule). and the functions are called PREDICTOR, SCANNER, or COMPLETER, respectively.

In case function PREDICTOR is called, then for the nonterminal symbol that is right of "$\bullet$" (nonterminal symbol C in pseudocode as dotted rule is B $\to \alpha \bullet$ C β), all grammar rules are traversed to select the rules that have this symbol at the left side of the rule (C $\to \delta$). The selected rules are then added to this set after placing the "$\bullet$" at the first position of the right side of the rule (C $\to \bullet \delta$).

When function SCANNER is called, if the terminal symbol that is right of "$\bullet$" (terminal symbol a in pseudocode as dotted rule is B $\to \gamma \bullet$ a δ) is equal to the current examined symbol of the input sting (input_string[i]), this state is added to the next set after moving "$\bullet$" one position to the right (B $\to \gamma$ a $\bullet \delta$).

When function COMPLETER is called, the states in the set where the completed state (A $\to \delta$ $\bullet$) was initially generated (x in pseudocode) are traversed in order to select the states that have the symbol (A in pseudocode) at the left side of the rule, one position after the "$\bullet$" (B $\to \gamma \bullet$Aβ). These states are added in Sets[i] after moving the dot one position to the right (B $\to \gamma$ A $\bullet \beta$).

Algorithm 1 Earley's Parser Algorithm

```
DECLARE ARRAY_OF_STATES Sets;

function INITIALIZE(input_string)
    n ← LENGTH(input_string)
    Sets ← CREATE_ARRAY(n + 1)
    for i ← from 0 to n
        Sets[i] ← EMPTY_SET
    endfor

function EARLEY_PARSER(input_string, grammar)
    INITIALIZE(input_string)
    n ← LENGTH(input_string)
    ADD_TO_SET((Start → •S, 0), Sets[0])
    for i ← from 0 to n
        for each state in Sets[i]
            if (state is not completed)
                if (RIGHT_TO_DOT(state) is a nonterminal)
                    PREDICTOR(state, i, grammar)
                else
                    SCANNER(state, i, input_string)
                endif
            else
                COMPLETER(state, i)
            endif
        endfor
    endfor
    return Sets

function PREDICTOR((B → α•Cβ, j), i, grammar)
    for each (C → δ) in GRAMMAR_RULES
        ADD_STATE_TO_SET((C → •δ, i), Sets[i])
    endfor

function SCANNER((B → γ•aδ, j), i, input_string)
    if (a is input_string[i])
        ADD_STATE_TO_SET((B → γa•δ, j), Sets[i+1])
    endif

function COMPLETER((A → δ•, x), i)
    for each (B → γ•Aβ, j) in Sets[x]
        ADD_STATE_TO_SET((B → γA•β, j), Sets[i])
    endfor
```

5.3 Related work

Due to the complexity of the prediction of a tertiary representation and its significant computational cost, many studies focus on predicting the earlier stage of the secondary structure. Not only is the prediction of pseudoknotted RNA structures a challenge for bioinformatics but also the annotation of pseudoknots based on the secondary or tertiary structure of RNA molecule is not an easy task [APZ⁺18]. Spotting pseudoknots, specifically, is known to be a challenging puzzle, considering the short experimentally verified RNA-structures-to-sequences ratio. The majority of well-known algorithms make use of dynamic programming techniques, trying to predict the lowest free-energy structure, tackling the problem in a thermodynamic approach. Typical implementations that use thermodynamic models are RNAfold [LBHzS⁺11] and manifold [Zuk03], while others such as RNAalifold [BHW⁺08] utilize the ViennaRNA package [LBHzS⁺11] to calculate energy minimization. Although these methods are time-consuming, requiring an exponential amount of time relative to the input sequence length; that is, the problem is NP-complete [Aku00, LP00].

In this context, the bioinformatics industry adopts various approaches to overcome the efficiency barrier [LMT10]. Stochastic methods are applied to simulate folding pathways or to sample structures [VBGP95, IS00, MM07]. An update of the specific framework utilizes the folding pathway to locate free-energy structures by determining base pairs in a deterministic way [DFK07]. The second methodology relies on a heavily constrained dynamic programming approach. In that case, the predicted structures' possible topologies are limited based on certain criteria [RE99, DP03, RG04]. Another proposed alternative is to build structures iteratively or to even utilize graph-theory techniques. Such a worth-mentioning example is the nuclear magnetic resonance (NMR)-assisted prediction of the RNA secondary-structure (NAPSS) algorithm, which includes constraints from simple NMR experiments to improve predictions [TCGS98, WHS04, RSZ04, RRCH05]. Recently, other software tools—like RNAthor [JTK⁺20] and RNAProbe [WMB⁺20]—have appeared to facilitate the incorporation of experimental data into RNA structure prediction.

Many heuristic methods using different approaches have also been developed in the literature to overcome that computational barrier. Knotty [JWMW18] computes the secondary structure with a low runtime, using MFE prediction algorithm CCJ with sparsification. Knotty introduces a new class of structures called three-groups-of-band (TGB) and can predict a wide range of pseudoknots such as H-type pseudoknotted structures, kissing hairpins, and chains of four interleaved base pairs by overlaying TGB structures. Through the incorporation of sparsification, which improves space demands during the execution, keeping only a fraction of dynamic programming matrices, the overall need for memory remains significant. Next, ProbKnot [BM10] is a general secondary-structure prediction method that includes pseudoknots. It predicts base pair probabilities leveraging a partition function of any sub sequence not including pseudoknotted structures and then assembles a maximum-expected-accuracy structure from these probabilities without using dynamic programming. In that manner, it performs well in the structure prediction of pseudoknots and shows a fast execution time. A more-sophisticated variation of the latter, TheshKnot [ZZMH20], outperforms its results by discarding pairs with probabilities below a given threshold. Despite that, it performs the fastest prediction and scales almost linearly

to a sequence's length, and it sacrifices accuracy compared with specialized methods for pseudoknots. In particular, IPknot [SKH$^+$11] outperforms the systems mentioned above in terms of accuracy. It boosts the expected accuracy of a predicted structure using a thresholding variation of integer programming. It also approximates the base-pairing probabilities in order to decrease the inference time of prediction.

On the other hand, implementations with SCFGs have been proposed in the literature. These approaches present accuracy, which largely depends on the chosen SCFG that describes the secondary-structure prediction. A typical SCFGs prediction example is Pfold [KH99, KH03], which receives RNA alignment input to produce a consensus secondary structure of that as output. Additionally, a multithreaded version of Pfold, the PPfold [SKV$^+$11] , has been released. RNA-Decoder [PMF$^+$04] predicts the secondary structure of alignments using a SCFG and also taking into account the known protein-coding context of RNAs. Various implementations exist such as Contrafold [DWB06], Evfold [PBS$^+$06], Infernal [NKE09], Oxfold [AHM$^+$13], Stemloc [BPH08], TRNAnscan-SE [LE97], Xrate [KUB$^+$06], etc., all exploiting the SCFG model. All of the above implementations are software ones, while only two implementations are hardware ones. The one in reference [XDZL10] was designed and executed on a field programmable gate array (FPGA), and the other in reference [CKO10] was executed on a GPU using a CUDA [Nvi07] implementation of the Nussinov algorithm [NPGK78]. The two different approaches of thermodynamic models and SCFGs are more alike than different in some ways. It has been shown that Zuker's thermodynamic model can be translated to a SCFG, by calculating the probabilities of productions from the thermodynamic constants [IS00]. SCFG-based as well as thermodynamic approaches aim to the optimization of an objective function; thermodynamic methods try to minimize free energy of a structure, while SCFG methods are dealing with the maximization of the corresponding probability. In these approaches, optimization lies on recursion relations and is resolved using dynamic programming techniques, leading to a computational complexity of $O(n^3)$. Nevertheless, these approaches are diametrically opposite regarding the scientific concepts and assumptions used. Thermodynamic methods incorporate a biologically oriented, energy-driven model for RNA folding and obtain their parameters from experiments on specific short RNA molecules. SCFG-based methods, on the other hand, pursue a machine-learning orientation, by targeting on modeling the complete structures observed in nature. Afterwards, these systems reproduce similar structures, based on patterns and detected similarities. Considering that, it is obvious that SCFG-based prediction is inherently probabilistic, leveraging the advantage of probability and statistics as the background. They can also be combined with other models within the same probabilistic context to become more efficient and to improve their predictions.

Recent research also suggests the utilization of pure machine-learning approaches towards the prediction of RNA secondary structure. In [SHPZ19], the authors propose using deep contextual learning for base-pair prediction, including those non-canonical and non-nested (pseudoknot) base pairs stabilized by tertiary interactions. However, the lack of sufficiently large datasets may question the quality of such deep-learning methodologies. In [WLZ$^+$19], the proposed deep-learning framework DMfold predicts the secondary structure of RNA sequences, including pseudoknots. DMfold consists of a bidirectional-LSTM network as an encoder and a fully connected layer as a decoder. The system predicts an initial dot-bracket representation for each RNA sequence given as input, using the encoder−decoder framework. Afterwards, DMfold applies the

improved base pair maximization principle (IBPMP) to select the base pairs in the dot-bracket sequence and create three pseudoknot-free substructures, which in turn, are combined to calculate the secondary structures with pseudoknots. Inspired by DMfold, 2dRNA [KJY20] proposes a coupled two-staged deep neural network, leveraging the advantages of a bidirectional LSTM with a U-net architecture. In the first stage, the two-level bidirectional LSTM encodes sequence information in higher dimensions, while a fully connected network decodes that data and predicts the dot-bracket representation. This procedure consists of the coarse-grained dot-bracket prediction (CGDBP). The second stage, called fine-grained dot-plot prediction (FGDPP), feeds that representation to a fully convolutional network, which constructs a dot-plot matrix. However, the output shows mismatches between brackets because of the inherent ability of LSTM to reveal sequential information. This problem is countered by introducing a U-net architecture, which receives that structure and predicts base-pairing, providing at the same time significant structural information. In that same context, a recent approach, ATTfold [WLW+20], predicts the secondary structure of RNA with pseudoknots. The framework utilizes deep-learning techniques based on an attention mechanism. It calculates the base-pairing score matrix via an encoder with an attention mechanism and a convolutional neural network as the decoder. Finally, the resulted matrix is enforced to comply with the hard constraints of RNA folding, and the overall architecture is trained with respect to those biological restrictions.

5.4 Overview of our approach - An Illustrative example

5.4.1 The Proposed Methodology

In the current section, an overview of the proposed methodology will be presented. The procedure of RNA pseudoknots recognition is split into the following three tasks: (i) RNA sequence is parsed using a CFG parser in order all trees that include a pseudoknot to be produced; (ii) all derived trees are then traversed to identify additional base pairs around the pseudoknot; and (iii) the optimal tree is selected via the criteria of minimum energy and the maximum number of base pairs of the pseudoknot. These three main tasks of the proposed methodology (see Figure 5.2) are thoroughly described in Sections 5.4.1, 5.4.1 and 5.4.1, respectively.
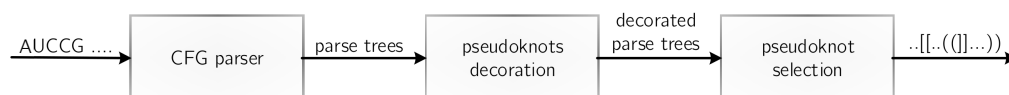
Figure 5.2: Tasks of the proposed methodology.

As exhibited in Figure 5.2, the presented implementation given an input RNA in the form of a string representing a sequence of nitrogenous bases produces the base pairing of the given string in extended dot-bracket notation. A separate software module was developed to implement each task, and all the implementation details are described in Section 5.5. A more-extensive representation of our approach is shown in Figure 5.3.
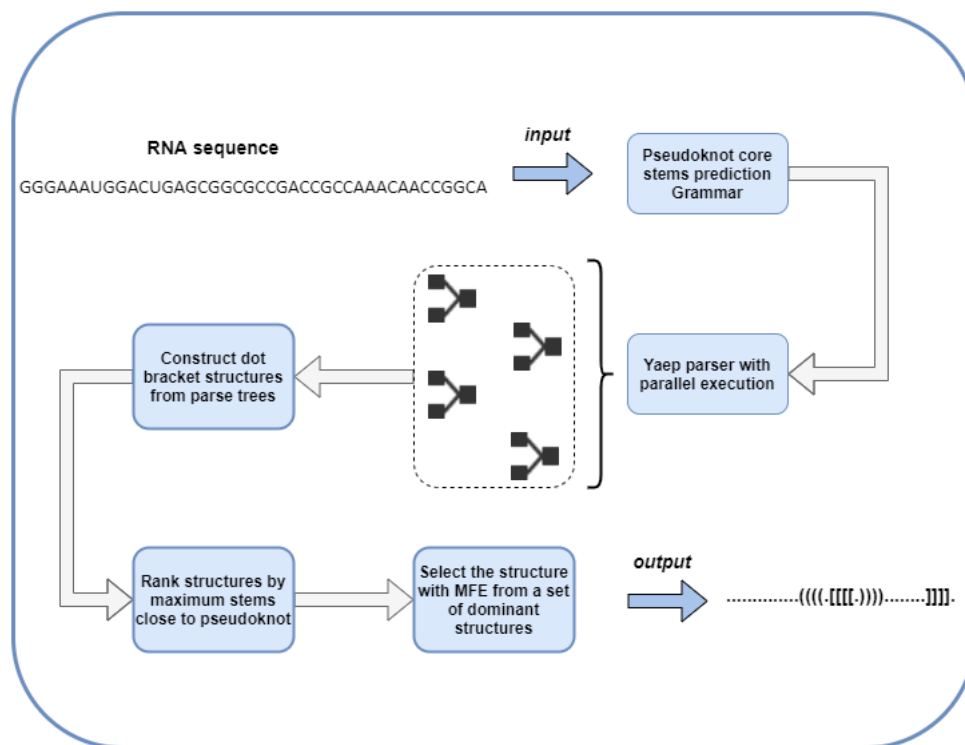
Figure 5.3: A more extensive representation of our approach.

CFG to Identify Pseudoknots

The proposed methodology of detecting pseudoknots in sequences of nitrogenous bases representing RNA rests on syntactic-pattern-recognition techniques and specifically on an efficient CFG parser. Consequently, it is vital to select the right primitive patterns. In the case of RNA recognition, the most-typical option is to represent the nitrogenous bases adenine, cytosine, guanine, and uracil as single characters "A," "C," "G," and "U," respectively. These characters in sequence constitute an RNA representation. Hence, in our case, where a grammar parser is proposed to recognize pseudoknots in RNA, the proposed grammar vocabulary contains only the four terminal symbols T = {"A," "C," "G," "U"}, with each one representing a distinct base: adenine, cytosine, guanine, and uracil, respectively. Therefore, every RNA sequence may linguistically be represented as a string containing the terminal symbols, e.g., UAGGC or AUGGCCGUACG.

The task to syntactically recognize a given pattern may actually be converted into using an appropriate pattern grammar, in order to parse the linguistic representation of the original patterns. The design of the pattern grammar may have a significant impact on the recognition's result. Therefore, the formation of the CFG to be used is an important subtask in implementation having as the underlying model syntactic-pattern-recognition techniques. Hence, the design of an efficient grammar is indispensable in order to describe the syntax of the pseudoknot within any arbitrary RNA sequence. It is well known that CFGs are adequate to represent structural features. The G_{RNA} shown in Table 5.1 is utilized to recognize pseudoknots in RNA.

Table 5.1: Description of AG_{RNA}.

#	Syntactic Rules
0	S → "A" L "A" D "U" L "U"
1	S → "U" L "A" D "A" L "U"
2	S → "C" L "A" D "G" L "U"
3	S → "G" L "A" D "C" L "U"
4	S → "A" L "U" D "U" L "A"
5	S → "U" L "U" D "A" L "A"
6	S → "C" L "U" D "G" L "A"
7	S → "G" L "U" D "C" L "A"
8	S → "A" L "C" D "U" L "G"
9	S → "U" L "C" D "A" L "G"
10	S → "C" L "C" D "G" L "G"
11	S → "G" L "C" D "C" L "G"
12	S → "A" L "G" D "U" L "C"
13	S → "U" L "G" D "A" L "C"
14	S → "C" L "G" D "G" L "C"
15	S → "G" L "G" D "C" L "C"
16	L → "A" L
17	L → "U" L
18	L → "C" L
19	L → "G" L
20	L → "A"
21	L → "U"
22	L → "C"
23	L → "G"
24	D → K N
25	K → "A"
26	K → "U"
27	K → "C"
28	K → "G"
29	K → ϵ
30	N → "A"
31	N → "u"
32	N → "C"
33	N → "G"
34	N → ϵ

The second column of Table 5.1 highlights all the grammar's syntactic rules. G_{RNA} contains the five non-terminal symbols of set NT = {S, L, D, K, N}. S is the start symbol; all syntactic rules having S on their left side, e.g., rule 0 to rule 15, aim to detect a possible pseudoknot in the

input string. A pseudoknot consists of at least two base pairs in which half of one base pair is intercalated between the two halves of another base pair. For instance, rule 6: S → "C" L "U" D "G" L "A" specifies the existence of a pseudoknot of the form C..U..G..A where the base pairs C–G and U–A are intercalated. These base pairs for the rest of the article will be mentioned as **core stems**. Figure 5.4 depicts the core stems C–G and U–A of this example, while half of base pair U–A is intercalated between base pair C–G, i.e., base U is between base pair C–G. Base G belonging in base pair C–G is also intercalated between base pair U–A—that is, the paradigm of the detected interference, leading to the prediction of the pseudoknot.
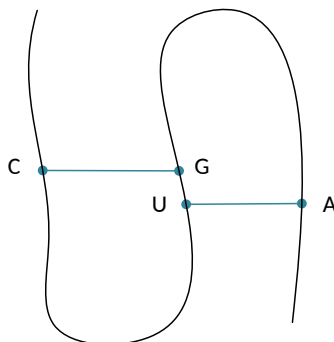
Figure 5.4: Pseudoknot detected by rule S → "C" L "U" D "G" L "A."

L is the non-terminal symbol that will produce sequences of bases forming the two interior loops of the pseudoknot, i.e., sequences of bases between C and U as well as between G and A. Non-terminal L may produce strings belonging to set $(T)^* \neq \emptyset$, where T is the set of terminal symbols, and $\emptyset$ is the empty set. Hence, L may produce strings of length greater than zero as A, UA, CCGGAU, etc.

D is the non-terminal symbol that will produce sequences of bases between the two crossing base pairs, i.e., between bases U and G in this example. Using non-terminal symbols K and N, L may recognize substrings of terminal symbols of length zero to two, i.e., ϵ, A, U, C, G, AU, UA, AC, CA, etc. where ϵ is the empty string. The length of the sub-string between crossing base pairs may easily be extended to more than two via simple grammar modifications. The maximum length of this sub-string may be defined by the user as explained in the development environment section at knotify github repo [cod].

G_{RNA} may detect pseudoknots in strings where the first and last symbols of the sequence belong to the core stems group. In the examined example of a pseudoknot detected by the sixth rule, the pseudoknot exists in a substring starting from terminal symbol C and ending with terminal symbol A. However, that should not be considered a limitation since the parser is extensively executed in subparts of the strings using the sliding-windows technique.

The parse tree produced by the parsing of substring "C G C C U G A U U U G A" is shown in Figure 5.5. Following the previous example, syntax rule 6 was used to detect the pseudoknot of the form **C** ...**U G** ......**A**. Then, rules 19, 18, and 22 were used to recognize the bases between C and U, i.e., C **G C C** U G ......A. After that, rules 24, 29, and 34 were used to recognize the empty string between the U G bases of the core stems. Finally, rules 16, 17, 17, 17, and 23 were used to recognize the bases between C and U, i.e., C G C C U G **A U U U G** A. The integration

of this substring in the initial RNA sequence and the process of decorating the pseudoknot with additional base pairs is explained in Section 5.4.1.

The proposed methodology parses all substrings, beginning with the one that starts with the first sequence symbol and features the minimum potential length. Iteratively, the length is extended by one symbol to include the entire initial RNA sequence finally. In the same iterative fashion, string starting points are augmented to exclude the previous set starting symbol. The parsing is over when the substring's length to be parsed deteriorates further than a predefined threshold (i.e., the minimum length of the pseudoknot). This methodology, considering that G_{RNA} is ambiguous, leads to the creation of a considerable number of parse trees. The selection methodology of the optimal tree is analyzed in Section 5.4.1. The CFG parser selected is that of YAEP [yae], which is a highly efficient CFG parser based on Earley's algorithm [Ear70], and according to the literature, it can handle ambiguity in grammars.

Context-free grammar was selected with the view of augmenting it with attributes (forming an attribute grammar) in order to store probabilities and to manage to prune parse trees during the parse-tree construction process, in future work. In order to enhance the performance of the proposed system, an alternative implementation of the first task was proposed using a brute-force algorithm. This approach traverses the input string in order to spot all possible base pairs and then traverses all base pairs so as to identify couples of base pairs, which potentially form the core stems of a pseudoknot, i.e., consisting of two helical segments that are bound by two single-stranded sections or loops. The brute-force approach achieved a performance speed of 2.55 compared to the implementation that makes use of the context-free grammar at the first task. The speedup, though, comes at the cost of the overall extensibility that the grammar approach provides. In any case, both proposed implementations are faster than two well-known platforms [JWMW18, SKH$^+$11]. Performance evaluation is analyzed in Section 5.6.
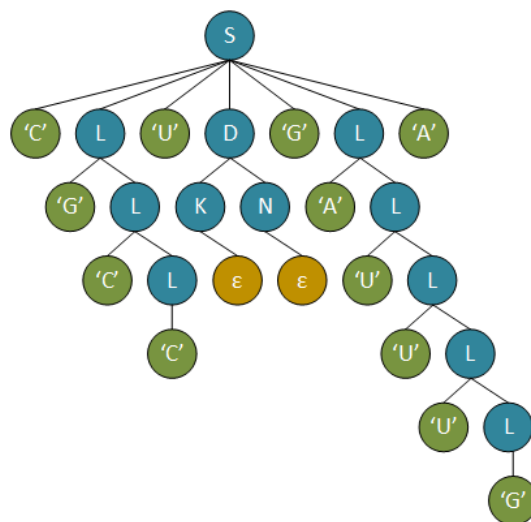
Figure 5.5: Parse-tree-recognizing pseudoknot is substring "C G C C U G A U U U G A."

Decorate Core Stems

Once the parse trees are generated following the methodology described in the previous subsection, all trees are traversed to decorate the pseudoknot with additional base pairs. G_{RNA} is designed to identify only the pseudoknot's core stems to enhance the CFG parser's performance. This fact leads to a CFG with few syntactic rules and enables the parser to behave efficiently. However, the drawback is that all parse trees should be traversed to detect the rest of the base pairs surrounding and framing the pseudoknot's core stems. All bases in each of the two pseudoknot loops are sequentially examined if they may form a base pair with another base located in a suitable position. Table 5.2 exhibits the part of the algorithm called **decoration**. Having detected the core stems U–A and C–G at positions 9 and 11 and 5 and 10, respectively, all bases in both pseudoknot loops, (i.e., bases at positions 6 to 9 (left loop) and 11 to 15 (right loop)) were examined to identify against forming base pairs with bases outside the pseudoknot loops. Hence, base pairs belonging in the left loop were checked out if they may form base pairs with bases at positions 17 to 19, while bases belonging in the right loop were examined to match the ones at positions 1 to 4. In both the left and the right loops of the pseudoknot, the base pairs at positions 8–17, 7–18 and 4–11, and 3–12 were sequentially detected, respectively. Table 5.2 presents the entire procedure in detail.

Table 5.2: Decoration of core stems of pseudoknot.

String enumeration	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
String	C	C	A	U	C	G	C	C	U	G	A	U	U	U	G	A	G	G	A
Parser output:	.	.	.	.	[	.	.	.	(	]	.	.	.	.	.	)	.	.	.
Step 1	.	.	.	.	[	.	.	(	(	]	.	.	.	.	.	)	)	.	.
Step 2	.	.	.	.	[	.	(	(	(	]	.	.	.	.	.	)	)	)	.
Step 3	.	.	.	[	[	.	(	(	(	]	]	.	.	.	.	)	)	)	.
step 4	.	.	[	[	[	.	(	(	(	]	]	]	.	.	.	)	)	)	.

The proposed system permits the user to optionally select the existence of base pairs U–G in the pseudoknot's loops. At the same time, the existence of bulges or interior loops in the loops of the pseudoknot is part of our future work.

Optimal Tree Selection

According to the literature, multiple methodologies have been presented that tackle the problem of RNA base-pair prediction, with the prevalent ones being the (i) method of minimum free energy [E.14], which detects the RNA sequence that features the lowest amount of free energy. It is synonymous with the natural-mode structure, but it is not necessarily the structure that forms in nature. The perception of minimum free energy is basically a restatement of the second law of thermodynamics. (ii) The method of maximum pairing [NJ80] is a technique based on the number of base pairs formed around the pseudoknot's core stems. The dot-bracket notation with the maximum number of base pairs around the pseudoknot will probably lead to the minimum free energy. Next to the row, the (iii) method of partition function [H.04] is founded on the fact

that the actual base pairs should entitle a high base-pairing likelihood in the estimated minimum free energy distribution. The method boosts any positive predictive value of the actual base pairs by considering their nearest neighbors' parameters for formed free energy at a given temperature. Finally, the (iv) method of comparative sequence analysis [RE01] is about testing the pattern of substitutions observed in a pairwise alignment of two homologous sequences, while the (v) method of ohysical experiments [CR12] focuses on providing insights through actual wet experiments.

The proposed system employs a hybrid model of optimal tree selection, combining principles originating from the two most-prevalent techniques, i.e., the method of maximum pairing and that of MFE, to predict the pseudoknot pattern of any RNA secondary structure accurately yet in an efficient way. MFE is cost-effective in terms of performance; initially, all trees are sorted by the number of base pairs around the detected pseudoknot, and MFE applies only to the ones ranking the top score of the base-pairs count. This is a heuristic that outperforms the MFE original approach.

Minimum-Free-Energy Calculation

In order to select the best candidate from the set of secondary structures, our method chooses the one with the minimum free energy. To carry out this important task, a module from HotKnots [RRCH05] was incorporated, to compute the energy of each structure and then supply it to our framework for the final selection. This module is based on an algorithm introduced by Mathews [MSZT99], which has been extended for pseudoknots by Dirks [DP03]. Specifically, the energy of the pseudoknot is given by the relation below:

$$G^{pseudo} = \beta_1 + \beta_2 * B^p + \beta_3 * U^p \tag{5.1}$$

where β_1 is the weight for the existence of pseudoknot; B^p is the number of core stems; and U^p is the number of unpaired bases inside the pseudoknot. The parameters β_2 and β_3 were set to 0.1, as computed experimentally in [RRCH05], and refer to the core stems and unpaired bases, respectively. The β_1 weight was set to 9.6. Figure 5.6 provides an illustrative example for the weights-costs β_1, β_2, and β_3 in an indicative $H - type$ pseudoknot.
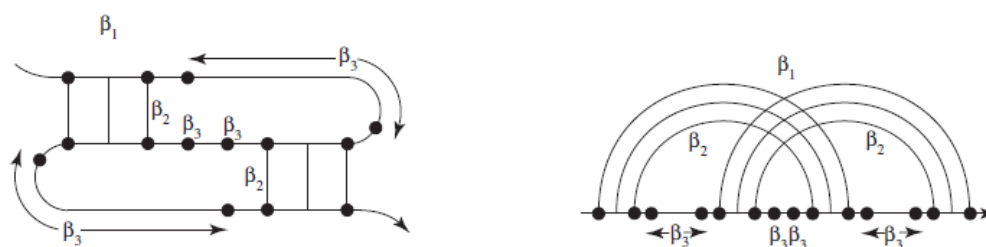
Figure 5.6: The cost of forming a pseudoknot is β_1, while core stems contribute a cost β_2 and unpaired bases inside the pseudoknot a cost β_3. The energies associated with the stacked base pairs were computed with respect to the standard model [MSZT99] (after [RRCH05]).

5.5 Materials and methods

Implementation Details

According to the literature, the prediction of pseudoknots of any arbitrary RNA sequence is a NP-complete problem. On the one hand, free-energy-minimization algorithms proposed to provide pseudoknot predictions invoke dynamic programming to rank high regarding their computational cost, while their precision decreases proportionally to the length of the input sequence. On the other hand, existing heuristic approaches lack generalization capabilities when being tested under different datasets. In this setting, we introduced a novel hybrid strategy to pick the RNA sub-sequence that is the most-probable pseudoknot expression. According to Figure 5.3, the proposed methodology initially creates a sub-space of all potential pseudoknot expressions, i.e., successfully parsed trees that satisfy some minimum length criteria and then solves a global defined optimization problem by picking the pseudoknot representation that features (i) the maximum number of base pairs around the pseudoknot and (ii) the minimum free energy.

The proposed implementation is hybrid itself as well. By invoking Python and C code routines, a performant, scalable yet easy-to-use, and extended software package was implemented. Python was used to provide high-level flexibility and out-of-the box features such as flexible parallelization capabilities, sub-process supervision, and file management, while C nailed the parsing task by minimizing the memory footprint and maximizing CPU utilization.

The input sequence was sliced into multiple sub-sequences (this procedure is also described in Section 5.4.1). Since the tasks of parsing the derived sub-sequences are orthogonal to each other, we can parallelize the particular workload. A pool of tasks is spawned to constitute a parallel CFG parser that evaluates all the produced sub-sequences in a completely parallel fashion. The size of the pool is proportional to the CPU logical cores to max out CPU utilization, while every task is a YAEP-parser [yae] instance implemented in C to guarantee optimal resource allocation and blazing-fast parsing (Figure 5.7).

Each CFG parser instance produces a pseudoknot structure that describes each potential pseudoknot within the CFG domain. If some parsed sub-sequences do not represent a pseudoknot, the CFG parser will fail, resulting in no pseudoknot structure. Next, all pseudoknot structures are serialized to a CSV format to be efficiently analyzed through the Pandas package [M$^+$11]. Given that the resulted data frame consists of all potential solutions to our problem (i.e., valid pseudoknots), we should pick the most likely one. Assuming that the most-suitable prediction may be the one that affords the least free energy, we ended up solving a minimum-free-energy optimization problem. However, calculating the free energy for every single potential RNA folding is a computational- and memory-intensive task that sets our entire pipeline as relatively inefficient. Our methodology tackles this highly computationally intensive task by adapting the observation that free-energy minimization is directly associated with the maximization of the base-pairs count of any potential RNA folding. Consequently, instead of calculating the minimum free energy for all pseudoknot structures, a maximum stem count look-up that features O(n) time and O(n) space complexity proportional to the input sequence's length was performed.

As mentioned in Section 5.2, the first task, that of pseudoknot core stems prediction, was accomplished by two different approaches: one based on the YAEP parser (knotify_yaep) and a second based on a brute-force algorithm (knotify_bruteforce). The first implementation needs
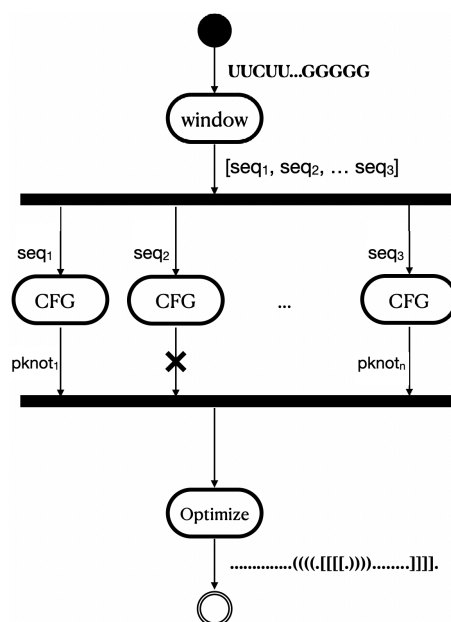
Figure 5.7: Pipeline parallelization.

$O(n^5)$ time: the complexity of the Earley parser [Ear70] is for ambiguous grammars plus the complexity to traverse all direct acyclic graphs (DAG), and the YAEP parser produces as compact representation of all possible parse trees of an ambiguous grammar. On the other hand, the second one needs $O(n^2) + O(n^4) \approx O(n^4)$: $O(n^2)$ to traverse the input string in order to spot all possible base pairs (the maximum number of base pairs is n^2) and then $O(n^4)$ to traverse all base pairs so as to identify couples of base pairs that may form the core stems of a pseudoknot. The implementation source code is publicly available under the *knotify* github repo [cod].

5.6 Performance Evaluation

5.6.1 Dataset Presentation

A dataset [dat] of 262 RNA sequences was used to evaluate our methodology's accuracy against other methodologies. It is composed of well-known RNA sequences; thus, it should be considered a perfect fit to compare our methodology against other highly respected implementations proposed in the literature, i.e., Hotknots, Iterative HFold (IHFold), IPknot, and Knotty [RRCH05, SKH$^+$11, JWMW18]. The dataset of 262 RNA sequences was divided into four groups regarding their length. Consequently, there was a group of 75 RNA sequences of length smaller than 30, a group of 68 RNA sequences of length greater equal than 30 and smaller than 40, a group of 55 RNA sequences of length grater equal than 40 and smaller than 50, and a group of 64 RNA sequences of length grater equal than 50. The above-mentioned groups are notated as L<30 (#75), 30<=L<40 (#68), 40<=L<50 (#55), and L>=50 (#64), respectively, in the Tables and Figures of this section. For all methods of evaluation, the results are presented for the entire dataset and per groups.

5.6.2 Methods of Evaluation

To evaluate our methodology, we decided to go with three metrics: we asserted i) the accuracy of the pseudoknot's core stems prediction, ii) the capability to predict the base pairs existing in the ground truth dot-bracket (confusion matrix), and iii) the execution time.

Predicting Pseudoknot location

Table 5.3 provides a compact comparison among our solution and the aforementioned platforms, by summarizing the capability of predicting the core stems of the pseudoknot. The comparison was made against both methodologies proposed in this article, i.e., knotify_yaep and knotify_bruteforce (see Sections 5.4 and 5.5). Our methodology succeeded in perfectly detecting the core stems of the pseudoknot in 143 out of 262 sequences, while Knotty in 121 sequences, Hot-Knots in 75, IPknot in 38 sequences, and IHFold in 0 sequences. Towards the calculation of core stems' location, we allowed one base of each pair to be located one position on the right or left, i.e., pair *(i,j)* is equivalent to *(i-1,j)*, *(i+1,j)*, *(i,j-1)*, and *(i,j+1)*, as proposed in [MSZT99].

Table 5.3: Predicting pseudoknot location in entire dataset.

Platform	Exact Matches	Exact Matches (%)
IHFold	0	0
HotKnots	75	28.6
IPknot	38	14.5
Knotty	121	46.1
knotify_yaep	143	54.5
knotify_bruteforce	144	54.9

The measurements divided per RNA sequences' length are presented in Table 5.4, where it is shown that our methodology succeeded in predicting exactly the core stems in more pseudoknots compared to the other implementations in three out of four groups, while in the groups where the length is between 30 and 40, our methodology predicted one less than Knotty.

Table 5.4: Predicting pseudoknot location per RNA sequence length.

Platform	L < 30 (#75)		30 <= L < 40 (#68)		40 <= L <50 (#55)		L >= 50 (#64)	
	Exact Matches	Exact Matches (%)	Exact Matches	Exact Matches (%)	Exact Matches	Exact Matches (%)	Exact Matches	Exact Matches (%)
IHFold	0	0.00	0	0.00	0	0.00	0	0.00
Hotknots	20	26.67	18	26.47	13	23.64	24	37.5
IPknot	20	26.67	7	10.29	4	7.27	7	10.94
Knotty	52	69.33	28	41.18	13	23.64	28	43.75
knotify_yaep	63	84.00	27	39.71	24	43.64	29	45.31
knotify_bruteforce	62	82.67	27	39.71	26	47.27	29	45.31

The percentage of pseudoknot's core stems exact prediction per platform is also shown in Figure 5.8 while 5.9 depicts the per platform prediction accuracy for all groups of length of RNA sequences
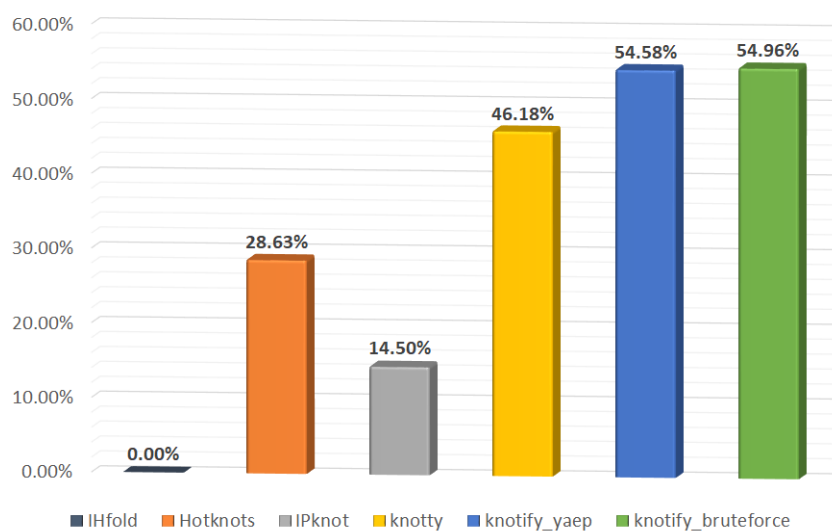
Figure 5.8: Percentage of pseudoknot's core stems exact prediction per platform.
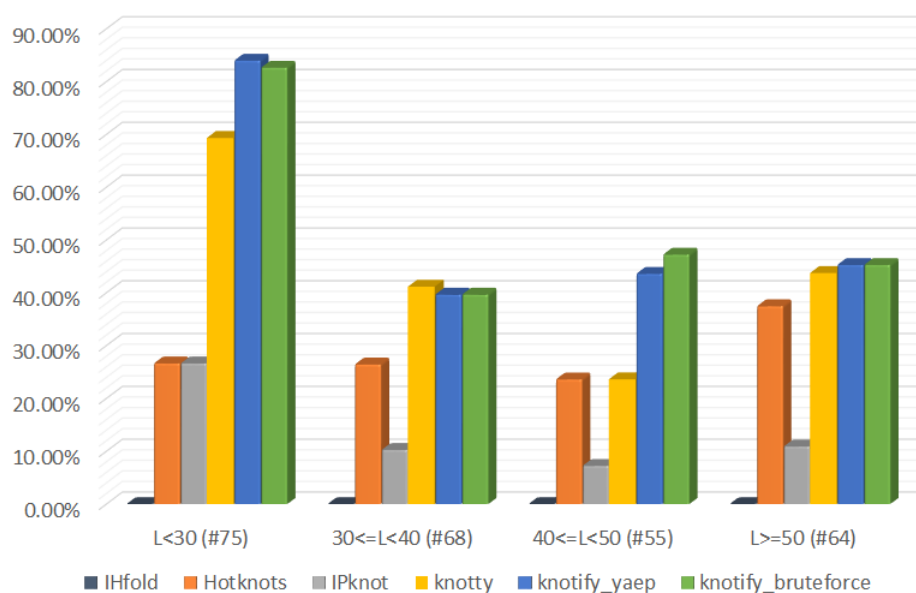
Figure 5.9: Percentage of pseudoknot's core stems exact prediction per platform and sequence length.

Confusion Matrix

We benchmarked our framework on the sequences of Table 5.5 along with the same state-of-the-art methods IHFold, HotKnots, IPknot, and Knotty. Table 5.5 presents the performance for each method in terms of the positive predicted value (PPV), the recall, the F1-score, and the Matthews correlation coefficient (MCC). Equations (5.2)–(5.5) provide the definitions, where TP refers to the number of correctly predicted base pairs, FP to the number of incorrectly predicted base pairs, FN to the number of base pairs that were not predicted, and TN to the number of the bases that were not correctly predicted from the system.

$$PPV = \frac{TP}{TP + FP} \tag{5.2}$$

$$Recall = \frac{TP}{TP + FN} \tag{5.3}$$

$$F1 - score = \frac{2 \times PPV \times Recall}{PPV + Recall} \tag{5.4}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5.5}$$

To evaluate the overall performance, we focused on the precision, the MCC, and the F1-score. The latter is the harmonic mean of the PPV and the recall. Our methodology outperformed on average all methods in regards to precision metric having 0.784, while Knotty was 0.729, IPknot 0.718, Hotknots 0.706, and IHFold 0.608. Regardingthe F1-score and the MCC metrics, Knotty outperformed on average all methods with a F1-score equal to 0.807 and a MCC equal to 0.569. The proposed methodology had an accuracy very close to Knotty (F1-score = 0.774; MCC = 0.543). In addition, HotKnots scored F1-score = 0.738 and MCC = 0.452, while IPknot (F1-score = 0.712; MCC = 0.418) and IHFold (F1-score = 0.595; MCC = 0.226) had lower accuracy on both the F1-score and the MCC.

Table 5.5: Precision, recall, F1-score, and MCC per platform in entire dataset.

Platform	tp	tn	fp	fn	Precision	Recall	F1-score	MCC
IHFold	3056	3556	1968	2196	0.608	0.582	0.595	0.226
Hotknots	4180	3632	1744	1220	0.706	0.774	0.738	0.452
IPknot	3872	3767	1522	1615	0.718	0.706	0.712	0.418
Knotty	5026	3352	1870	528	0.729	0.905	0.807	0.569
knotify_yaep	4212	4102	1162	1300	0.784	0.764	0.774	0.543
knotify_bruteforce	4214	4101	1160	1301	0.784	0.764	0.774	0.543

The above-mentioned results are also shown in Figure 5.10.
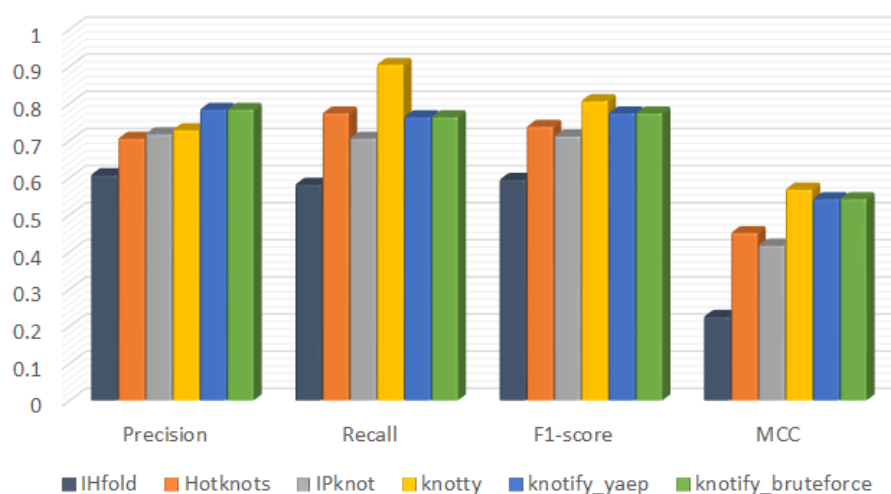
Figure 5.10: Precision, recall, F1-score, and MCC per platform.

In Tables 5.6–5.9, the precision, the recall, the F1-score, and the MCC metrics per platform are exhibited for the four groups of different RNA sequences' length. In these Tables it is shown that our methodology outperformed on average all methods in regards to the precision metric for all ranges of length, while Knotty outperformed our methodology in regards to the F1-score and MCC metrics mainly when RNA sequences were of larger size. When the length was smaller than 30, our methodology had a higher F1-score and MCC than Knotty. As shown in Table 5.3, our methodology was more accurate in predicting the core stems of the pseudoknot. The increased MCC of the Knotty platform in larger RNA sequences probably correlates with the fact that larger RNA sequences incorporate multiple structures (i.e., hairpins) that do not relate explicitly to the psuedoknot one. The latter possibly augmented the overall true-positive (tp) score. One of our future work tasks is to enhance our methodology to cover even more complex patterns such as pseudoknots enclosing bulges or hairpins.

Table 5.6: Precision, Recall, F1-score, and MCC per platform for sequences of length < 30.

Platform	tp	tn	fp	fn	Precision	Recall	F1-score	MCC
IHFold	738	522	118	513	0.862	0.590	0.701	0.386
Hotknots	904	492	156	339	0.853	0.727	0.785	0.465
IPknot	916	514	124	337	0.881	0.731	0.799	0.510
Knotty	1196	469	146	80	0.891	0.937	0.914	0.722
knotify_yaep	1244	486	134	27	0.903	0.979	0.939	0.805
knotify_bruteforce	1242	485	136	28	0.901	0.978	0.938	0.802

Table 5.7: Precision, recall, F1-score, and MCC per platform for sequences of length >= 30 and < 40.

Platform	tp	tn	fp	fn	Precision	Recall	F1-score	MCC
IHFold	550	832	352	587	0.610	0.484	0.539	0.191
Hotknots	922	851	294	254	0.758	0.784	0.771	0.528
IPknot	824	823	314	360	0.724	0.696	0.710	0.420
Knotty	1078	802	324	117	0.769	0.902	0.830	0.628
knotify_yaep	988	893	296	144	0.769	0.873	0.818	0.627
knotify_bruteforce	988	893	296	144	0.769	0.873	0.818	0.627

Table 5.8: Precision, recall, F1-score, and MCC per platform for sequences of length >= 40 and < 50.

Platform	tp	tn	fp	fn	Precision	Recall	F1-score	MCC
IHFold	612	864	478	418	0.561	0.594	0.577	0.237
Hotknots	792	857	510	213	0.608	0.788	0.687	0.412
IPknot	764	911	410	287	0.651	0.727	0.687	0.414
Knotty	904	817	524	127	0.633	0.877	0.735	0.492
knotify_yaep	764	1010	298	300	0.719	0.718	0.719	0.490
knotify_bruteforce	772	1012	290	298	0.727	0.721	0.724	0.499

Table 5.9: Precision, recall, F1-score, and MCC per platform for sequences of length >= 50.

Platform	tp	tn	fp	fn	Precision	Recall	F1-score	MCC
IHFold	1156	1338	1020	678	0.531	0.63	0.577	0.196
Hotknots	1562	1432	784	414	0.666	0.790	0.723	0.439
IPknot	1368	1519	674	631	0.670	0.684	0.677	0.377
Knotty	1848	1264	876	204	0.678	0.901	0.774	0.515
knotify_yaep	1216	1713	434	829	0.737	0.595	0.658	0.402
knotify_bruteforce	1212	1711	438	831	0.735	0.593	0.656	0.398

Execution-Time Comparison

The last metric that was used in order to compare the proposed methodology with other platforms is that of execution time. In Table 5.10, the execution time required per platform to predict an existing pseudoknot in RNA sequences is provided. The third column of this depicts the total execution time required by each platform to analyze all 262 RNA sequences, while the second column depicts the average execution time per RNA sequence. Our methodology outperformed Knotty, which had worse results regarding the core stems prediction and the precision but better results regarding the F1-score and MCC. Knotify_bruteforce required 33.894 secs; knotify_yaep required 85.756 secs; and Knotty required 263.303 secs. The methodology we introduced achieved

a speed of 7.76 (1.004/0.129) compared to the Knoty platform. The IPknot and Hotknots per-
formed at values of 3.45 and 1.31, respectively. Finally, IHFold recorded the lowest execution
time; nonetheless, it had the poorest accuracy-evaluation profile.

Table 5.10: Execution time required per platform in entire dataset.

Platform	Average Time (sec)	Total Time (sec)
IHFold	0.030	8.096
Hotknots	0.169	44.432
IPknot	0.447	117.246
Knotty	1.004	263.303
knotify_yaep	0.327	85.756
knotify_bruteforce	0.129	33.894

The execution time required per platform is also shown in Figure 5.11.
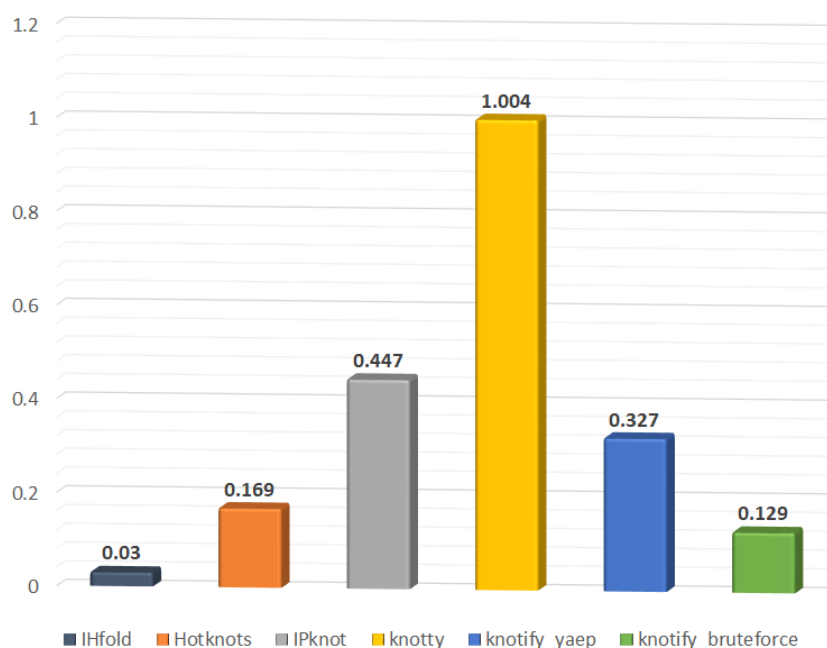
Figure 5.11: Average execution time (sec) required per platform.

In Tables 5.11–5.14, the average and total execution time per platform for the four groups
of different RNA sequences' length are shown. It is worth noting that the execution time of the
Knotty platform increased significantly in proportion to the length of the input RNA sequence,
while Hotknots, knotify_yaep, and knotify_bruteforce seemed to increase similarly as the length
of the RNA sequence became larger, keeping a quite steady ratio.

Table 5.11: Execution time required per platform in for RNA sequences of length < 30.

Platform	Average Time (sec)	Total Time (sec)
IHFold	0.0233	1.748
Hotknots	0.0709	5.314
IPknot	0.0143	1.070
Knotty	0.0212	1.590
knotify_yaep	0.0697	5.226
knotify_bruteforce	0.0427	3.204

Table 5.12: Execution time required per platform in for RNA sequences of length >= 30 and < 40.

Platform	Average Time (sec)	Total Time (sec)
IHFold	0.0248	1.689
Hotknots	0.0982	6.680
IPknot	0.0408	2.777
Knotty	0.0692	4.703
knotify_yaep	0.1589	10.808
knotify_bruteforce	0.0964	6.555

Table 5.13: Execution time required per platform in for RNA sequences of length >= 40 and < 50.

Platform	Average Time (sec)	Total Time (sec)
IHFold	0.0274	1.507
Hotknots	0.1503	8.264
IPknot	0.107	5.886
Knotty	0.1918	10.546
knotify_yaep	0.2331	12.821
knotify_bruteforce	0.111	6.103

Table 5.14: Execution time required per platform in for RNA sequences of length > 50.

Platform	Average Time (sec)	Total Time (sec)
IHFold	0.0492	3.151
Hotknots	0.3777	24.172
IPknot	1.679	107.511
Knotty	3.851	246.462
knotify_yaep	0.8891	56.900
knotify_bruteforce	0.2817	18.030

Total execution time required per platform for the four groups of different RNA sequences' length is also shown in Figure 5.12. Please note that a logarithmic scale was used.
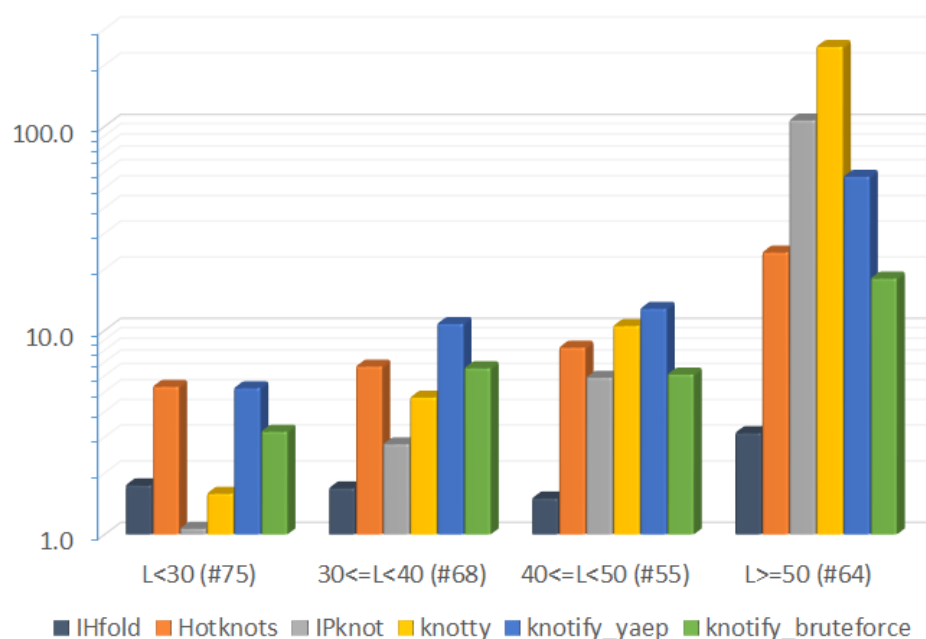
Figure 5.12: Average execution time (sec) required per platform.

5.7   Discussion and future work

Recently, COVID-19 evolved into a severe global threat. This virus' connection to RNA is one of the most-prominent artifacts justifying the requirement of concrete and accurate RNA-analysis toolkits. This study introduced an innovative methodology to detect H-type pseudoknots in RNA secondary structures accurately and performantly. The method is based on Early's parser, which, given an RNA sequence in the form of a string, produces a space of all possible parse-trees, each one expressing a potential pseudoknot structure. The optimal tree is picked through a hybrid model combining pairing maximization and free-energy minimization for each context–tree pair. The evaluation of the proposed algorithm demonstrated its outperformance to the task of pseudoknot prediction, paving the way for our future work endeavours: (i) discovering even more complex patterns such as pseudoknots enclosing bulges or hairpins, (ii) building advanced searching algorithms leveraging common patterns of the secondary structure of RNA, and (iii) creating an open web platform to make our work accessible to all researchers. The work in [ARTM18, ARTM15b] may be extended further to facilitate enhanced RNA-analysis services to fulfill our vision of unified collaboration among the members of an interdisciplinary team of healthcare experts since the future of medicine will be genomics.

CHAPTER 6

Conclusion & Future work

This thesis introduces a novel design, implementation and deployment of a coherent and agile CSCW environment for teleconsultations over medical imaging data. The proposed system enhances communication among different specialists and overcomes MDTM barriers in a robust way. Table 3.2 provides a qualitative assertion of the system's functionality compared with the related high-valued works found in literature. The meetings are now easily integrated into the work schedule of specialists, while small or regional hospitals need no special facilities for teleconferencing to support remote MDTMs. Its rich functionality may also be utilized as a valuable tool to support teaching medical students and junior doctors [Gao13] or it can further empower research groups to collaborate efficiently to resolve issues such as image-based co-registration [WEC+13] or co-segmentation [HFW+11] [SNH+09] [Må06].

The growth of the older adult population in the world is surprising and it will have a great impact on the healthcare system. In the current digital era, the volume of radiological imaging exams is constantly growing due to the (1) the increased living standards and (2) the constantly decreasing cost of healthcare infrastructure because of the recent technological advances. In this equation the only constant factor is the available human intelligence and labour. To meet this increased workload demand, the average radiologist may have to interpret an image every 3–4 s over an 8-h workday which contributes to fatigue, burnout, and increased error-rate [MSE+15]. Deep learning in healthcare is proliferating due to the potential for successful automated systems to either augment or offload cognitive work from busy physicians [DJJ+15, BSY+19, SKSJ20]. In this context, The introduced platform is extended further into to inter-domain fashion. The integration of the novel approach of urinary bladder reconstruction 4 paves the way for a hybrid collaboration schema among domain experts and computer scientists as depicted in the work in [HPS+20]. The results depicted at tables 4.1 and 4.2 highlight the significant value added to the platform by the integration of novel algorithms to assist medical data analysis. The accuracy of the discussed empowers early diagnosis of urine bladder cancer. Any tumour of size larger than 0.05 mm is obvious in the 3D reconstructed model. The non-invasively acquired imaging data enables the periodical administration of the examination due to the zero impact to the physical or mental health of the patient. Despite, the fact that our model considers only the pixel data as a single modality for input it provides an output that enhances the clinical translation of combined distinct imaging data sources.

Considering classic computer vision application to biomedical imaging data, or custom analysis workflows that are build to operate on genomics. The bibliography lists a plethora of advanced algorithms that potentially add great value the wide community of (bio)medical data analysis that never found their way to the wild. To extend our work further to that direction, we introduced a novel ribonucleic acid analysis (RNA) service since obtaining valuable clues for RNA remains a significant challenge that impacts the broader community of healthcare provisioners. The integrated service addresses the problem of predicting pseudoknots of the first order (H-type) utilizing a context-free grammar (CFG). Featuring a polynomial-time complexity and boosted by a lean parallelization technique proportionally to the deployed hardware, the proposed methodology does succeed in predicting the core stems of any RNA pseudoknot of the test dataset by performing a 76.4% recall ratio. The methodology achieved a F1-score equal to 0.774 and MCC equal 0.543 in discovering all the stems of an RNA sequence; it outperforms the task compared to most of the suggested implementations that literature suggests.

Having incorporated novel approaches related to both medical imaging and genomics domains and integrating them to a sustainable collaborating platform we serve our vision of *creating a space to let the theoritical, clinical and algorithmic prospective of medical data analysis to co-exist.* Having said that and given the fact that the AI and the ML class of computer algorithms have proven very effective to recognition, classification and segmentation tasks, our focus shifts to that direction, i.e. the extension of the introduced platform to empower human-machine collaboration in the broader field of medicine.

In chapter A of the Appendix we are present a novel embedding architecture to tackle the problem of detecting pneumonia in chest radiographs (CXRs) which has attracted recently the interest of several research groups from the academia and the industry especially due to the recent Covid outbreak. In the framework of research challenges and competitions enabled by the public dataset published by the National Institutes of Health Clinical Center. In this work, we aspire to introduce a novel three-step pipeline for pneumonia detection on CXR imaging data. CXR images are initially distinguished to target and control cases, then the lung area is extracted using a simple network and finally, a complex Deep Neural Network architecture discovers inconspicuous patterns such as pneumonia artifacts. The chapter (A) discusses the technical details of the proposed methodology and provides some initial results.

Another interest topic is the semantic ROI annotation though taxonomies and ontologies. Given an image and ROI of it, a list of related studies should be retrieved from the PubMed paper silo [CW13]. Linking selected ROIs to semantically organized taxonomies should enhance the discovery of novel knowledge through generation of new hypotheses [SKI+14]. In this way, the available resources would enable clinical personnel and researchers across different disciplines to discover associations among their areas of interest and those that have already been studied and classified.

Another interesting concept it the one of "artifact-rooms". Practitioners and researchers would be able to setup "workspaces", consisting of medical data accompanied by diagnosis workflows and invite some colleagues to assist them at a later time. Ex post assistance would probably reduce distractions and increase concentration levels when it is required.

We argue that our collaboration technology ecosystem has the potential to decisively contribute in many novel ways that remain to be explored. Telemedicine or CAD oriented medicine

are the most obvious ones; we imagine through the potential application of its own to education when fresh physicians would be able to participate in meetings that are going to assist their expertise growth. Another application to be explored is the systematic and massive data annotation. One of the most challenging aspects of the newly introduced machine learning paradigm is the creation of high-quality reference datasets (ground-truth annotations). A slight extension of the introduced would arguably contribute to medical data annotation by experts on the medical field. The latter is one of the core concepts that we propose for further research steps.

However, the biggest challenge would be the evolution of the platform in a way that would enable researchers to add functionality to be consumed directly by specialists; consider dedicated biomedical image analysis operations or machine learning modules to form an open ecosystem of software components generated by members of some community, evaluated by healthcare specialists and then applied on real medical data to extend and enrich existing medical services or to provide novel ones or even assist education. Nowadays the testing or use of such an approach into a real life case requires serious programming skill and system knowledge. This fact contradicts the recent trend of user-friendly tools, which allow users to envision the implementation and define overall complex business logic through drag and drop interfaces. The so-called "low-" or "no-code" frameworks are software design systems that enable even non-technical people to execute software without composing a line of code that ends to two key benefits:

- The healthcare specialists should focus on the details related to their existing domain of knowledge. Any no-code generated application signifies that no time has been spent on technical details essential for its existence. On the contrary, the team behind that particular project concentrated on the application's heart.

- By offering rich models to utilize and configure, no-code components (i.e. computational units with standardized input and output formats) abstract a lot of the complexity in designing the front-end user interface. There is no need to start from scratch with UI/UX, and the time saved may be the core concept of the development effort.

- There is a significant reduction in the overall cost of development. Some studies suggest even an 80% lower development costs when in the no-code approach —eliminating the need to recruit special technical skills such as front-end and back-end developers, financial experts, security consultants, and software and cloud experts. The nuances of front-end and back-end development are abstracted for no-code application development. They will develop faster, and they don't have to write any code from the ground up or think about underlying systems like cases, databases, or security when you're building on a network.

Apparently, "no-code" benefits do not come to a zero cost. Going with the "no" or "low" code approach one has to trade with restrictions in the overall computational unit interface. To move on that way a common glossary should be defined. This is something we have already developed and defined in 3.2.2. What we miss is a common protocol ot invoke such computational units.

Customization of software in no-code platforms is restricted. To put it another way, you'll have to adapt your company procedures to reach the capabilities of the no-code network. Since you wrote the code, you know you will depend on it heavily while creating it. However, since you don't have full authority when dealing with no-code, you can take some risks. It can be an excellent opportunity for security breaches, and once your no-code platform is compromised,
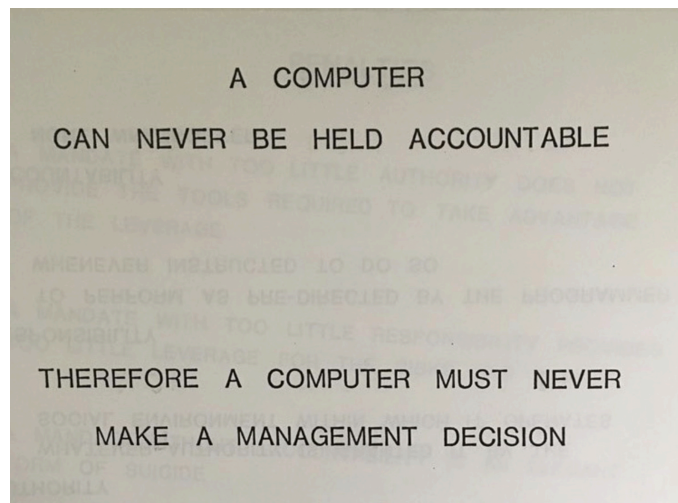
Figure 6.1: IBM training slight, 1979. What has changed? Only the domain we argue ...

your program will become insecure as well. Especially in the field of the healthcare should be a topic of further research due to the sensitive nature of the (bio)medical data.

Finally our work highlights both the promise and the challenges faced by the introduced collaboration platform. Hopefully, this thesis will provide a good reference for future work, analysis algorithms, and machine learning of medical images. However, there are still a key message that emerge from our research experience in the specific domain: digitalization and automation have been showed up to stay and disrupt the healthcare operations; they are here to support the work of health care experts. They are not here to substitute the doctors. And in case you consider the opposite just recall the well known IBM 1979 slight 6.1.

Bibliography

[AAB+17]   A Amoroso, F Astorino, S Bagnasco, N A Balashov, F Bianchi, M Destefa-
             nis, S Lusso, M Maggiora, J Pellegrino, L Yan, T Yan, X Zhang, and X Zhao.
             A modular (almost) automatic set-up for elastic multi-tenants cloud (mi-
             cro)infrastructures. *journal of Physics: Conference Series*, 898(8):082031, 2017.

[ABL13]   Foteini Gr Andriopoulou, Konstantinos D Birkos, and Dimitrios K Lymberopou-
             los. P2care: A dynamic self-organized healthcare network for ubiquitous
             healthcare environment. In *e-Health Networking, Applications & Services
             (Healthcom), IEEE 15th International Conference on*, pages 294–298. IEEE, 2013.

[ABL15]   Foteini Gr Andriopoulou, Konstantinos Birkos, and Dimitrios Lymberopou-
             los. P2care: A dynamic peer-to-peer network for collaboration in personalized
             healthcare service delivery. *Computers in Industry*, 69:45–60, 2015.

[Ack00]   Mark S. Ackerman. The intellectual challenge of cscw: The gap between social
             requirements and technical feasibility. *Hum.-Comput. Interact.*, 15(2):179–203,
             sep 2000.

[AHM+13]   J. W. Anderson, P. A. Haas, L.-A. Mathieson, V. Volynkin, R. Lyngsø, P. Tataru,
             and J. Hein. Oxfold: kinetic folding of rna using stochastic context-free gram-
             mars and evolutionary information. *Bioinformatics*, 29(6):704–710, 2013.

[Ahm17]   R. Ahmed. Systems and methods for medical diagnostic collaboration. `https:
             //www.google.com/patents/US20170195377`, July 2017. US
             Patent App. 15/462,336.

[Aku00]   Tatsuya Akutsu. Dynamic programming algorithms for rna secondary struc-
             ture prediction with pseudoknots. *Discrete Applied Mathematics*, 104(1-3):45–62,
             2000.

[ALSU06]   Alfred V. Aho, Monica S. Lam, Ravi Sethi, and Jeffrey D. Ullman. *Compilers:
             Principles, Techniques, and Tools (2nd Edition)*. Addison Wesley, August 2006.

[AMB+14]   Christos Andrikos, Ilias Maglogiannis, Efthymios Bilalis, George Spyroglou,
             and Panayiotis Tsanakas. An intelligent platform for hosting medical collabo-
             rative services. In *Hellenic Conference on Artificial Intelligence*, pages 354–359.
             Springer, 2014.

[AMK$^+$22]   Christos Andrikos, Evangelos Makris, Angelos Kolaitis, Georgios Rassias, Christos Pavlatos, and Panayiotis Tsanakas. Knotify: An efficient parallel platform for rna pseudoknot prediction using syntactic pattern recognition. *Methods and Protocols*, 5(1):14, 2022.

[APZ$^+$18]   Maciej Antczak, Mariusz Popenda, Tomasz Zok, Michal Zurkowski, Ryszard W. Adamiak, and Marta Szachniuk. New algorithms to represent complex pseudoknotted rna structures in dot-bracket notation. *Bioinformatics*, 34:1304 – 1312, 2018.

[ARTM15a]   Christos Andrikos, Georgios Rassias, Panayiotis Tsanakas, and Ilias Maglogiannis. Real-time medical collaboration services over the web. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1393–1396. IEEE, 2015.

[ARTM15b]   Christos Andrikos, Georgios Rassias, Panayiotis Tsanakas, and Ilias Maglogiannis. Real-time medical collaboration services over the web. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1393–1396. IEEE, 2015.

[ARTM18]   Christos Andrikos, Georgios Rassias, Panayiotis Tsanakas, and Ilias Maglogiannis. An enhanced device-transparent real-time teleconsultation environment for radiologists. *IEEE journal of biomedical and health informatics*, 23(1):374–386, 2018.

[ASK$^+$04]   Peter L Adams, Mary R Stahley, Anne B Kosek, Jimin Wang, and Scott A Strobel. Crystal structure of a self-splicing group i intron with both exons. *Nature*, 430(6995):45–50, 2004.

[AZDR10]   Saad Alahmari, Ed Zaluska, and David De Roure. A service identification framework for legacy system migration into soa. In *2010 IEEE International Conference on Services Computing*, pages 614–617. IEEE, 2010.

[B$^+$96]   John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[BAAH16]   Riadh Bouslimi, Jalel Akaichi, Mouhamed Gaith Ayadi, and Hana Hedhli. A medical collaboration network for medical image analysis. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 5(1):10, 2016.

[BBG$^+$15]   S Bagnasco, D Berzano, A Guarise, S Lusso, M Masera, and S Vallero. Monitoring of iaas and scientific applications on the cloud using the elasticsearch ecosystem. *journal of Physics: Conference Series*, 608(1):012016, 2015.

[BD09]   Bernd Bruegge and Allen H Dutoit. Object−oriented software engineering. using uml, patterns, and java. *Learning*, 5(6):7, 2009.

[BHH03]   Martin Beer, Wei Huang, and Richard Hill. Designing community care systems with auml. 2003.

[BHPVS97]   W Dean Bidgood, Steven C Horii, Fred W Prior, and Donald E Van Syckle. Understanding and using dicom, the data interchange standard for biomedical imaging. *journal of the American Medical Informatics Association*, 4(3):199–212, 1997.

[BHW⁺08]   Stephan Bernhart, Ivo Hofacker, Sebastian Will, Andreas Gruber, and Peter Stadler. Rnaalifold: Improved consensus structure prediction for rna alignments. *BMC Bioinformatics*, 9, 11 2008.

[BM10]   Stanislav Bellaousov and David H Mathews. Probknot: fast prediction of rna secondary structure including pseudoknots. *RNA*, 16:1870–80, 2010.

[BPH08]   R. K. Bradley, L. Pachter, and I. Holmes. Specific alignment of structured rna: stochastic grammars and sequence annealing. *Bioinformatics*, 24(23):2677–2683, 2008.

[BPYM05]   Daniel P. Barboriak, Anthony O. Padua, Gerald E. York, and James R. MacFall. Creation of dicom—aware applications using imagej. *journal of Digital Imaging*, 18(2):91–99, 2005.

[BSH17]   Marcus D Bloice, Christof Stocker, and Andreas Holzinger. Augmentor: an image augmentation library for machine learning. *arXiv preprint arXiv:1708.04680*, 2017.

[BSY⁺19]   Imon Banerjee, Miji Sofela, Jaden Yang, Jonathan H Chen, Nigam H Shah, Robyn Ball, Alvin I Mushlin, Manisha Desai, Joseph Bledsoe, Timothy Amrhein, et al. Development and performance of the pulmonary embolism result forecast model (perform) for computed tomography clinical decision support. *JAMA network open*, 2(8):e198719–e198719, 2019.

[CC]   Evan Czaplicki and Stephen Chong. Asynchronous functional reactive programming for guis. In *ACM SIGPLAN Notices*, volume 48, pages 411–422. ACM.

[CCCD93]   Vicent Caselles, Francine Catté, Tomeu Coll, and Françoise Dibos. A geometric model for active contours in image processing. *Numerische mathematik*, 66(1):1–31, 1993.

[CCK⁺02]   Elaine M Caoili, Richard H Cohan, Melvyn Korobkin, Joel F Platt, Isaac R Francis, Gary J Faerber, James E Montie, and James H Ellis. Urinary tract abnormalities: initial experience with multi–detector row ct urography. *Radiology*, 222(2):353–360, 2002.

[CF84]   Y.T. Chiang and K.S. Fu. Parallel parsing algorithms and VLSI implementations for syntactic pattern recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:302–314, 1984.

[CFB⁺11]   Carlos Costa, Carlos Ferreira, Luís Bastião, Luís Ribeiro, Augusto Silva, and José Luís Oliveira. Dicoogle-an open source peer-to-peer pacs. *journal of digital imaging*, 24(5):848–856, 2011.

[CFS17]   Dalen Chan, Chao Feng, and Robert Spitale.   Measuring rna structure
               transcriptome-wide with icshape. *Methods*, 120, 03 2017.

[che]   Nih chest xray dataset of 14 common thorax disease categories.
               Available   at   `http://academictorrents.com/details/`
               `557481faacd824c83fbf57dcf7b6da9383b3235a`.   Accessed
               an 2019-02-12.

[Che18a]   Lianping Chen. Continuous delivery at scale: challenges and opportunities.
               In *Proceedings of the 4th International Workshop on Rapid Continuous Software
               Engineering*, pages 42–42, 2018.

[Che18b]   Lianping Chen. Microservices: Architecting for continuous delivery and de-
               vops. In *2018 IEEE International Conference on Software Architecture (ICSA)*,
               pages 39–397. IEEE, 2018.

[Cho56]   N. Chomsky. Three models for the description of language. *IRE Transactions on
               Information Theory*, 2(3):113 –124, september 1956.

[Cho17]   François Chollet. Xception: Deep learning with depthwise separable convolu-
               tions. *arXiv preprint*, pages 1610–02357, 2017.

[CKO10]   D.-J. Chang, C. Kimmer, and M. Ouyang. Accelerating the nussinov rna fold-
               ing algorithm with cuda/gpu. *in Signal Processing and Information Technology
               (ISSPIT)*, 2010:120–125, 2010.

[CMT⁺16]   Ana Conesa, Pedro Madrigal, Sonia Tarazona, David Gomez-Cabrero, Alejandra
               Cervera, Andrew McPherson, Michał Wojciech Szcześniak, Daniel J Gaffney,
               Laura L Elo, Xuegong Zhang, et al. A survey of best practices for rna-seq data
               analysis. *Genome biology*, 17(1):1–19, 2016.

[Coc07]   Alistair Cockburn. Hexagonal architecture, 2007.

[cod]   `https://github.com/ntua-dslab/knotify`.

[CR12]   Yongjun Chu and David R.Corey. Rna sequencing: Platform selection, experi-
               mental design, and data interpretation. *Nucleic Acid Therapeutics*, 22:271–274,
               2012.

[CS14]   Thomas Cech and Joan Steitz. The noncoding rna revolution—trashing old rules
               to forge new ones. *Cell*, 157:77–94, 03 2014.

[CSDL17]   Juliana Costa-Silva, Douglas Domingues, and Fabricio Martins Lopes. Rna-seq
               differential expression analysis: An extended review and a software tool. *PloS
               one*, 12(12):e0190152, 2017.

[CW13]   Kathi Canese and Sarah Weis. Pubmed: the bibliographic database. *The NCBI
               handbook*, 2(1), 2013.

[DAB⁺01]   Robert H Dolin, Liora Alschuler, Calvin Beebe, Paul V Biron, Sandra Lee Boyer, Daniel Essin, Elliot Kimber, Tom Lincoln, and John E Mattison. The hl7 clinical document architecture. *journal of the American Medical Informatics Association*, 8(6):552–569, 2001.

[dat]   `https://drive.google.com/file/d/13_cfX4pVSSYBsWxbrco1Xmv_bwAQ_Dzj/view?usp=sharing`.

[DC02]   Jennifer Doudna and Thomas Cech. The chemical repertoire of natural ribozymes. *Nature*, 418:222–8, 08 2002.

[DDS⁺09]   Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. IEEE, 2009.

[dea]   Deaths final data for 2015. supplemental tables. tables i-21, i-22. `https://www.cdc.gov/nchs/data/nvsr/nvsr66/nvsr66_06_tables.pdf`. Accessed on: 2019-02-12.

[DFK07]   Wayne K Dawson, Kazuya Fujiwara, and Gota Kawai. Prediction of rna pseudoknots using heuristic modeling with mapping and sequential folding. *PloS one*, 2(9):905, 2007.

[DHT⁺04]   Willem MLLG Deserno, Mukesh G Harisinghani, Matthias Taupitz, Gerrit J Jager, J Alfred Witjes, Peter F Mulders, Christina A Hulsbergen van de Kaa, D Kaufmann, and Jelle O Barentsz. Urinary bladder cancer: preoperative nodal staging with ferumoxtran-10–enhanced mr imaging. *Radiology*, 233(2):449–456, 2004.

[DJ10]   Pieter Duysburgh and An Jacobs. *Collaboration through ICT between Healthcare Professionals: The Social Requirements of Health 2.0 Applications*, pages 165–172. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[DJJ⁺15]   Nathan C Dean, Barbara E Jones, Jason P Jones, Jeffrey P Ferraro, Herman B Post, Dominik Aronsky, Caroline G Vines, Todd L Allen, and Peter J Haug. Impact of an electronic clinical decision support tool for emergency department patients with pneumonia. *Annals of emergency medicine*, 66(5):511–520, 2015.

[DLB⁺10]   Chaijie Duan, Zhengrong Liang, Shangliang Bao, Hongbin Zhu, Su Wang, Guangxiang Zhang, John J Chen, and Hongbing Lu. A coupled level set framework for bladder wall segmentation with application to mr cystography. *IEEE transactions on medical imaging*, 29(3):903–915, 2010.

[DLL⁺17]   Nicola Dragoni, Ivan Lanese, Stephan Thordal Larsen, Manuel Mazzara, Ruslan Mustafin, and Larisa Safina. Microservices: How to make your application scale. In *International Andrei Ershov Memorial Conference on Perspectives of System Informatics*, pages 95–104. Springer, 2017.

[DP03]   Robert Dirks and Niles Pierce. Introduction a partition function algorithm for nucleic acid secondary structure including pseudoknots. *Journal of computational chemistry*, 24:1664–77, 10 2003.

[DWB06]   C. B. Do, D. A. Woods, and S. Batzoglou. Contrafold: Rna secondary structure prediction without physics-based models. *Bioinformatics*, 22:e90–e98, 2006.

[DZL⁺17]   Chunhua Dong, Xiangyan Zeng, Lanfen Lin, Hongjie Hu, Xianhua Han, Masoud Naghedolfeizi, Dawit Aberra, and Yen-Wei Chen. An improved random walker with bayes model for volumetric medical image segmentation. *Journal of healthcare engineering*, 2017, 2017.

[E.14]   Trotta E. On the normalization of the minimum free energy of rnas by sequence length. *PLoS One.*, 9(11), 2014.

[Ear70]   Jay Earley. An efficient context-free parsing algorithm. *Commun. ACM*, 13(2):94–102, 1970.

[Edd02]   Sean Eddy. Non-coding rna genes and the modern rna world. *Nature reviews. Genetics*, 2:919–29, 01 2002.

[EKN⁺17]   Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115, 2017.

[Eva04]   Eric Evans. *Domain-driven design: tackling complexity in the heart of software.* Addison-Wesley Professional, 2004.

[FD96]   Aaron Fenster and Donal B Downey. 3-d ultrasound imaging: A review. *IEEE Engineering in Medicine and Biology magazine*, 15(6):41–51, 1996.

[Fow15]   Martin Fowler. Microservicepremium. *Retrieved Jan*, 25:2019, 2015.

[Fow18]   Martin Fowler. *Microservices.* Archived, 2018.

[Fra]   T. Franquet. Imaging of community-acquired pneumonia. *J Thorac Imaging*.

[Gao13]   Fei Gao. A case study of using a social annotation tool to support collaboratively learning. *The Internet and Higher Education*, 17(Supplement C):76–83, 2013.

[GHR80]   Susan L. Graham, Michael A. Harrison, and Walter L. Ruzzo. An improved context-free recognizer. *ACM Trans. Program. Lang. Syst.*, 2(3):415–462, 1980.

[GKK⁺02]   Mark O Gueld, Michael Kohnen, Daniel Keysers, Henning Schubert, Berthold B Wein, Joerg Bredno, and Thomas M Lehmann. Quality of dicom header information for image categorization. In *Medical Imaging*, pages 280–287. International Society for Optics and Photonics, 2002.

[GKQ16]   Charles Gawad, Winston Koh, and Stephen Quake. Single-cell genome sequencing: Current state of the science. *Nature reviews. Genetics*, 17, 01 2016.

[GLW⁺01]   Gary D Grossfeld, Mark S Litwin, J Stuart Wolf, Hedvig Hricak, Cathryn L Shuler, David C Agerter, and Peter R Carroll. Evaluation of asymptomatic microscopic hematuria in adults: the american urological association best practice policy—part i: definition, detection, prevalence, and etiology1. *Urology*, 57(4):599–603, 2001.

[GPS05]   RNJ Graham, RW Perriss, and AF Scarsbrook. Dicom demystified: a review of digital file formats and their use in radiological practice. *Clinical radiology*, 60(11):1133–1140, 2005.

[GXM⁺14]   Teng Geng, Fu Xu, Han Mei, Wei Meng, Zhibo Chen, and Changqing Lai. A practical GLR parser generator for software reverse engineering. *JNW*, pages 769–776, 2014.

[H.04]   Mathews D. H. Using an rna secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA (New York, N.Y.)*, 10(8):1178–1190, 2004.

[HAB17]   Sara Hassan, Nour Ali, and Rami Bahsoon. Microservice ambients: An architectural meta-modelling approach for microservice granularity. In *2017 IEEE International Conference on Software Architecture (ICSA)*, pages 1–10. IEEE, 2017.

[Har]   Derrick Harris. Microservices, monoliths and laser nail guns: Etsy tech boss on finding the right focus.

[HFW⁺11]   R.S. Huennekens, S.M. Fry, B.D. Walker, J.D. Klingensmith, N.P. Pool, V.J. Burgess, and W.R. Kanz. Vascular image co-registration. `https://www.google.com/patents/US7930014`, 2011.

[HGS⁺17]   Benjamin Hippchen, Pascal Giessler, Roland Steinegger, Michael Schneider, and Sebastian Abeck. Designing microservice-based applications by using a domain-driven design approach. *International Journal on Advances in Software*, 10(3&4):432–445, 2017.

[HGZ⁺17]   Han Hu, Jiayuan Gu, Zheng Zhang, Jifeng Dai, and Yichen Wei. Relation networks for object detection. *arXiv preprint arXiv:1711.11575*, 2017.

[HH13]   Shaista Hafeez and Robert Huddart. Advances in bladder cancer imaging. *BMC medicine*, 11(1):1–10, 2013.

[HLS⁺99]   Frits J Huyse, John S Lyons, Friedrich C Stiefel, Joris PJ Slaets, Peter de Jonge, Per Fink, Rijk OB Gans, Patrice Guex, Thomas Herzog, and Antonio Lobo. "intermed": a method to assess health service needs: I. development and reliability. *General hospital psychiatry*, 21(1):39–48, 1999.

[HLVDMW17] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[HP17]   Olaf Hartig and Jorge Pérez. An initial analysis of facebook's graphql language.
             2017.

[HPB05]   Richard Hill, Simon Polovina, and Martin Beer. From concepts to agents: to-
             wards a framework for multi-agent system modelling. In *Proceedings of the
             fourth international joint conference on Autonomous agents and multiagent sys-
             tems*, pages 1155–1156, 2005.

[HPD16]   Daniel Haak, Charles-E Page, and Thomas M Deserno. A survey of dicom
             viewer software to integrate clinical research and medical imaging. *Journal
             of digital imaging*, 29(2):206—215, April 2016.

[HPKD15]   Daniel Haak, Charles E Page, Klaus Kabino, and Thomas M Deserno. Evaluation
             of dicom viewer software for workflow integration in clinical trials. In *SPIE
             Medical Imaging*, pages 94180O–94180O–9. International Society for Optics and
             Photonics, 2015.

[HPS+20]   Shih-Cheng Huang, Anuj Pareek, Saeed Seyyedi, Imon Banerjee, and
             Matthew P Lungren. Fusion of medical imaging and electronic health records
             using deep learning: a systematic review and implementation guidelines. *NPJ
             digital medicine*, 3(1):1–9, 2020.

[HRX+12]   Wei He, Peng Ran, Zheng Xu, Bing Li, and Song-nong Li. A 3d visualization
             method for bladder filling examination based on eit. *Computational and Math-
             ematical Methods in Medicine*, 2012, 2012.

[HSS+17]   Nikolai Hecker, Stefan Seemann, Asli Silahtaroglu, Walter Ruzzo, and Jan
             Gorodkin. Associating transcription factors and conserved rna structures with
             gene regulation in the human brain. *Scientific Reports*, 7, 12 2017.

[htm]   Leadtools html5 zero-footprint medical viewer. `https://www.
             leadtools.com/sdk/medical/html5`.

[HU69]   John E. Hopcroft and Jeffrey D. Ullman. *Formal Languages and Their Relation to
             Automata*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA,
             1969.

[Hus16]   Tom Huston. What is microservices architecture? *SmartBear,[Online]. Avail-
             able: https://smartbear. com/learn/api-design/what-are-microservices*, 2016.

[IAH+15]   Ahsan Ikram, Ashiq Anjum, Richard Hill, Nick Antonopoulos, Lu Liu, and Ste-
             lios Sotiriadis. Approaching the internet of things (iot): a modelling, analysis
             and abstraction framework. *Concurrency and Computation: Practice and Expe-
             rience*, 27(8):1966–1984, 2015.

[IMK+14]   Forrest Iandola, Matt Moskewicz, Sergey Karayev, Ross Girshick, Trevor Dar-
             rell, and Kurt Keutzer. Densenet: Implementing efficient convnet descriptor
             pyramids. *arXiv preprint arXiv:1404.1869*, 2014.

[IS00]   Hervé Isambert and Eric D Siggia. Modeling rna folding paths with pseudo-knots: application to hepatitis delta virus ribozyme. *Proceedings of the National Academy of Sciences*, 97(12):6515–6520, 2000.

[JBD+]   Sébastien Jodogne, Claire Bernard, Magali Devillers, Eric Lenaerts, and Philippe Coucke. Orthanc-a lightweight, restful dicom server for healthcare and medical research. In *2013 IEEE 10th International Symposium on Biomedical Imaging*, pages 190–193. IEEE.

[JCA+]   S. Jaeger, S. Candemir, S. Antani, Y.X. Wáng, P.X. Lu, and G. Thoma. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. In *Quantitative imaging in medicine and surgery*, volume 4, page 475–7.

[JCA+14]   Stefan Jaeger, Sema Candemir, Sameer Antani, Yì-Xiáng J Wáng, Pu-Xuan Lu, and George Thoma. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative imaging in medicine and surgery*, 4(6):475, 2014.

[JHY+17]   Dai Jifeng, Qi Haozhi, Xiong Yuwen, Li Yi, Zhang Guodong, Hu Han, and Wei Yichen. Deformable convolutional networks. *arXiv preprint arXiv:1703.06211*, 2017.

[JLSH+19]   Ilaria Jansen, Marit Lucas, C Dilara Savci-Heijink, Sybren L Meijer, Esmee IML Liem, Onno J de Boer, Ton G van Leeuwen, Henk A Marquering, and Daniel M de Bruin. Three-dimensional histopathological reconstruction of bladder tumours. *Diagnostic pathology*, 14(1):1–7, 2019.

[JTK+20]   Gumna J, Zok T, Figurski K, Pachulska-Wieczorek K, and Szachniuk M. Rnathor - fast, accurate normalization, visualization and statistical analysis of RNA probing data resolved by capillary electrophoresis. *PLoS One. 2020 Oct 1;15(10):e0239287. doi:*, 10., 2020.

[JWMW18]   Hosna Jabbari, Ian Wark, Carlo Montemagno, and Sebastian Will. Knotty: efficient and accurate prediction of complex RNA pseudoknot structures. *Bioinformatics*, 34(22):3849–3856, 2018.

[JYKJ16]   Dai Jifeng, Li Yi, He Kaiming, and Sun Jian. R-FCN: Object detection via region-based fully convolutional networks. 2016.

[KBR+17]   Maximilian C Kriegmair, Tobias Bergen, Manuel Ritter, Philipp Mandel, Maurice S Michel, Thomas Wittenberg, and Christian Bolenz. Digital mapping of the urinary bladder: Potential for standardized cystoscopy reports. *Urology*, 104:235–241, 2017.

[KDP+15]   Su Yeon Kim, James Diggans, Dan Pankratz, Jing Huang, Moraima Pagan, Jeffrey L Myers, Anna-Luise A Katzenstein, Ganesh Raghu, and Giulia C Kennedy. Classification of usual interstitial pneumonia in patients with interstitial lung disease: assessment of a machine learning approach using high-dimensional transcriptional data. *The Lancet Respiratory Medicine*, 3(6):473 – 482, 2015.

[KF97]   David G Kilman and David W Forslund. An international collaboratory based
            on virtual patient records. *Communications of the ACM*, 40(8):110–117, 1997.

[KGS+13]   Shin-ya Katayama, Takushi Goda, Shun Shiramatsu, Tadachika Ozono, and
            Toramatsu Shintani. A fast synchronization mechanism for collaborative web
            applications based on html5. In *Software Engineering, Artificial Intelligence,
            Networking and Parallel/Distributed Computing (SNPD), 14th ACIS International
            Conference on*, pages 663–668. IEEE, 2013.

[KH99]   B Knudsen and J Hein. RNA secondary structure prediction using stochastic
            context-free grammars and evolutionary history. *Bioinformatics*, 15(6):446–454,
            1999.

[KH03]   Bjarne Knudsen and Jotun Hein. Pfold: Rna secondary structure prediction
            using stochastic context-free grammars. *Nucleic acids research*, 31:3423–8, 08
            2003.

[KHSQ16]   Marcel Kuchařík, Ivo L Hofacker, Peter F Stadler, and Jing Qin. Pseudoknots in
            rna folding landscapes. *Bioinformatics*, 32(2):187–194, 2016.

[KJY20]   Mao Kangkun, Wang Jun, and Xiao Yi. Prediction of rna secondary structure
            with pseudoknots using coupled deep neural networks. *Biophysics Reports*,
            6:146–154, 2020.

[KL06]   George V Koutelakis and Dimitrios K Lymperopoulos. Pacs through web com-
            patible with dicom standard and wado service: advantages and implementation.
            In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual In-
            ternational Conference of the IEEE*, pages 2601–2605. IEEE, 2006.

[KOK+17]   Nur Diyana Kamarudin, Chia Yee Ooi, Tadaaki Kawanabe, Hiroshi Odaguchi,
            and Fuminori Kobayashi. A fast svm-based tongue's colour classification aided
            by k-means clustering identifiers and colour attributes as computer-assisted
            tool for tongue diagnosis. *Journal of healthcare engineering*, 2017, 2017.

[KS03]   Vikas Kundra and Paul M Silverman. Imaging in the diagnosis, staging, and
            follow-up of cancer of the urinary bladder. *American Journal of Roentgenology*,
            180(4):1045–1054, 2003.

[KTS15]   Miles Kubota, Catherine Tran, and Robert Spitale. Progress and challenges for
            chemical probing of rna structure inside living cells. *Nature chemical biology*,
            11:933–941, 11 2015.

[KUB+06]   P. S. Klosterman, A. V. Uzilov, Y. R. Benda na, R. K. Bradley, S. Chao, C. Kosiol,
            N. Goldman, and I. Holmes. Xrate: a fast prototyping, training and annotation
            tool for phylo-grammars. *BMC bioinformatics*, 7:428, 2006.

[KZD+04]   Ailong Ke, Kaihong Zhou, Fang Ding, Jamie HD Cate, and Jennifer A Doudna. A
            conformational switch controls hepatitis delta virus ribozyme catalysis. *Nature*,
            429(6988):201–205, 2004.

[KZH⁺14]   Maulik R Kamdar, Dimitris Zeginis, Ali Hasnain, Stefan Decker, and Helena F
            Deus. Reveald: A user-driven domain-specific interactive search platform for
            biomedical research. *journal of biomedical informatics*, 47:112–130, 2014.

[LBHzS⁺11] Ronny Lorenz, Stephan Bernhart, Christian Höner zu Siederdissen, Hakim
            Tafer, Christoph Flamm, Peter Stadler, and Ivo Hofacker. Viennarna package
            2.0. *Algorithms for molecular biology : AMB*, 6:26, 11 2011.

[LC87]   William E Lorensen and Harvey E Cline. Marching cubes: A high resolution
            3d surface construction algorithm. *ACM siggraph computer graphics*, 21(4):163–
            169, 1987.

[LCQ⁺18]   Xiaomeng Li, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, and Pheng-Ann
            Heng. H-denseunet: Hybrid densely connected unet for liver and tumor seg-
            mentation from ct volumes. *IEEE transactions on medical imaging*, 37(12):2663–
            2674, 2018.

[LE97]   T. M. Lowe and S. R. Eddy. trnascan-se: a program for improved detection of
            transfer rna genes in genomic sequence. *Nucleic acids research*, 25(5):955–964,
            1997.

[Lew09]   Kelly D Lewis. Web single sign-on authentication using saml. *aInternational
            Journal of Computer Science Issues (IJCSI)*, 1:41–48, August 2009.

[LFC00]   Yu S Lim, David Dugan Feng, and Tom Weidong Cai. A web-based collabora-
            tive system for medical image analysis and diagnosis. In *Selected papers from
            the Pan-Sydney workshop on Visualisation*, volume 2, pages 93–95. Australian
            Computer Society, Inc., 2000.

[LGG⁺18]   Tsung-Yi Lin, Priyal Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Fo-
            cal loss for dense object detection. *IEEE transactions on pattern analysis and
            machine intelligence*, 2018.

[LHA14]   Michael I Love, Wolfgang Huber, and Simon Anders. Moderated estimation
            of fold change and dispersion for rna-seq data with deseq2. *Genome biology*,
            15(12):1–21, 2014.

[LLH09]   Anh H. T. Le, Brent Liu, and H. K. Huang. Integration of computer-aided di-
            agnosis/detection (cad) results in a pacs environment using cad–pacs toolkit
            and dicom sr. *International journal of Computer Assisted Radiology and Surgery*,
            4(4):317–329, 2009.

[LMT10]   Biao Liu, David H Mathews, and Douglas H Turner. Rna pseudoknots: folding
            and finding. *F1000 biology reports*, 2, 2010.

[LP00]   Rune B Lyngsø and Christian NS Pedersen. Rna pseudoknot prediction in
            energy-based models. *Journal of computational biology*, 7(3-4):409–427, 2000.

[M⁺11]   Wes McKinney et al. pandas: a foundational python library for data analysis and statistics. *Python for High Performance and Scientific Computing*, 14(9):1–9, 2011.

[Mar02]   Robert C Martin. *Agile software development: principles, patterns, and practices.* Prentice Hall, 2002.

[MBA13]   Abir Messaoudi, Riadh Bouslimi, and Jalel Akaichi. Indexing medical images based on collaborative experts reports. *International Journal of Computer Applications*, page 9, May 2013.

[MDK06]   Ilias Maglogiannis, Constantinos Delakouridis, and Leonidas Kazatzopoulos. Enabling collaborative medical diagnosis over the internet via peer-to-peer distribution of electronic health records. *journal of Medical Systems*, 30(2):107–116, 2006.

[MEM02]   Peter Mildenberger, Marco Eichelberg, and Eric Martin. Introduction to the dicom standard. *European radiology*, 12(4):920–927, 2002.

[MHSM11]   Mojtaba Mehrara, Po-Chun Hsu, Mehrzad Samadi, and Scott Mahlke. Dynamic parallelization of javascript applications using an ultra-lightweight speculation mechanism. In *IEEE 17th International Symposium on High Performance Computer Architecture*, pages 87–98. IEEE, 2011.

[MJMT11]   Zhen Ma, Renato Natal Jorge, T Mascarenhas, and João Manuel RS Tavares. Novel approach to segment the inner and outer boundaries of the bladder wall in t2-weighted magnetic resonance images. *Annals of biomedical engineering*, 39(8):2287–2297, 2011.

[MJMT13]   Zhen Ma, Renato Natal Jorge, T Mascarenhas, and João Manuel RS Tavares. A level set based algorithm to reconstruct the urinary bladder from multiple views. *Medical Engineering & Physics*, 35(12):1819–1824, 2013.

[MM07]   Irmtraud M Meyer and István Miklós. Simulfold: simultaneously inferring rna structures including pseudoknots, alignments, and trees using a bayesian mcmc framework. *PLoS computational biology*, 3(8):149, 2007.

[MSE⁺15]   Robert J McDonald, Kara M Schwartz, Laurence J Eckel, Felix E Diehn, Christopher H Hunt, Brian J Bartholmai, Bradley J Erickson, and David F Kallmes. The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. *Academic radiology*, 22(9):1191–1198, 2015.

[MSZT99]   David Mathews, Jeffrey Sabina, Michael Zuker, and Douglas Turner. Expanded sequence dependence of thermodynamic parameters improves prediction of rna secondary structure1. *Journal of molecular biology*, 288:911–40, 06 1999.

[MT15]   Scott Millett and Nick Tune. *Patterns, principles, and practices of domain-driven design.* John Wiley & Sons, 2015.

[MV08]   Dagmar Krefting Michal Vossberg, Thomas Tolxdorff. Dicom image communication in globus-based medical grids. *IEEE Transactions on Information Technology in Biomedicine*, 12:145–153, 2008.

[MVV⁺01]   Paul JA Michiels, Alexandra AM Versleijen, Paul W Verlaan, Cornelis WA Pleij, Cornelis W Hilbers, and Hans A Heus. Solution structure of the pseudoknot of srv-1 rna, involved in ribosomal frameshifting. *Journal of molecular biology*, 310(5):1109–1123, 2001.

[MY14]   Jorge Millan and Leonardo Yunda. An open-access web-based medical image atlas for collaborative medical image sharing, processing, web semantic searching and analysis with uses in medical training, research and second opinion of cases. *Nova*, 12(22):143–150, 2014.

[Må06]   Per Måseide. The deep play of medicine: Discursive and collaborative processing of evidence in medical problem solving. *Communication & medicine*, 3(1):43–54, 2006.

[Nes19]   Donka Nesheva. Introduction to health information technologies. *IOP Conference Series: Materials Science and Engineering*, 618:012033, oct 2019.

[New15a]   Sam Newman. *Building Microservices*. O'Reilly Media, Inc., 1st edition, 2015.

[New15b]   Sam Newman. *Building microservices: designing fine-grained systems.* " O'Reilly Media, Inc.", 2015.

[NJ80]   Ruth Nussinov and Ann B Jacobson. Fast algorithm for predicting the secondary structure of single-stranded rna. *Proceedings of the National Academy of Sciences*, 77(11):6309–6313, 1980.

[NKE09]   E. P. Nawrocki, D. L. Kolbe, and S. R. Eddy. Infernal 1.0: inference of rna alignments. *Bioinformatics*, 25(10):1335–1337, 2009.

[NMMA16]   Irakli Nadareishvili, Ronnie Mitra, Matt McLarty, and Mike Amundsen. *Microservice architecture: aligning principles, practices, and culture.* " O'Reilly Media, Inc.", 2016.

[NOKT14]   Atsushi Nambu, Katsura Ozawa, Noriko Kobayashi, and Masao Tago. Imaging of community-acquired pneumonia: Roles of imaging examinations, imaging diagnosis of specific pathogens and discrimination from noninfectious diseases. *World journal of radiology*, 6(10):779, 2014.

[NPGK78]   R. Nussinov, G. Pieczenik, J. R. Griggs, and D. J. Kleitman. Algorithms for loop matchings. *SIAM Journal on Applied mathematics*, 35(1):68–82, 1978.

[NRK⁺02]   Paul L Nixon, Anupama Rangan, Y-G Kim, Alexander Rich, David W Hoffman, Mirko Hennig, and David P Giedroc. Solution structure of a luteoviral p1–p2 frameshifting mrna pseudoknot. *Journal of molecular biology*, 322(3):621–633, 2002.

[Nvi07]   C. Nvidia. Compute unified device architecture programming guide, 2007.

[OM11]   F. Ozsolak and Patrice Milos. Rna sequencing: Advances, challenges and opportunities. *Nat Rev*, 12:87–98, 01 2011.

[PBS⁺06]   J. S. Pedersen, G. Bejerano, A. Siepel, K. Rosenbloom, K. Lindblad-Toh, E. S. Lander, J. Kent, W. Miller, and D. Haussler. Identification and classification of conserved rna secondary structures in the human genome. *PLoS computational biology*, 2:e33, 2006.

[PDK⁺09]   Christos Pavlatos, Alexandros C. Dimopoulos, Andrew Koulouris, Theodore Andronikos, Ioannis Panagopoulos, and George Papakonstantinou. Efficient reconfigurable embedded parsers. *Comput. Lang. Syst. Struct.*, 35(2):196–215, July 2009.

[PMF⁺04]   J. S. Pedersen, I. M. Meyer, R. Forsberg, P. Simmonds, and J. Hein. A comparative method for finding and folding rna secondary structures within protein-coding regions. *Nucleic acids research*, 32(16):4925–4936, 2004.

[PWK⁺20]   Alan W Partin, Alan J Wein, Louis R Kavoussi, Craig A Peters, and Roger R Dmochowski. *Campbell Walsh Wein Urology, E-Book*. Elsevier Health Sciences, 2020.

[PXSY18]   Chen Pan, Wenlong Xu, Dan Shen, and Yong Yang. Leukocyte image segmentation using novel saliency detection based on positive feedback of visual perception. *Journal of healthcare engineering*, 2018, 2018.

[PZA⁺17]   Cesare Pautasso, Olaf Zimmermann, Mike Amundsen, James Lewis, and Nicolai Josuttis. Microservices in practice, part 1: Reality check and service design. *IEEE Software*, (1):91–98, 2017.

[RB13]   ER Ranschaert and FH Binkhuysen. European teleradiology now and in the future: results of an online survey. *Insights into imaging*, 4(1):93–102, 2013.

[RBOC96]   Toolika Rastogi, Tara L Beattie, Joan E Olive, and Richard A Collins. A long-range pseudoknot is required for activity of the neurospora vs ribozyme. *The EMBO journal*, 15(11):2820–2825, 1996.

[RE99]   Elena Rivas and Sean R Eddy. A dynamic programming algorithm for rna structure prediction including pseudoknots. *Journal of molecular biology*, 285(5):2053–2068, 1999.

[RE01]   Eddy S.R. Rivas E. Noncoding rna gene detection using comparative sequence analysis. *BMC Bioinformatics 2*, 8, 2001.

[Rei]   Emily Reinhold. Lessons learned on uber's journey into microservices.

[RG04]   Jens Reeder and Robert Giegerich. Design, implementation and evaluation of a practical pseudoknot folding algorithm based on thermodynamics. *BMC bioinformatics*, 5(1):104, 2004.

[RIZ⁺17]   Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brand on Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, and Andrew Y. Ng. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *arXiv e-prints*, page arXiv:1711.05225, Nov 2017.

[RK]   P. Rui and K. Kang. National ambulatory medical care survey: 2015 emergency department summary tables, table 27. Available at `https://www.cdc.gov/nchs/data/nhamcs/web_tables/2015_ed_web_tables.pdf`. Accessed on: 2019-02-12.

[RML13]   Justin Ritz, Joshua Martin, and Alain Laederach. Evolutionary evidence for alternative structure in rna sequence co-variation. *PLoS computational biology*, 9:e1003152, 07 2013.

[RRCH05]   Jihong Ren, Baharak Rastegari, Anne Condon, and Holger H Hoos. Hotknots: heuristic prediction of rna secondary structures including pseudoknots. *Rna*, 11(10):1494–1504, 2005.

[RRTN06]   Thomas Ruhstaller, Helen Roe, Beat ThÃŒrlimann, and Jonathan J. Nicoll. The multidisciplinary meeting: An indispensable aid to communication between different specialities. *European journal of Cancer*, 42(15):2459 – 2462, 2006.

[RSN]   Rsna pneumonia detection challenge. Available at `https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/`. Accessed on: 2019-02-12.

[RSR04]   Antoine Rosset, Luca Spadola, and Osman Ratib. Osirix: an open-source software for navigating in multidimensional dicom images. *journal of digital imaging*, 17(3):205–216, 2004.

[RSZ04]   Jianhua Ruan, Gary D Stormo, and Weixiong Zhang. An iterated loop matching approach to the prediction of rna secondary structures with pseudoknots. *Bioinformatics*, 20(1):58–66, 2004.

[Ruz78]   Walter Lawrence Ruzzo. *General context-free language recognition*. PhD thesis, Univ. of California, Berkeley, 1978.

[RVPP⁺82]   Krijn Rietveld, R Van Poelgeest, Cornelis WA Pleij, JH Van Boom, and Leendert Bosch. The trna-uke structure at the 3' terminus of turnip yellow mosaic virus rna. differences and similarities with canonical trna. *Nucleic acids research*, 10(6):1929–1946, 1982.

[SB05]   David W Staple and Samuel E Butcher. Pseudoknots: Rna structures with diverse functions. *PLoS Biol*, 3(6):e213, 2005.

[SCF⁺12]   Robert Spitale, Pete Crisalli, Ryan Flynn, Eduardo Torre, Eric Kool, and Howard Chang. Rna shape analysis in living cells. *Nature chemical biology*, 9, 11 2012.

[SdJH+99]   Frederich C Stiefel, Peter de Jonge, Frits J Huyse, Patrice Guex, Joris PJ Slaets, John S Lyons, Jacques Spagnoli, and Marco Vannotti. "intermed": a method to assess health service needs: Ii. results on its validity and clinical use. *General hospital psychiatry*, 21(1):49–56, 1999.

[SDO10]   Piter T. de Boer Sílvia D. Olabarriaga, Tristan Glatard. A virtual laboratory for medical image analysis. *IEEE Transactions on Information Technology in Biomedicine*, 14:979–985, 2010.

[sha]   Healthcare professional's collaboration space (share4health). `https://distrinet.cs.kuleuven.be/research/projects/Share4Health`. Accessed: 2017-11-30.

[Shi14]   Yigong Shi. A glimpse of structural biology through x-ray crystallography. *Cell*, 159:995–1014, 11 2014.

[SHPZ19]   Jaswinder Singh, Jack Hanson, Kuldip Paliwal, and Yaoqi Zhou. Rna secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nature communications*, 10(1):1–13, 2019.

[SIKW+20]   Rodrigo Suarez-Ibarrola, Maximilian Kriegmair, Frank Waldbillig, Britta Grüne, Misgana Negassi, Ujwala Parupalli, Annette Schmitt, Alexander Reiterer, Christoph Müller, Alexander Scheurer, et al. A novel endoimaging system for endoscopic 3d reconstruction in bladder cancer patients. *Minimally Invasive Therapy & Allied Technologies*, pages 1–8, 2020.

[Sip06]   Michael Sipser. *Introduction to the Theory of Computation*, volume 2. Thomson Course Technology Boston, 2006.

[SIVA17]   Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *AAAI*, volume 4, page 12, 2017.

[SJC+19]   Eugene Shkolyar, Xiao Jia, Timothy C Chang, Dharati Trivedi, Kathleen E Mach, Max Q-H Meng, Lei Xing, and Joseph C Liao. Augmented bladder tumor detection using deep learning. *European urology*, 76(6):714–718, 2019.

[SKH+11]   Kengo Sato, Yuki Kato, Michiaki Hamada, Tatsuya Akutsu, and Kiyoshi Asai. Ipknot: fast and accurate prediction of rna secondary structures with pseudoknots using integer programming. *Bioinformatics [ISMB/ECCB]*, 27(13):85–93, 2011.

[SKI+14]   Muhammad Saleem, Maulik R Kamdar, Aftab Iqbal, Shanmukha Sampath, Helena F Deus, and Axel-Cyrille Ngonga Ngomo. Big linked cancer data: Integrating linked tcga and pubmed. *Web Semantics: Science, Services and Agents on the World Wide Web*, 27:34–41, 2014.

[SKSJ20]   E Sandeep Kumar and Pappu Satya Jayadev. Deep learning for clinical deci-
             sion support systems: A review from the panorama of smart healthcare. *Deep
             Learning Techniques for Biomedical and Health Informatics*, pages 79–99, 2020.

[SKV⁺11]   Z. Sukosd, B. Knudsen, M. Vaerum, J. Kjems, and E. S. Andersen. Multithreaded
             comparative rna secondary structure prediction using stochastic context-free
             grammars. *BMC bioinformatics*, 12:103, 2011.

[SLAJ09]   Stuart G Silverman, John R Leyendecker, and E Stephen Amis Jr. What is the
             current role of ct urography and mr urography in the evaluation of the urinary
             tract? *Radiology*, 250(2):309–323, 2009.

[SNH⁺09]   Jérôme Schmid, Niels Nijdam, Seunghyun Han, Jinman Kim, and Nadia
             Magnenat-Thalmann. Interactive segmentation of volumetric medical images
             for collaborative telemedicine. In *3D Physiological Human Workshop*, pages 13–
             24. Springer, 2009.

[Sny93]   Alan Snyder. The essence of objects: Common concepts and terminology. *IEEE
             Software*, 1993.

[SQV⁺14]   Quan Z. Sheng, Xiaoqiang Qiao, Athanasios V. Vasilakos, Claudia Szabo, Scott
             Bourne, and Xiaofei Xu. Web services composition: A decade's overview. *In-
             formation Sciences*, 280(Supplement C):218 – 238, 2014.

[STJ95]   Ling X Shen and Ignacio Tinoco Jr. The structure of an rna pseudoknot that
             causes efficient frameshifting in mouse mammary tumor virus. *Journal of
             molecular biology*, 247(5):963–978, 1995.

[sun]   Synapse 3d, radiology visualization software. `http://3dimaging.`
             `fujimed.com`.

[SZLD17]   QingZeng Song, Lei Zhao, XingKe Luo, and XueChen Dou. Using deep learning
             for classification of lung nodules on computed tomography images. *Journal of
             healthcare engineering*, 2017, 2017.

[TBF05]   Carla A Theimer, Craig A Blois, and Juli Feigon. Structure of the human telom-
             erase rna pseudoknot reveals conserved tertiary interactions essential for func-
             tion. *Molecular cell*, 17(5):671–682, 2005.

[TCGS98]   Jack E. Tabaska, Robert B. Cary, Harold N. Gabow, and Gary D. Stormo. An
             rna folding method capable of identifying pseudoknots and base triples. *Bioin-
             formatics (Oxford, England)*, 14(8):691–699, 1998.

[TH15]   Abdel Aziz Taha and Allan Hanbury. Metrics for evaluating 3d medical image
             segmentation: analysis, selection, and tool. *BMC medical imaging*, 15(1):1–28,
             2015.

[TNV+20]   Ashitosh Tilve, Shrameet Nayak, Saurabh Vernekar, Dhanashri Turi, Pratiksha R Shetgaonkar, and Shailendra Aswale. Pneumonia detection using deep learning approaches. In *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, pages 1–8. IEEE, 2020.

[UKT+19]   Hannes Ulrich, Jori Kern, Deniz Tas, Ann-Kristin Kock-Schoppenhauer, Frank Ückert, Josef Ingenerf, and Martin Lablans. Ql 4 mdr: a graphql query language for iso 11179-based metadata repositories. *BMC medical informatics and decision making*, 19(1):1–7, 2019.

[VACK11]   Oliver Vogel, Ingo Arnold, Arif Chughtai, and Timo Kehrer. *Software architecture: a comprehensive framework and guide for practitioners*. Springer Science & Business Media, 2011.

[VBGP95]   FHD Van Batenburg, Alexander P Gultyaev, and Cornelis WA Pleij. An apl-programmed genetic algorithm for the prediction of rna secondary structure. *Journal of theoretical Biology*, 174(3):269–280, 1995.

[VDMCML+08] Aart J Van Der Molen, Nigel C Cowan, Ullrich G Mueller-Lisse, Claus CA Nolte-Ernsting, Satoru Takahashi, and Richard H Cohan. Ct urography: definition, indications and techniques. a guideline for clinical practice. *European radiology*, 18(1):4–17, 2008.

[Ver13]   Vaughn Vernon. *Implementing domain-driven design*. Addison-Wesley, 2013.

[VG11]   M. Valdiserri and W. Goble. Medical collaboration system and method. `https://www.google.com/patents/US20110238618`, 2011.

[VICBGPn17] Andrea Vázquez-Ingelmo, Juan Cruz-Benito, and Francisco J. García-Peñalvo. Improving the oeeu's data-driven technological ecosystem's interoperability with graphql. In *Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality*, TEEM 2017, New York, NY, USA, 2017. Association for Computing Machinery.

[VNCLO12]   C Vasconcelos-Nóbrega, A Colaco, C Lopes, and PA Oliveira. Bbn as an urothelial carcinogen. *In Vivo*, 26(4):727–739, 2012.

[VRP+12]   Sadhna Verma, Arumugam Rajesh, Srinivasa R Prasad, Krishnanath Gaitonde, Chandana G Lall, Vladimir Mouraviev, Gunjan Aeron, Robert B Bracken, and Kumaresan Sandrasegaran. Urinary bladder cancer: role of mr imaging. *Radiographics*, 32(2):371–387, 2012.

[Wag]   Tim Wagner. Microservices without the servers.

[WB08]   Ligang Wu and Joel Belasco. Let me count the ways: Mechanisms of gene regulation by mirnas and sirnas. *Molecular cell*, 29:1–7, 02 2008.

[WC03]   J.D. Watson and F.H.C. Crick. Molecular structure of nucleic acids. *American Journal of Psychiatry*, 160:623–624, 04 2003.

[WEC+13]   Peng Wang, Olivier Ecabert, Terrence Chen, Michael Wels, Johannes Rieber, Martin Ostermeier, and Dorin Comaniciu. Image-based co-registration of angiography and intravascular ultrasound images. *IEEE transactions on medical imaging*, 32(12):2238–2249, 2013.

[WHS04]   Christina Witwer, Ivo L Hofacker, and Peter F Stadler. Prediction of consensus rna secondary structures including pseudoknots. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 1(2):66–77, 2004.

[WLW+20]   Yili Wang, Yuanning Liu, Shuo Wang, Zhen Liu, Yubing Gao, Hao Zhang, and Liyan Dong. Attfold: Rna secondary structure prediction with pseudoknots based on attention mechanism. *Frontiers in Genetics*, 11:1564, 2020.

[WLZ+19]   Linyu Wang, Yuanning Liu, Xiaodan Zhong, Haiming Liu, Chao Lu, Cong Li, and Hao Zhang. Dmfold: A novel method to predict rna secondary structure with pseudoknots based on deep learning and improved base pair maximization principle. *Frontiers in Genetics*, 10:143, 2019.

[WMB+20]   Tomasz K Wirecki, Katarzyna Merdas, Agata Bernat, Michał J Boniecki, Janusz M Bujnicki, and Filip Stefaniak. RNAProbe: a web server for normalization and analysis of RNA structure probing data. *Nucleic Acids Research*, 48(W1):W292–W299, 06 2020.

[Wol16]   Eberhard Wolff. *Microservices: flexible software architecture*. Addison-Wesley Professional, 2016.

[WPL+17]   Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, pages 3462–3471. IEEE, 2017.

[WWF13]   Randy Weaver, Dawn Weaver, and Dean Farwood. *Guide to network defense and countermeasures*. Cengage Learning, 2013.

[XDZL10]   F. Xia, Y. Dou, D. Zhou, and X. Li. Fine-grained parallel rna secondary structure prediction using scfgs on fpga. *Parallel Computing*, 36(9):516–530, 2010.

[Xie12a]   L. Xie. Adaptive distributed medical image viewing and manipulating systems. `https://www.google.com/patents/US20120084350`, 2012.

[Xie12b]   L. Xie. Adaptive distributed medical image viewing and manipulating systems. `https://www.google.com/patents/US20120084350`, 2012.

[yae]   `https://github.com/vnmakarov/yaep`. Accessed: 2020-03-25.

[You67]   D. H. Younger. Recognition and parsing of context-free languages in n^3. *Information and Control*, 10:189–208, 1967.

[ZAL15]   Dimitar V Zlatev, Emanuela Altobelli, and Joseph C Liao. Advances in imaging technologies in the evaluation of high-grade bladder cancer. *Urologic Clinics*, 42(2):147–157, 2015.

[Zuk00]   Michael Zuker. Calculating nucleic acid secondary structure. *Current opinion in structural biology*, 10:303–10, 07 2000.

[Zuk03]   Michael Zuker. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research*, 31:3406–15, 08 2003.

[ZZMH20]   Liang Zhang, He Zhang, David H. Mathews, and Liang Huang. Threshknot: Thresholded probknot for improved rna secondary structure prediction, 2020.

Glossary

AAI: Authentication & Authorization Infrastructure

AI: Artificial Intelligence

API: Application Programmable Interface, a logical or physical connection between computers or between computer programs

CAD: Computer Aided Diagnosis, systems that assist doctors in the interpretation of medical images

CGDBP: Coarse-grained Dot-Bracket Prediction, an intermediate step on predicting RNA secondary structure

CIS: Carcinoma in Situs, a group of abnormal cells that are found only in the place where they first formed in the body (see left panel)

CSCW: Computer-Supported Cooperative Work, the study of how people utilize technology collaboratively, often towards a shared goal

CT: Computed Tomography, a medical imaging technique used in radiology to obtain detailed internal images of the body noninvasively for diagnostic purposes

CTU: Computed Tomography Urogram, dedicated CT to examine examine the kidneys, ureters and bladder

DDD: Domain Driven Design, an approach to software development that centers the development on programming a domain model that has a rich understanding of the processes and rules of a domain

DICOM: Digital Imaging and Communications in Medicine, a standard for the communication and management of medical imaging information and related data

DIV: DICOM Image Viewer, a core component to assist medical image analyis

DNA: Deoxyribonucleic Acid, a polymer composed of two polynucleotide chains that coil around each other to form a double helix carrying genetic instructions for the development, functioning, growth and reproduction of all known organisms and many viruses

ETCD: Pronounced et-see-dee, an open source, distributed, consistent key-value store for shared configuration, service discovery, and scheduler coordination of distributed systems or clusters of machines

FDA: U.S. Food and Drug Administration

FGDPP: Fine-Grained Dot-Plot Prediction, the final step on predicting RNA secondary structure

GDPR: General Data Protection Regulation

GLSL: OpenGL Shading Language (GLSL), the principal shading language for OpenGL

GUI: Graphical User Interface, a form of user interface that allows users to interact with electronic devices through graphical icons and audio indicator such as primary notation, instead of text-based UIs, typed command labels or text navigation

GraphQL: GraphQL, an open-source data query and manipulation language for APIs, and a runtime for fulfilling queries with existing data

HL7: Health Level Seven International, a non-profit ANSI-accredited standards development organization dedicated to providing standards and solutions that empower global health data interoperability

IPT: Image Processing Toolkit, a set of image processing workflows

IoT: Internet of Things

JWT: JASON Web Token, a proposed Internet standard for creating data with optional signature and/or optional encryption whose payload holds JSON that asserts some number of claims.

MDT: Multidisciplinary Team, a group of people with different functional expertise working toward a common goal

MRI: Magnetic Resonance Imaging, a medical imaging technique used in radiology to form pictures of the anatomy and the physiological processes of the body. MRI scanners use strong magnetic fields, magnetic field gradients, and radio waves to generate images of the organs in the body

MSA: Micro-service Architecture, a logical structure for the design of a software program involving loosely-coupled modular components known as microservices.

MTDM: Multidisciplinary Team Meeting, a meeting of the group of professionals from one or more clinical disciplines who together make decisions regarding recommended treatment of individual patients

NMR: Nuclear magnetic resonance (NMR), a physical phenomenon in which nuclei in a strong constant magnetic field are perturbed by a weak oscillating magnetic field (in the near field) and respond by producing an electromagnetic signal with a frequency characteristic of the magnetic field at the nucleus

PACS: Picture Archiving and Communication System, a medical imaging technology which provides economical storage and convenient access to images from multiple modalities

PAI: Photoacoustic Imaging or optoacoustic imaging, a biomedical imaging modality based on the photoacoustic effect

PET: Positron Emission Tomography, a functional imaging technique that uses radioactive substances known as radiotracers to visualize and measure changes in metabolic processes, and in other physiological activities including blood flow, regional chemical composition, and absorption

POC: Proof of Concept, a subset of the end product to include most of the core functionalities to be delivered.

QoS: Quality of Service, the description or measurement of the overall performance of a service

RNA: Ribonucleic Acid, a nucleic acid present in all living cells that has structural similarities to DNA.

RW: Random Walkers, an optimization methodology

SAML: Security Assertion Markup Language, an open standard for exchanging authentication and authorization data between parties

SOA: Service Oriented Architecture, an architectural style that supports service orientation

SPA: Single Page Application, a web application or website that interacts with the user by dynamically rewriting the current web page with new data from the web server, instead of the default method of a web browser loading entire new pages

SPECT: Single-photon emission computed tomography, a nuclear medicine tomographic imaging technique using gamma rays

SUS: System Usability Scale, a reliable tool for measuring the usability

TT: Transdisciplinary Teams, teams of experts on different domains that aims tackling some serious challenge

UAC: User Access Control

UM: User Management

US: Ultrasound, an imaging method that uses sound waves to produce images of structures within your body.

U-net: A type of convolutional neural network

VBC: Value-Based Care, a framework for restructuring health care systems with the overarching goal of value for patients, with value defined as health outcomes per unit of costs.

WADO: Web Access to DICOM Object, a standard

WFMS: Workflow Management System, a software tool designed to help streamline routine business processes for optimal efficiency

WLC: White Light Cystoscopy

YAEP: Yet Another Early Parser, a concrete and performant open source implementation for the Early Parser Algorithm

pseudoknot: A pseudoknot is a nucleic acid secondary structure containing at least two stem-loop structures in which half of one stem is intercalated between the two halves of another stem.

stem-loop: Intramolecular base pairing, a pattern that can occur in single-stranded RNA

webGL: A JavaScript API for rendering interactive 2D and 3D graphics within any compatible web browser without the use of plug-ins

WebRTC: A free and open-source project providing web browsers and mobile applications with real-time communication via application programming

Journals

- C. Andrikos, G. Rassias, P. Tsanakas and I. Maglogiannis, "An Enhanced Device-Transparent Real-Time Teleconsultation Environment for Radiologists," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 1, pp. 374-386, Jan. 2019, doi: 10.1109/JBHI.2018.2824312.

- Andrikos, Christos, et al. "Knotify: An Efficient Parallel Platform for RNA Pseudoknot Prediction Using Syntactic Pattern Recognition." Methods and Protocols 5.1 (2022): 14.

Conferences

- C. Andrikos, G. Rassias, P. Tsanakas and I. Maglogiannis, "Real-time medical collaboration services over the web," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, 2015, pp. 1393-1396, doi: 10.1109/EMBC.2015.7318629.

- Maglogiannis I., Andrikos C., Rassias G., Tsanakas P. (2017) A DICOM Based Collaborative Platform for Real-Time Medical Teleconsultation on Medical Images. In: Vlamos P. (eds) GeNeDis 2016. Advances in Experimental Medicine and Biology, vol 989. Springer, Cham. https://doi.org/10.1007/978-3-319-57348-9_7

- G. Rassias, C. O. Andrikos, P. Tsanakas and I. Maglogiannis, "Versatile Cloud Collaboration Services for Device-Transparent Medical Imaging Teleconsultations", 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS), Thessaloniki, 2017, pp. 306-311, doi: 10.1109/CBMS.2017.79.

- Andrikos C., Maglogiannis I., Bilalis E., Spyroglou G., Tsanakas P. (2014) An Intelligent Platform for Hosting Medical Collaborative Services. In: Likas A., Blekas K., Kalles D. (eds) Artificial Intelligence: Methods and Applications. SETN 2014. Lecture Notes in Computer Science, vol 8445. Springer, Cham. https://doi.org/10.1007/978-3-319-07064-3_28

- Christos Andrikos et. All, "Location, location, location: Revisiting modeling and exploitation for location-based side channel leakages", nternational Conference on the Theory and Application of Cryptology and Information Security, 2019, Springer

113

- Christos Andrikos, et. Al. "Location-based leakages: New directions in modeling and exploiting", 2017 International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS), 2017, IEEE

APPENDIX A

An embedding based methodology for pneumonia detection on Computed Tomography images

A.1   Background information

Nowadays pneumonia still remains one of the leading causes of fatality worldwide, accounting for over 15% of all deaths of children under five years old. In 2015, almost a million children under the age of five have been reported passing away due to this specific disease, while in the USA alone pneumonia accounts for over 50,000 deaths [dea], a fact that places it on the list of top 10 death causes in the United States of America.

Despite that pneumonia is a common disease, its accurate diagnosis is an arduous task. On one hand the Chest radiograph (aka. Chest X-ray, CXR) examination remains the primary method used for diagnosis, in the most cases it requires to be assessed by highly-trained specialists and confirmation through clinical history, vital signs and laboratory exams.

To further complicate the matter, reading these images is problematic because pneumonia is usually manifested through an area or areas of **increased opacity**[Fra], which is not an obvious rule to follow even for highly skilled domain experts.

The evaluation of CXR is prone to erroneous diagnostic results due to numerous reasons. The positioning of the patient and the depth of inspiration are among the two most common factors able to alter the semantics of the artifacts [Fra].

In addition, clinicians are often faced with interpreting high volume of images during their shifts. Pneumonia has been reported to be the reason for over 500,000 visits to emergency departments in USA back in 2015 [RK].

In this work we suggest a novel analysis pipeline that highlights potential areas of pneumonia. It is based on an ensemble of neural networks leveraging the concept of deep attention. Briefly we split the pulmonary area of an CXR image into tiles that next are classified according to their (1) to their opacity and (2) according to the opacity of the neighboring ones. In that sense we achieve to build an embedding architecture able to "understand" the rule of "increased opacity" discussed in [Fra]. By achieving the remarkable F1 score of 0.487 we argue for the clear contribution of our owr work. In this settings, we move on step forward to incorporate the suggested embedding to the platform Hermes [ARTM18] we have discussed in previous sections.

Recently, deep learning has become very popular in the field of medical image analysis and computer aided diagnosis (CAD). This because it deep learning outperforms any typical approach

when it comes to the semantic segmentation task. Q. Song et al. in China "Using Deep Learning for Classification of Lung Nodules on Computed Tomography Images" employ a convolution neural network (CNN), deep neural network (DNN), and stacked autoencoder (SAE) for the early diagnosis of lung cancer to doctors. The experimental results suggest that the CNN archived the best performance when comparing with the DNN and SAE approaches [SZLD17]. N. D. Kamarudin et al. in Malaysia and Japan "A Fast SVM-Based Tongue's Colour Classification Aided by k-Means Clustering Identifiers and Colour Attributes as Computer-Assisted Tool for Tongue Diagnosis" propose a two-stage classification system for tongue color diagnosis aided with the devised clustering identifiers, and it can diagnose three tongue colors: red, light red, and deep red [KOK+17]. The diagnosis system is very useful for the early detection of imbalance condition inside the body. Experimental result shows that this novel classification system outperforms the conventional SVM by 20% computational time and 15% in terms of classification accuracy.

The challenge of lung disease detection has intrigued the broader community long before the SARS-COV-2 outbreak. Various works have been submitted on that topic in the recent years and some quite interesting datasets are now publicly accessible [JCA+], a fact which is essential for the development of any deep neural network approach. CheXNet [RIZ+17] is maybe one of the newest and perhaps the most inspiring work presenting a deep learning architecture trained on ChestX-ray14 [che], that is able to detect pneumonia at a level comparable to practicing radiologists. The authors in [KDP+15] presented a machine learning approach to classify usual interstitial pneumonia in patients with interstitial lung disease while the work in [] argues that machine learning is able to detect interstitial pneumonia transbronchial biopsies.

Back in 2018 the Radiological Society of North America (RSNA) issued a related challenge for automated Pneumonia Detection [RSN]. The top ranked solution suggested a two-stage classification-detection pipeline combined with a post-processing step. The classification part consisted of an ensemble [1] of 25 models of three different CNN architectures, i.e. Inception-ResNet v2 [SIVA17], Xception [Cho17] and DenseNet169 [HLVDMW17]. Each one of models had been initially trained on the ImageNet dataset, consecutively on the NIH ChestX-ray14 dataset [WPL+17] and finally they have been fine-tuned through various hyper-parameters tweaking (e.g. the number of classes, input size). The detection part was even more complex involving an ensemble of ensembles of RetinaNet [LGG+18], Deformable R-FCN [JHY+17, JYKJ16] and Relation Network [HGZ+17] models trained also across different configurations. Comparatively to the leaders of the challenge board the second and third ranked algorithms kept it more straightforward. They both relied on the RetinaNet architecture and presented an ensemble of 4 and 2 models respectively. They also both made use of heavy augmentation and early-stopping techniques, while maintaining a post-processing step for further fine-tuning the produced bounding boxes. Last but not least, the work in [TNV+20] lists the most recent advances in the field of lung CXR analysis along with a coherent list of the most up to date datasets.

In this work, we introduce a novel approach to pneumonia detection on CXR images. Our approach is based on ensemble of deep learning models yet it remains simple and explainable while performs an accuracy that exceeds 90%. Implementation-wise, the novelty of our work lies into the concept of neighboring tiles classification which aligns with the rule of "areas of "increased opacity" discussed before.

---

[1]Ensemble learning is a general meta approach to machine learning that seeks better predictive performance by combining the predictions from multiple models.
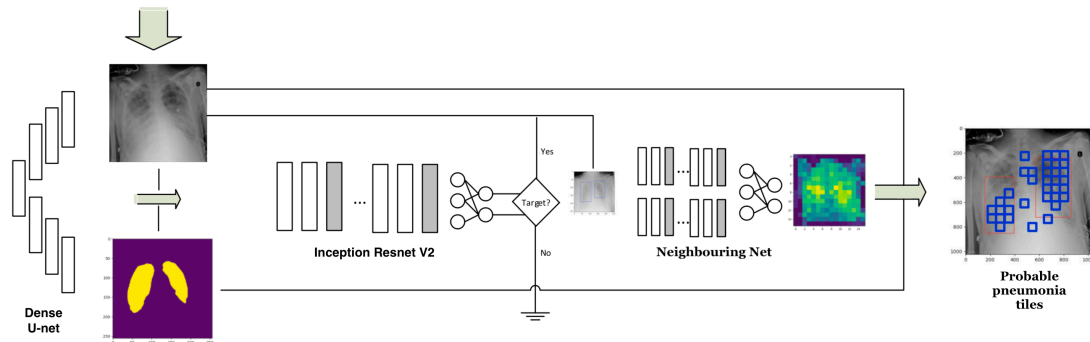
Figure A.1: Pneumonia detection pipeline: 1) lung segmentation, 2) pulmonary area classification, 3) tile classification

A.2 Proposed Methodology

The diagnosis of pneumonia is based on the presence of specific clinical features and is supported by imaging of the lung, usually by chest radiography (CXR). Infection of the lower respiratory tract typically presents radiologically as one of three patterns: (a) focal non-segmental or lobar pneumonia, (b) multi-focal bronchopneumonia or lobular pneumonia, and (c) focal or diffuse "interstitial" pneumonia [NOKT14]. All three cases manifest on CXR images as areas of increased opacity which should be considered to be an inconspicuous pattern.

To tackle the challenge of inconspicuous pattern recognition we introduce following three-step pipeline for pneumonia detection illustrated on Figure A.1: (1) initially images are classified in target and control (pneumonia and non-pneumonia accordingly). Consequently, (2) a custom U-Net [LCQ$^+$18] architecture generates a 2D probability distribution along the lung areas while a custom DNN classifies image segments to target and control ones according to their "level of opacity" providing a second 2D probability mapping. Finally, (3) the lower dimension confidence maps are combined to provide the final 2D probability distribution as follows:

$$P(``tile\_features\_opacity" \cap ``tile\_depicts\_lung\_area")$$

During the first step of the pipeline the CXR image is initially cropped to export the most discriminative pulmonary area. This is done utilizing the U-Net architecture described in the subsection A.2.2. The notion of segmentation in this step is instead of providing lung contours accurately, to generate a bounding box that contains the total thoracic area. The cropped thoracic area is then classified to target or control by the fine-tuned variation of Inception-ResNet v2 [SIVA17]. In case of a target image, the image is pieced in tiles of *64x64* pixels. Every wider area of 192x192 pixels enclosing the tile we focus along with the nine adjacent ones is fed to the Neighbouring Net (subsection A.2.3) which classifies the central tile of the neighborhood considering the opacity distribution of the entire neighborhood. The classification probabilities are structured as a 2D array. The final confidence map is provided by the multiplication of the initial CAM (Class Activation Mapping) provided by the U-Net and that last 2D probability distribution (Figure A.1).
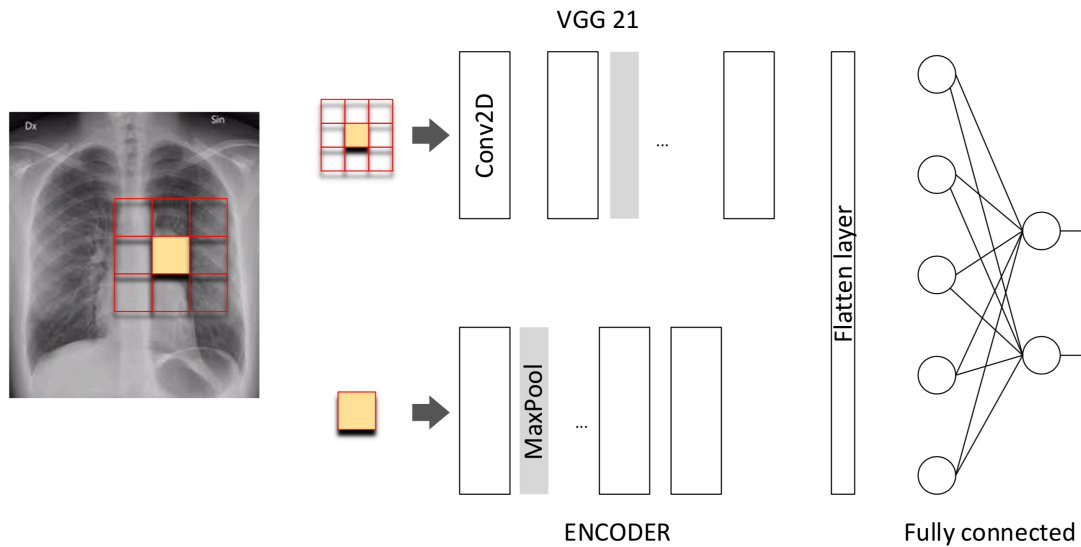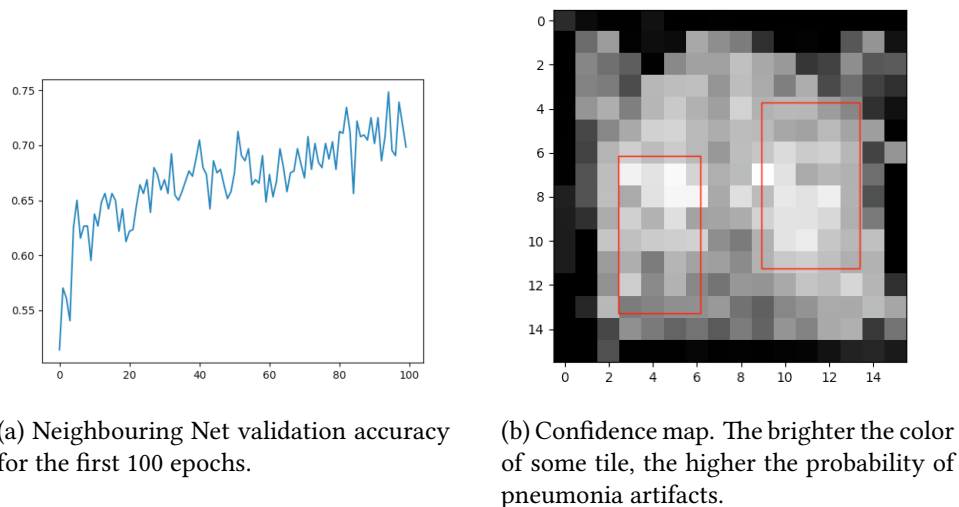
Figure A.2: Neighbouring Net: image is split into tiles that are classified into target and control.

(a) Neighbouring Net validation accuracy for the first 100 epochs.

(b) Confidence map. The brighter the color of some tile, the higher the probability of pneumonia artifacts.

A.2.1   CXR classification

After extended experiments, two of the most successful Convolutional Neural Networks (CNNs) (i.e. Xception and Inception-ResNet v2) have been conscripted to classify the provided imaging data into two classes (i.e. target and control). Other popular architectures (e.g. ResNet, Inception, VGG19) found to fall sort into learning meaningful patterns for that specific task and thus they were abandoned in the first place. Inception-ResNet v2 [SIVA17] was in particular the top performer, excelling in terms of accuracy, even by a close margin compared to the Xception [Cho17]. In this context it was chosen to be further optimized to classify medical images.

The pre-trained version of Inception-ResNet v2 on ImageNet [DDS$^+$09] was further trained on the *ChestX-ray14 dataset*. The latter contains over 120,000 CXR images of tens of thousands of patients which are annotated with 14 distinct pathologies, one of which is pneumonia. This is considered an essential training step due to the fact that RSNA dataset, while being smaller than ChestX-ray14, also suffers from severe data imbalance (i.e. 25% pneumonia cases vs 75% non-pneumonia cases). Finally, the produced model was trained on a heavily elastically augmented version of the RSNA dataset. The whole training process included extended hyperparameter tweaking ending to classification accuracy of **90.68%**.

In order to squeeze out the best possible performance of the given architecture,we trained a separate U-Net model that attempted to segment the wider pulmonary area, as it is depicted in the next section. Classifying only the cropped thoracic area improved the model accuracy to **93.74%** which is arguably a serious contribution.

A.2.2 Lung segmentation

Lung segmentation is considered a core task of the entire prediction pipeline as depicted in Figure A.1. To infer lung boundaries, a U-Net is deployed. It consists of 35 layers (convolutional and max pooling) interconnected in a dense fashion similar to the one introduced in [IMK$^+$14], featuring a 2D convolutional auto-encoder of input (CXR image) and output (lungs mask) size of (256, 256) pixels.

The U-Net was trained on a public CXR dataset generated by Shenzhen Hospital for computer-aided screening of pulmonary diseases [JCA$^+$14] and fine-tuned on the subset of RSNA challenge dataset. The Shenzhen Hospital dataset consisted of 661 CXR images while the corresponding lung masks have been retrieved from an older Kaggle competition on the topic of lung segmentation. It excelled in the particular task in terms of IoU achieving a score of **97.3%**.

A.2.3 Neighbouring Net, a tile classifier

The most difficult concept of the entire task is the way that the pneumonia is visually expressed. What actually means the term *"areas of increased opacity"*? To answer this question, we came up with the following idea: *"an area of increased opacity should be an arbitrary image tile that somehow features increased opacity compared with a set of adjacent tiles"*.

In this context we introduce a novel machine learning approach to tackle that indistinct definition. Figure A.2 depicts the Neighbouring Net, a Deep Neural Network (DNN) that is able to pick image tiles containing some sort of inconspicuous pattern such as the "increased opacity" ones. Neighbouring Net consumes an image subarea enclosing nine adjacent tiles. It consists of two branches that run in parallel to contribute with their corresponding predictions to a common flatten layer which next feeds a fully connected classifier. The final output express the probability that the center tile does feature the "increased opacity" pattern.

In the same figure (A.2), above VGG21 is the exact same well-known VGG19 architecture extended by three extra layers (two convolutional layers and a max-pooling one) while the custom encoder model is the first half (the encoder part) of the auto-encoder architecture described in subsection (A.2.2) but with decreased input dimensionality. That particular architecture focuses on extending the Field of View (FOV) considering the questioned tile as well as the expression power of its close adjacent to infer about the "increased opacity" pattern. The increased FOV does

provide a more concrete classification mechanism due to the exploitation of additional spatial information. Neighbouring Net have been trained on a balanced dataset of arbitrary 192x192 areas that were exported by the original RSNA dataset. Figure A.3a depicts the validation accuracy for the first 100 epochs. It is noteworthy that the observed accuracy oscillation does deteriorate after 100 epochs of training but it is not eliminated.

Regarding the size of the tile and the adjacent area, we have experimented with various setups (larger tiles and adjacent areas as well as smaller ones, i.e. 32x32 and 128x128 px). The results for the alternative experiments did not meet the performance of the one we present. This is a consequence of larger areas producing more misleading features, enclosing hidden correlations. On the contrary, smaller areas, have reduced FOV and thus they provide shallow classification (i.e. the classification will be based on quite rare patterns that do not correlate with the task). Moreover the 64x64 tile was chosen based on a brief inspection of the various annotation sizes provided by the RSNA experts.

subsection(Post processing output): Rank all tiles based on the number of the neighbors −> remove outliers: increased performance −> IOU 0.54

A.3   Discussion

The greatest problem regarding the specific task is that there is no obvious way to model pneumonia on CXR imaging data. An extra level of complexity provides the fuzzy way that specialists annotate pneumonia. The provided boxes due to the predefined rectangular geometry that they feature, lead specialists to enclose non-pneumonia areas and sometimes even non-lung areas into the majority of annotations. To tackle this level level on uncertainty, the proposed methodology leverages the idea of tiles classification to provide a highlight concrete pneumonia detection artifacts into any CXR image. Rather focusing on the pixel wise classification (the equivalent task of segmentation), our work focuses on classifying image tiles that potentially enclose disease artifacts. To evaluate the overall methodology performance we employ the typical classification metrics depicted in table A.1. Since the introduced methodology suggests an embeddings model with a post processing step, table A.1 highlights the accuracy increment coming from the architectural decision we made. Initially, we evaluated the score of the core architecture we introduced, the Neighboring Net, in the first column of the table. The next column highlights the overall performance boosting of ingesting information related to the generic task of pneumonia classification. Finally the last one indicates the performance increase due to a simple algorithmic post processing step that excludes any the outlier tiles (i.e. the tiles that do feature less than three neighbors). Our work achieves an increased F1 score of 0.52 that clearly states its contribution to the research community.

This particular study exhibited some encouraging results on the task of pneumonia detection. However, those do not completely meet our expectations —in fact, to the best of out knowledge, the methods presented so far (including ours) lags behind the ones developed for other disease detection tasks, such as melanoma detection [EKN⁺17]. This is due to two main reasons. Firstly, pneumonia detection requires extra information about patients health. Even experienced specialists face challenges in CXR review due to a number of other conditions in the lungs, such as fluid overload (pulmonary edema), bleeding, volume loss (atelectasis or collapse), lung cancer, post-radiation or surgical changes. Secondly, and this is our future work directive, a more

	Neighboring Net	**Full pipeline**
Recall	0.5826	0.6126
Precision	0.1915	0.4047
Sensitivity	0.775	0.82
F1 score	0.287	0.4874063

Table A.1: Evaluation of pneumonia region of interest extraction (tiles of 64x64 pixels) – a complete comparison proportionally to the embeddings model evolution

concrete modeling is required that can express the first-order logic that experts utilize when diagnosing pneumonia. Such modeling could be developed based on a structured learning approach. That however requires a closer collaboration of medical and computer vision experts than that of a data analysis competition. We argue that accurate pneumonia detection on CXR remains a quite challenging problem and worths the attention of the community.

Considering the widely accepted work in [TH15], we employ a custom flavor of the original metric of the Dice coefficient to communicate a clear performance retrospection of our work. Since we need to evaluate the outcome in terms of the overlapping between the ground truth and the extracted areas we define the following Dice coefficient variation:

$$\frac{1}{|thresholds|} \sum_t \frac{TP(t)}{TP(t) + FP(t) + FN(t)}$$

The threshold values range from 0.4 to 0.75 with a step size of 0.05. In other words, at a threshold of 0.5, a predicted object is considered a "hit" if its intersection over union with a ground truth object is greater than 0.5.

The first team managed to achieve a score of 0.25475 using a variety of ensembles.

It is worth mentioning that despite the long list of contestants specialized in radiology, machine learning and pattern recognition, there were not exhibited any serious breakthroughs as far as the object (pneumonia) detection part is concerned. More precisely the solutions that occupied the top positions of the leaderboard are based on quite complex ensemble architectures some of which do combine dozens of neural network models trained on various image datasets along with the provided one.