



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΕΤΦΥΩΝ ΣΥΣΤΗΜΑΤΩΝ

Αναγνώριση πυρκαγιάς σε πραγματικό χρόνο με
τεχνικές μηχανικής μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

**ΘΕΟΔΩΡΟΥ Σ.
ΤΑΒΟΥΛΑΡΕΑ**

Επιβλέπων: Ανδρέας-Γεώργιος Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2019



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΥΠΟΛΟΓΙΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΕΤΦΥΩΝ ΣΥΣΤΗΜΑΤΩΝ

Αναγνώριση πυρκαγιάς σε πραγματικό χρόνο με
τεχνικές μηχανικής μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

**ΘΕΟΔΩΡΟΥ Σ.
ΤΑΒΟΥΛΑΡΕΑ**

Επιβλέπων: Ανδρέας-Γεώργιος Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 26η Ιουνίου 2019.

.....
Ανδρέας-Γεώργιος Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

.....
Γεώργιος Στάμου
Αν.Καθηγητής Ε.Μ.Π.

.....
Παναγιώτης Τσανάκας
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2019

.....
Θεόδωρος Σ. Ταβουλαρέας

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © (2019) Θεόδωρος Σ. Ταβουλαρέας.
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η παρούσα διατριβή στοχεύει στην αντιμετώπιση του φαινομένου της πυρκαγιάς μέσω της έγκαιρης ανίχνευσης πιθανής εστίας φωτιάς. Το πρόβλημα των συνεχών δασικών πυρκαγιών, κυρίως κατά τους καλοκαιρινούς μήνες, αποτελεί μάλιστα του σύγχρονου κόσμου. Ειδικότερα, η Ελλάδα είναι η πλέον πληγμένη χώρα η οποία έχει θρηνήσει υπεράριθμα θύματα μέσα στον 21ο αιώνα. Ανά τους καιρούς, υπήρξαν πολλές προσπάθειες για την καταπολέμηση αυτού του φαινομένου με αξιολογα αποτελέσματα. Μερικές από αυτές περιλαμβάνουν χρήση αισθητήρων καπνού και εδάφους για την ανίχνευση μικρών εστιών, αλλά και δορυφόρων για την κάλυψη μεγαλύτερων εκτάσεων. Με την πρόοδο της τεχνολογίας όμως, είναι επιτακτική η ανάγκη για την έρευνα της συνεισφοράς της τεχνητής νοημοσύνης στο πρόβλημα αυτό. Τέτοιου είδους τεχνικές περιλαμβάνουν εξαγωγή χαρακτηριστικών και σημείων ενδιαφέροντος από εικόνες και βίντεο, καθώς και χρήση βαθύτερων αρχιτεκτονικών, όπως είναι τα σύγχρονα συνελικτικά νευρωνικά δίκτυα, με τα αποτελέσματα αυτών των προσεγγίσεων να είναι άκρως ενθαρρυντικά. Στην παρούσα διπλωματική εργασία ερευνώνται όλοι οι πιθανοί τρόποι προσέγγισης του προβλήματος της ανίχνευσης της φωτιάς, από μηχανές διανυσμάτων υποστήριξης μέχρι βαθιά συνελικτικά νευρωνικά δίκτυα, με τα αποτελέσματα των οποίων να βρίσκονται πολύ κοντά σε αυτά των σύγχρονων μεθόδων της βιβλιογραφίας, αλλά και σε ορισμένες περιπτώσεις μάλιστα να τα υπερνικούν. Τέλος, προτείνεται ένα σύστημα εντοπισμού πιθανής εστίας πυρκαγιάς σχεδιασμένο ώστε να λειτουργεί σε πραγματικό χρόνο, ενώ παρουσιάζονται τρόποι με τους οποίους μπορεί αυτό το μοντέλο να θέσει τις βάσεις για τη δημιουργία ενός καινοτόμου συστήματος ανίχνευσης φωτιάς μέσω της αξιοποίησης ήδη υπάρχοντων πόρων.

Λέξεις Κλειδιά

Συνελικτικά Νευρωνικά Δίκτυα, Μηχανική Μάθηση, Βαθιά Μάθηση, Μεταφορά Μάθησης, Όραση Υπολογιστών, Αναγνώριση Πυρκαγιάς, Αναγνώριση Καπνού, Αναγνώριση Δασικής Πυρκαγιάς σε Πραγματικό Χρόνο

Abstract

This dissertation aims at dealing with the fire phenomenon through the early detection of a possible fire spot. The problem of continual forest fires, especially during the summer months, is a puzzler of the modern world. In particular, Greece is the most affected country, which has sacrificed numerous victims in the 21st century. Over the years, there have been many attempts to prevent this phenomenon with remarkable results. Some of these include the use of smoke and ground sensors to detect small fires and satellites to cover larger areas. However, with the advancement of technology, it is necessary to investigate the contribution of artificial intelligence to this problem. Such techniques include the extraction of features and points of interest from images and videos, as well as the use of deeper architectures, such as modern convolutional neural networks (CNN), with the results of these approaches being particularly encouraging. In this diploma thesis, we investigate all possible ways of approaching the problem of fire detection, from supporting vector machines (SVM) to deep convolutional neural networks, the results of which are very close to those of the state-of-the-art methods in the literature, and in some cases, they outperform them. Finally, a fire detection system designed to operate in real time is proposed and the ways of how this model can lay the foundations to create an innovative fire detection system by exploiting the already available resources are presented.

Key Words

Convolutional Neural Networks, Machine Learning, Deep Learning, Transfer Learning, Computer Vision, Fire Detection, Smoke Detection, Real-Time Forest Fire Detection

Ευχαριστίες

Πρώτα απ'όλα, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Ανδρέα-Γεώργιο Σταφυλοπάτη, καθώς και τον κ. Γεώργιο Σιόλα για την καθοδήγησή τους κατά τη διεξαγωγή της διπλωματικής μου. Υπήρξαν πραγματική έμπνευση για μένα για να συνεχίσω τα όνειρά μου στην έρευνα.

Επιπρόσθετα, θα ήθελα να ευχαριστήσω τον κ. Γεώργιο Στάμου και τον κ. Παναγιώτη Τσανάκα για την συμμετοχή τους στην επιτροπή εξέτασης της διπλωματικής μου καθώς και για τα υψηλού επιπέδου μαθήματα που μας προσέφεραν στο Εθνικό Μετσόβιο Πολυτεχνείο (ΕΜΠ).

Στη συνέχεια, ένα μεγάλο ευχαριστώ στον κ. Γεώργιο Λυμπερόπουλο, επικεφαλή του τμήματος R&D του ομίλου ΟΤΕ για την βοήθεια που μού παρείχε όλους αυτούς τους μήνες, καθώς και για τη συμβολή του ώστε η διπλωματική να ενταχθεί σε πραγματική εφαρμογή.

Θα ήθελα επίσης να εκφράσω την ευγνωμοσύνη μου στους γονείς μου που με στήριξαν τόσο πνευματικά όσο και οικονομικά κατά την διάρκεια του ακαδημαϊκού μου ταξιδιού. Χωρίς τη βοήθειά τους και την απεριόριστη υποστήριξη τους, αυτό το ταξίδι δεν θα ήταν δυνατό.

Τέλος, θα ήθελα να ευχαριστήσω όλους τους φίλους μου για το χρόνο που περάσαμε μαζί και την ενθάρρυνση που μού παρείχαν όλα αυτά τα χρόνια για να επιτύχω τους στόχους μου.

Στους γονείς μου,

Θεόδωρος Σ. Ταβουλαρέας

Αθήνα, Ιούνιος 2019

Περιεχόμενα

Περίληψη	5
Abstract	7
Ευχαριστίες	9
Περιεχόμενα	11
Κατάλογος Σχημάτων	13
Κατάλογος Πινάκων	15
1 Εισαγωγή	17
1.1 Κίνητρο	17
1.2 Αντικείμενο της Διπλωματικής	18
1.2.1 Το Πρόβλημα	18
1.2.2 Συνεισφορά	19
1.3 Δομή της Εργασίας	20
2 Αντιμετώπιση Φωτιάς	21
2.1 Φωτιά: Το Φαινόμενο	21
2.1.1 Σύσταση	22
2.1.2 Αστικό Περιβάλλον	24
2.1.3 Δασικό Περιβάλλον	26
2.1.4 Φωτιά και Άνθρωπος: Από τον Προμηθέα στο Σήμερα	30
2.2 Ανάγκη για Εξέλιξη	31
3 Μηχανική Μάθηση	33
3.1 Ορισμός	33
3.2 Μάθηση με Επίβλεψη	33
3.3 Μάθηση χωρίς Επίβλεψη	34
3.4 Νευρωνικά Δίκτυα	35
3.5 Γραμμική Ταξινόμηση	36
3.5.1 Μηχανές Διανυσμάτων Υποστήριξης	36
3.5.2 Λογιστική Παλινδρόμηση	38
3.6 Ταξινόμηση Εικόνων	39
3.6.1 Επίπεδα Συνελικτικών Νευρωνικών Δικτύων	40
3.6.2 Εξαγωγή Χαρακτηριστικών	41
3.6.3 Συνήθεις Αρχιτεκτονικές Συνελικτικών Νευρωνικών Δικτύων	41

3.7	Μεταφορά Μάθησης	43
4	Σχετικές Εργασίες	46
4.1	Ανίχνευση Φωτιάς	46
4.1.1	Κλασικές Προσεγγίσεις	46
4.1.2	Τεχνικές με Βαθιά Μάθηση	55
4.2	Ανίχνευση Καπνού	62
4.3	Real-Time Ανίχνευση	70
5	Βάσεις Δεδομένων	74
5.1	Corsican Database	74
5.1.1	Προέλευση των Εικόνων	74
5.1.2	Περιγραφητές των Εικόνων	75
5.2	MIVIA Dataset	76
5.2.1	Φωτιά	76
5.2.2	Καπνός	76
5.3	Λοιπά Datasets	76
6	Προτεινόμενο Σύστημα	81
6.1	Χαρακτηριστικά Εικόνων	81
6.1.1	Είδη Χαρακτηριστικών	81
6.2	Δημιουργία Βάσης	82
6.2.1	Εξαγωγή Χαρακτηριστικών	83
6.2.2	Περιγραφή του Dataset	83
6.3	Προσέγγιση με ΣΝΔ	84
6.3.1	Αρχιτεκτονική του Συστήματος	84
6.3.2	Περιγραφή του Πειράματος	85
6.4	Real-Time	85
6.5	Εργαλεία Υλοποίησης	86
6.6	Αποτελέσματα	87
7	Εφαρμογή	91
7.1	Κάμερες	91
7.2	Επεξεργασία Δεδομένων	91
7.2.1	Επεξεργασία σε Κεντρικό Server	91
7.2.2	Επεξεργασίας στο Νέφος	92
8	Συμπεράσματα	93
8.1	Αξιολόγηση	93
8.2	Μελλοντικές Επεκτάσεις	93
	Βιβλιογραφία	95

Κατάλογος Σχημάτων

1.1	Διαφορετικές Απεικονίσεις Φωτιάς	19
2.1	Το τετράπτυχο της φωτιάς	22
2.2	Επίπεδα Χρώματος της Φωτιάς	24
2.3	Αφίσα της Αμερικάνικης εκστρατείας για την πρόληψη ενάντια στις πυρκαγιές	27
2.4	Οι πυρκαγιές στη Νότια Ελλάδα την 15η Αυγούστου 2007 (από δορυφόρο της NASA)	30
2.5	Οι καμμένες περιοχές (25 Αυγούστου - 5 Σεπτέμβρη 2007)	30
2.6	Μάτι Ατικής πριν και μετά την καταστροφική πυρκαγιά του 2018	32
3.1	Κάτοψη Νευρωνικού Δικτύου Πρόσθιας Τροφοδότησης	35
3.2	Συνήθεις Συναρτήσεις Ενεργοποίησης	36
3.3	Παράδειγμα δυαδικού διαχωρισμού δύο γραμμικά διαχωρίσιμων κλάσεων.	37
3.4	Εξαγωγή χαρακτηριστικών από CNN	41
3.5	Μέθοδος Μεταφοράς Μάθησης	44
3.6	Οι 3 τρόποι με τους οποίους η μεταφορά μάθησης βελτιώνει την μάθηση	44
4.1	Παραδείγματα ανίχνευσης φωτιάς σε patches	48
4.2	Δομή του δικτύου για patches	49
4.3	Η συνολική δομή του δικτύου: Ο τοπικός ταξινομητής εκπαιδεύεται με τα χαρακτηριστικά που εξάγονται από το μικρό μη γραμμικό νευρωνικό δίκτυο (πράσινο πλαίσιο)	51
4.4	Το σύστημα ανίχνευσης φωτιάς των Jiang et al.	52
4.5	Αρχιτεκτονική του δικτύου: Το πρώτο μέρος αποτελείται από 6 στρώματα: 1 έως 6 (στρώματα συνένωσης και max pooling), το δεύτερο μέρος είναι τα πλήρως συνδεδεμένα στρώματα	56
4.6	VGG16 Architecture	60
4.7	Modified VGG16 Architecture	60
4.8	ResNet50 Architecture	60
4.9	Modified ResNet50 Architecture	61
4.10	Παραδείγματα εικόνων με φωτιά	62
4.11	Παραδείγματα εικόνων χωρίς φωτιά δύσκολες να κατηγοριοποιηθούν	62
4.12	Συνελικτικό νευρωνικό δίκτυο για αναγνώριση καπνού	65
4.13	Framework για αναγνώριση καπνού σε περιβάλλον IoT	67
4.14	Αρχιτεκτονική του CNN. Ο ταξινομητής Softmax παρέχει τις τελικές προβλέψεις για τις τέσσερις προβλεπόμενες κατηγορίες, όπως "Smoke", "Smoke with fog", "Non-Smoke" και "Non-Smoke with fog".	68
4.15	Διαφορά μεταξύ 2D και 3D συνελικτικού φίλτρου	69
4.16	Προτεινόμενος αλγόριθμος ανίχνευσης φωτιάς των Muhammad et al.	72

5.1	Παραδείγματα εικόνων τραβηγμένες με multi-spectral camera	75
5.2	Παραδείγματα βίντεο φωτιάς του MIVIA Dataset	77
5.3	Παραδείγματα βίντεο καπνού του MIVIA Dataset	78
5.4	Παραδείγματα βίντεο φωτιάς του Firesense dataset	79
5.5	Παραδείγματα βίντεο καπνού του Firesense dataset	80
6.1	Χρωματικός χώρος LAB	83
6.2	Παραδείγματα εικόνων με και χωρίς πυρκαγιά του dataset για τη λύση του συστήματος βάσης	84
6.3	Η αρχιτεκτονική του δικτύου CNN-SVM	85
6.4	Η αρχιτεκτονική του δικτύου Inception-V1	86
6.5	Παράδειγμα εξόδου του συστήματος ανίχνευσης πυρκαγιάς σε πραγματικό χρόνο	88
6.6	Παράδειγμα εξόδου του συστήματος ανίχνευσης πυρκαγιάς σε πραγματικό χρόνο	89
6.7	Παράδειγμα εξόδου του συστήματος ανίχνευσης πυρκαγιάς σε πραγματικό χρόνο	90
7.1	Προτεινόμενο σύστημα παρακολούθησης και ανίχνευσης φωτιάς	92

Κατάλογος Πινάκων

2.1	Οι 15 πιο θανατηφόρες πυρκαγιές της σύγχρονης ιστορίας	31
4.1	Περιγραφή του dataset των Zhang et al.	47
4.2	Σύγκριση των αποτελεσμάτων των τεσσάρων αρχιτεκτονικών σύμφωνα με τους Sharma et al.	62
4.3	Σύγκριση των αποτελεσμάτων των τριών αρχιτεκτονικών σύμφωνα με τους Khan et al.	68
6.1	Σύγκριση των αποτελεσμάτων του προτεινόμενου συστήματος βάσης με αυτό των Jiang et al.	87
6.2	Result metrics για τα δύο CNN	88

Κεφάλαιο 1

Εισαγωγή

1.1 Κίνητρο

Ένας πατέρας θέλοντας να αξιολογήσει την εξυπνάδα του παιδιού του, ζήτησε από αυτό να γεμίσει το κουβαδάκι του, φροντίζοντας όμως να κάνει κάποιες τρύπες στη βάση του, με τη βοήθεια ενός μικρότερου σκεύους. Λογικά το παιδί ανάλογα με την ηλικία του και την εξυπνάδα του, αργά ή γρήγορα θα καταλάβει ότι κάτι δεν πάει καλά, όσο προσπάθεια και να καταβάλει, το κουβαδάκι δε θα γεμίσει με νερό. Θα αρχίσει, λοιπόν, να ψάχνει και τελικά, θα βρει την αιτία και είτε θα ζητήσει από το γονιό του να κλείσει τις τρύπες ή θα παρατήσει τη δοκιμασία.

Ως κοινωνία λοιπόν, θέλουμε και απαιτούμε να μην καίγονται τα δασικά μας οικοσυστήματα. Έτσι, κάθε χρόνο ξοδεύονται όλο και περισσότερα χρήματα, για να δημιουργηθεί ένα άριστο σύστημα κατάσβεσης δασικών πυρκαγιών. Το αποτέλεσμα δεν είναι αυτό που θα περιμέναμε. Αυτό που συμβαίνει είναι ότι, αντί να εξαλειφθεί το πρόβλημα των δασικών πυρκαγιών, έχουμε σε κάποιες χρονιές (π.χ. 2007) μεγάλες δασικές πυρκαγιές που το σύστημα μας δεν μπορεί να τις αντιμετωπίσει, καθώς και πολλές χαμένες ανθρώπινες ζωές (Μάτι, 2018). Για το λόγο αυτό, είναι απαραίτητη η δημιουργία ενός έξυπνου συστήματος με χρήση Μηχανικής Μάθησης προκειμένου να αποφευχθούν αντίστοιχα γεγονότα στο μέλλον.

Την τελευταία δεκαετία, υπήρξε ταχεία ανάπτυξη στην απόδοση διαφόρων σύνθετων εργασιών στον τομέα της Όρασης Υπολογιστών, και γενικά της Μηχανικής Μάθησης, όπως η αναγνώριση αντικειμένων. Οι κύριοι λόγοι αυτής της ανάπτυξης ήταν η διαθεσιμότητα λογισμικού και υλισμικού για μαζικά παράλληλους υπολογισμούς, που χρησιμοποιήθηκαν για την βελτιστοποίηση σύνθετων μοντέλων μηχανικής μάθησης, γνωστά ως Βαθιά Νευρωνικά Δίκτυα (ΒΝΔ), μαζί με την επινόηση νέων ΒΝΔ αρχιτεκτονικών. Τα αντίστοιχα ΒΝΔ για δεδομένα από εικόνες είναι τα Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ). Τα ΣΝΔ βελτιστοποιούνται - εκπαιδεύονται χρησιμοποιώντας επισημασμένα σύνολα δεδομένων. Τα ΣΝΔ θεωρούνται συνήθως συστήματα μαύρου κουτιού, που δέχονται ως είσοδο εικόνα και δίνουν κατηγορική ή ποσοτική έξοδο. Μέσα σε αυτά τα μαύρα κουτιά, υπολογίζονται ενδιάμεσοι μετασχηματισμοί της εικόνας, καλύπτοντας ολόκληρο το γεωμετρικό-σημασιολογικό φάσμα της εικόνας. Σε αυτή τη διπλωματική εργασία, χρησιμοποιείται αυτή η αρχιτεκτονική για την αναγνώριση δασικής πυρκαγιάς και καπνού.

Η αναγνώριση και παρακολούθηση εστίας πυρκαγιάς αποτελείται από τα μέρη του εντοπισμού και αναγνώρισης ενός συνόλου παραδειγμάτων εικόνων από ροές δεδομένων, όπως δεδομένα από ραντάρ, βίντεο και εικόνες. Η ανάλυση βίντεο από εναέρια

κάμερα είναι μια κατάλληλη εφαρμογή για αναγνώριση πυρκαγιάς, που μπορεί να χρησιμοποιηθεί ως βοήθεια στην κατάσβεση δασικών πυρκαγιών. Η τρέχουσα τάση στην αναγνώριση αντικειμένων με όραση υπολογιστών βασίζεται στη σύγκριση δειγμάτων μεταξύ καρέ χρησιμοποιώντας μετρικές ομοιότητας, η οποία υιοθετείται σε αυτή την εργασία.

1.2 Αντικείμενο της Διπλωματικής

1.2.1 Το Πρόβλημα

Οι δασικές καταστροφές προκαλούν πάντα μεγάλες απώλειες για τον άνθρωπο. Ως εκ τούτου, τα ισχυρά συστήματα συναγερμού πυρκαγιάς έγκαιρης προειδοποίησης χρειάζονται πάντοτε για την προστασία της ασφάλειας και των περιουσιών των ανθρώπων. Μια γρήγορη και ακριβής μέθοδος ανίχνευσης πυρκαγιάς είναι η τεχνολογία που απαιτείται σε τέτοια συστήματα. Ως αποτέλεσμα, η ανίχνευση πυρκαγιάς ή φλόγας έχει λάβει μεγάλη προσοχή από την ερευνητική κοινότητα εδώ και δεκαετίες [1],[2]. Τα τελευταία χρόνια, η ανίχνευση πυρκαγιάς βρίσκει επίσης εφαρμογές στην ανάκτηση γεγονότων σε πυρκαγιά σε αρχεία ψηφιακής εικόνας ή βίντεο [3].

Οι υπάρχουσες μέθοδοι ανίχνευσης πυρκαγιάς μπορούν να ταξινομηθούν σε δύο κατηγορίες:

- μεθόδους βασισμένες σε αισθητήρες.
- μεθόδους βασισμένες σε όραση υπολογιστών.

Οι μέθοδοι που βασίζονται σε αισθητήρες χρησιμοποιούν πληροφορίες που συλλαμβάνονται από ειδικά όργανα όπως αισθητήρες υπέρυθρης ακτινοβολίας και αισθητήρες καπνού, οι οποίοι είναι δαπανηροί και δεν είναι εύκολοι να αποκτηθούν [4]. Επιπλέον, αυτοί οι αισθητήρες τυπικά ανιχνεύουν την παρουσία ορισμένων σωματιδίων που παράγονται από τον ιονισμό φωτιάς ή φωτομετρία και όχι εξαιτίας της ίδιας της καύσης. Επομένως, τα συστήματα συναγερμού πυρκαγιάς που βασίζονται σε αυτούς τους αισθητήρες παρουσιάζουν συνήθως χρονική καθυστέρηση και έχουν λανθασμένα αποτελέσματα [5]. Αντίθετα, οι μέθοδοι βασισμένες σε όραση χρησιμοποιούν τις κανονικές εικόνες RGB ή τα βίντεο και έχουν πολλά πλεονεκτήματα [6]. Πρώτον, οι κάμερες καθίστανται πιο λειτουργικές με τη σταθερή πτώση των τιμών. Έτσι, οι εικόνες και τα βίντεο μπορούν να παρέχουν πιο λεπτομερείς οπτικές πληροφορίες για την ανίχνευση πυρκαγιάς. Επιπλέον, οι κάμερες παρακολούθησης που είναι ήδη εγκατεστημένες σε δημόσιους χώρους μπορούν επίσης να χρησιμοποιηθούν για ανίχνευση πυρκαγιάς. Δεύτερον, ο χρόνος απόκρισης μπορεί να είναι γρηγορότερος από τους παραδοσιακούς αισθητήρες, καθώς οι κάμερες δεν χρειάζεται να περιμένουν να διαχυθεί ο καπνός ή η θερμότητα. Τέλος, σε σύγκριση με τους παραδοσιακούς αισθητήρες σημείου, οι κάμερες μπορούν να παρακολουθήσουν μια ευρύτερη περιοχή, δημιουργώντας μεγαλύτερη πιθανότητα ανίχνευσης πυρκαγιάς σε πρώιμο στάδιο. Το σύστημα που παρουσιάζεται σε αυτή την εργασία ανήκει στην τελευταία κατηγορία και αποσκοπεί στην ανίχνευση πυρκαγιάς σε βίντεο κλιπ καθώς και σε πραγματικό χρόνο.

Παρά τις αυξανόμενες ανάγκες και ενδιαφέροντα στην ανίχνευση πυρκαγιάς, εξακολουθεί να μην υπάρχει μεγάλος αριθμός εργασιών για το συγκεκριμένο πρόβλημα στη βιβλιογραφία της όρασης υπολογιστών. Η οικοδόμηση ενός ισχυρού συστήματος ανίχνευσης πυρκαγιάς είναι δύσκολη ως προς δύο πτυχές: (1) η φωτιά ή η φλόγα είναι

εύκαμπτη σε σχήμα και ένταση και δεν έχει σταθερή δομή ή εμφάνιση. Αν και το χρώμα είναι ένα σχετικά ξεχωριστό χαρακτηριστικό που χρησιμοποιείται ευρέως για την ανίχνευση πυρκαγιάς, εμφανίζεται στις εικόνες και τα βίντεο επηρεάζονται από την ποιότητα της κάμερας (π.χ. ανάλυση, ευκρίνεια) και ρυθμίσεις (π.χ. ισορροπία λευκού). Η φωτιά εμφανίζεται επίσης ευρέως σε κλίμακα. Το Σχήμα 1.1 δείχνει έξι εικόνες πυρκαγιάς με διαφορετικά το σχήμα, την ένταση, το χρώμα και την κλίμακα.



Σχήμα 1.1: Διαφορετικές Απεικονίσεις Φωτιάς

Έτσι, είναι δύσκολο να υλοποιηθεί ένα γενικό και ακριβές μοντέλο πυρκαγιάς. Και (2) η πραγματική εφαρμογή απαιτεί μια γρήγορη μέθοδο ανίχνευσης πυρκαγιάς που μπορεί να λειτουργήσει σε πραγματικό χρόνο. Οι πολύπλοκες μέθοδοι ή μοντέλα δεν μπορούν να βρουν μια θέση σε πραγματικές εφαρμογές. Ορισμένες μέθοδοι ανίχνευσης πυρκαγιάς που χρησιμοποιούνται στις τράπεζες φίλτρων [1], οι μετασχηματισμοί συχνότητας [7] και οι τεχνικές ανίχνευσης κίνησης χρειάζονται περισσότερο χρόνο επεξεργασίας υπολογιστών και δεν είναι κατάλληλοι για εφαρμογή σε πραγματικό χρόνο.

1.2.2 Συνεισφορά

Οι κύρια συνεισφορά αυτής της διπλωματικής εργασίας είναι η χρήση state-of-the-art μεθόδων μηχανικής μάθησης και όρασης υπολογιστών ώστε να κατασκευαστεί ένα σύγχρονο και λειτουργικό σύστημα ανίχνευσης πιθανής εστίας πυρκαγιάς.

Ξεκινώντας από τις κλασικές μεθόδους της όρασης υπολογιστών, όπως είναι η εξαγωγή χαρακτηριστικών μικρής και μεγάλης κλίμακας από εικόνες, αλλά και η χρήση περιγραφητών μέχρι την υλοποίηση αρχιτεκτονικών βαθειών συνελικτικών νευρωνικών δικτύων, οι τεχνικές που εφαρμόζονται στην παρούσα διπλωματική εργασία πετυχαίνουν παρόμοια αποτελέσματα με αυτά των σύγχρονων μεθόδων που παρουσιάζονται στη βιβλιογραφία, και μάλιστα σε ορισμένες περιπτώσεις υπερτερούν.

Τα αποτελέσματα αυτά είναι πολύ ενθαρρυντικά, αφού προσφέρουν μία επαρκή λύση στο χρόνιο πρόβλημα των πυρκαγιών στην Ελλάδα, και κυρίως στην αποτροπή νέων θυμάτων. Το σύστημα που προτείνεται στα πλαίσια της διπλωματικής έχει και πραγματική εφαρμογή. Για παράδειγμα, ο όμιλος ΟΤΕ διαθέτει πλήθος πύργων και κεραιών σε ολόκληρη την ηπειρωτική Ελλάδα όπου θα μπορούσαν να τοποθετηθούν ειδικές κάμερες (π.χ. GoPro) συνδεδεμένες με κάποια διάταξη επιπέδου hardware, στην οποία το σύστημα θα πραγματοποιεί ανίχνευση καπνού ή φωτιάς σε πραγματικό χρόνο.

Αξίζει να αναφερθεί στο σημείο αυτό ότι στην υλοποίηση και διεκπεραίωση της διπλωματικής αυτής συνέβαλε το τμήμα Έρευνας και Ανάπτυξης της Cosmote.

1.3 Δομή της Εργασίας

Η παρούσα διπλωματική εργασία είναι οργανωμένη σε οκτώ κεφάλαια. Στο Κεφάλαιο 2 παρουσιάζεται το πρόβλημα των πυρκαγιών τόσο σε δασικό όσο και αστικό περιβάλλον, καθώς και τρόποι πρόληψης έως τώρα.

Στο Κεφάλαιο 3 περιγράφονται τα Συνελικτικά Νευρωνικά Δίκτυα, ένας τύπος βαθιών μοντέλων μηχανικής μάθησης με επίβλεψη. Παρουσιάζεται το θεωρητικό υπόβαθρο αλλά και οι εφαρμογές τους στον τομέα της όρασης υπολογιστών.

Το Κεφάλαιο 4 ορίζει τις εργασίες που έχουν γίνει πάνω στο αντικείμενο της διπλωματικής.

Στο Κεφάλαιο 5 παρατίθενται οι βάσεις δεδομένων (εικόνες, βίντεο, ακολουθίες εικόνων) που χρησιμοποιήθηκαν στην εκπαίδευση και στην αξιολόγηση των μοντέλων.

Το Κεφάλαιο 6 αναλύει τα μοντέλα και τους αλγόριθμους που χρησιμοποιήθηκαν για την επίλυση του προβλήματος. Επίσης, παρουσιάζονται τα πειραματικά αποτελέσματα.

Στο Κεφάλαιο 7 παρουσιάζονται πιθανές συνδέσεις της διπλωματικής με πραγματικές εφαρμογές.

Τέλος, το Κεφάλαιο 8 περιλαμβάνει τα συμπεράσματα της εργασίας. Επίσης, συζητούνται πιθανές κατευθύνσεις επέκτασης και βελτιστοποίησης, καθώς και μελλοντικές εργασίες.

Κεφάλαιο 2

Αντιμετώπιση Φωτιάς

2.1 Φωτιά: Το Φαινόμενο

Η πυρκαγιά είναι η ταχεία οξειδωση ενός υλικού στην εξωθερμική χημική διαδικασία καύσης, απελευθερώνοντας θερμότητα, φως και διάφορα προϊόντα αντίδρασης. Οι πιο αργές οξειδωτικές διεργασίες όπως η σκουριά ή η πέψη δεν περιλαμβάνονται στον ορισμό αυτό.

Η φωτιά είναι θερμή επειδή η μετατροπή του ασθενούς διπλού δεσμού στο μοριακό οξυγόνο, O_2 , στους ισχυρότερους δεσμούς στα προϊόντα καύσης διοξειδίου του άνθρακα και νερού απελευθερώνει ενέργεια (418kJ ανά 32g O_2). Οι ενεργειακές δεσμεύσεις του καυσίμου παίζουν μόνο έναν δευτερεύοντα ρόλο εδώ [8]. Σε κάποιο σημείο της αντίδρασης καύσης, που ονομάζεται σημείο ανάφλεξης, παράγονται φλόγες. Η φλόγα είναι το ορατό τμήμα της φωτιάς. Οι φλόγες αποτελούνται κυρίως από διοξείδιο του άνθρακα, υδρατμούς, οξυγόνο και άζωτο. Εάν είναι αρκετά ζεστά, τα αέρια μπορεί να ιονιστούν για να παράγουν πλάσμα. Ανάλογα με τις επικίνδυνες ουσίες και τυχόν εξωτερικές ακαθαρσίες, το χρώμα της φλόγας και η ένταση της φωτιάς θα είναι διαφορετικές.

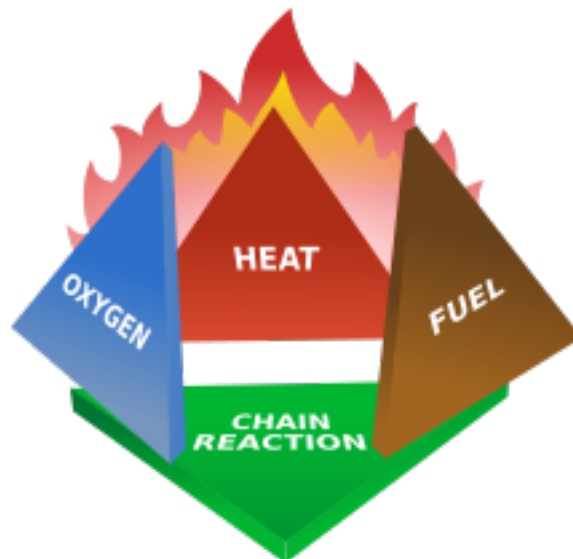
Η φωτιά στην πιο κοινή μορφή της μπορεί να προκαλέσει πυρκαγιά, η οποία έχει τη δυνατότητα να προκαλέσει φυσική βλάβη από την καύση. Η πυρκαγιά είναι μια σημαντική διαδικασία που επηρεάζει τα οικολογικά συστήματα σε όλο τον κόσμο. Τα θετικά αποτελέσματα της πυρκαγιάς περιλαμβάνουν την τόνωση της ανάπτυξης και τη διατήρηση διαφόρων οικολογικών συστημάτων.

Οι αρνητικές επιπτώσεις της πυρκαγιάς περιλαμβάνουν κίνδυνο για τη ζωή και την ιδιοκτησία, την ατμοσφαιρική ρύπανση και τη μόλυνση του νερού. Εάν η φωτιά καταστρέψει την προστατευτική βλάστηση, οι έντονες βροχοπτώσεις μπορεί να οδηγήσουν σε αύξηση της διάβρωσης του εδάφους από το νερό. Επίσης, όταν καίγεται η βλάστηση, το άζωτο που περιέχει απελευθερώνεται στην ατμόσφαιρα, σε αντίθεση με στοιχεία όπως το κάλιο και ο φώσφορος που παραμένουν στην τέφρα και ανακυκλώνονται γρήγορα στο έδαφος. Αυτή η απώλεια αζώτου που προκαλείται από πυρκαγιά προκαλεί μακροπρόθεσμη μείωση της γονιμότητας του εδάφους, η οποία ανακάμπτει πολύ αργά καθώς το άζωτο "σταθεροποιείται" από την ατμόσφαιρα με κεραυνούς και με οσπριοειδή φυτά όπως το τριφύλλι.

Η πυρκαγιά έχει χρησιμοποιηθεί από ανθρώπους σε τελετουργίες, στη γεωργία, για εκκαθάριση γης, για μαγείρεμα, παραγωγή θερμότητας και φωτός, για σηματοδότηση, σκοπούς πρόωσης, τήξη, σφυρηλάτηση, αποτέφρωση αποβλήτων, αποτέφρωση αλλά και ως όπλο ή τρόπο καταστροφής [9].

2.1.1 Σύσταση

Οι πυρκαγιές ξεκινούν όταν ένα εύφλεκτο υλικό, σε συνδυασμό με επαρκή ποσότητα οξειδωτικού όπως το οξυγόνο ή άλλη πλούσια σε οξυγόνο ένωση (αν και υπάρχουν οξειδωτές χωρίς οξυγόνο), εκτίθεται σε πηγή θερμότητας ή θερμοκρασίας περιβάλλοντος άνω του σημείου ανάφλεξης του μίγματος καυσίμου (οξειδωτή) και είναι σε θέση να διατηρεί ένα ρυθμό ταχείας οξείδωσης που παράγει αλυσιδωτή αντίδραση. Αυτό ονομάζεται συνήθως *τετράπτυχο της φωτιάς*, το οποίο φαίνεται και στο παρακάτω Σχήμα 2.1. Η πυρκαγιά δεν μπορεί να υπάρξει χωρίς όλα αυτά τα στοιχεία στη θέση τους και στις σωστές αναλογίες. Για παράδειγμα, ένα εύφλεκτο υγρό θα αρχίσει να καίει μόνο εάν το καύσιμο και το οξυγόνο είναι στη σωστή αναλογία. Ορισμένα μίγματα καυσίμου-οξυγόνου μπορεί να απαιτούν έναν καταλύτη, μια ουσία που δεν καταναλώνεται, όταν προστίθεται, σε οποιαδήποτε χημική αντίδραση κατά τη διάρκεια της καύσης, αλλά που επιτρέπει στα αντιδραστήρια να καούν πιο εύκολα. Αφού αναφλεγεί, πρέπει να πραγματοποιηθεί αλυσιδωτή αντίδραση, κατά την οποία οι φωτιές μπορούν να διατηρήσουν τη δική τους θερμότητα με την περαιτέρω απελευθέρωση θερμικής ενέργειας στη διαδικασία καύσης και μπορούν να διαδοθούν, υπό την προϋπόθεση ότι υπάρχει συνεχής παροχή οξειδωτικού και καυσίμου. Εάν ο οξειδωτής είναι οξυγόνο από τον περιβάλλοντα αέρα, είναι απαραίτητη η παρουσία βαρύτητας ή κάποιας παρόμοιας δύναμης που προκαλείται από την επιτάχυνση, ώστε να παράγεται η μεταφορά, η οποία απομακρύνει τα προϊόντα καύσης και φέρνει μια παροχή οξυγόνου στη φωτιά. Χωρίς βαρύτητα, μια φωτιά περιβάλλεται γρήγορα με τα δικά της προϊόντα καύσης και τα μη οξειδωτικά αέρια από τον αέρα, τα οποία αποκλείουν το οξυγόνο και σβήνουν τη φωτιά. Εξαιτίας αυτού, ο κίνδυνος πυρκαγιάς σε ένα διαστημικό σκάφος είναι μικρός όταν περνάει με αδρανειακή πτήση [10]. Αυτό δεν ισχύει αν το οξυγόνο τροφοδοτηθεί στη φωτιά με κάποια διαδικασία εκτός της θερμικής μεταφοράς.



Σχήμα 2.1: Το τετράπτυχο της φωτιάς

Η πυρκαγιά μπορεί να σβήσει απομακρύνοντας οποιοδήποτε από τα στοιχεία του τετράπτυχου φωτιάς. Εξετάστε μια φλόγα φυσικού αερίου, όπως από έναν καυστήρα με καμινάδα. Η φωτιά μπορεί να σβήσει από οποιοδήποτε από τα ακόλουθα:

- απενεργοποιώντας την παροχή αερίου, η οποία αφαιρεί την πηγή καυσίμου.
- καλύπτοντας πλήρως τη φλόγα, κάτι το οποίο την πνίγει, καθώς η καύση χρησιμοποιεί και τον διαθέσιμο οξειδωτή (το οξυγόνο στον αέρα) και το μετατοπίζει από την περιοχή γύρω από τη φλόγα με CO₂.
- εφαρμογή νερού, το οποίο απομακρύνει τη θερμότητα από τη φωτιά ταχύτερα από ό,τι η φωτιά μπορεί να την παράγει (ομοίως, το φυσικό φούσκωμα σε μια φλόγα θα μετατοπίσει τη θερμότητα του σημερινού καυσίμου από την πηγή καυσίμου του στο ίδιο άκρο), ή
- εφαρμογή επιβραδυντικού χημικού όπως Halon 1211 στη φλόγα (υλικό το οποίο χρησιμοποιείται και στους πυροσβεστήρες), η οποία καθυστερεί την ίδια τη χημική αντίδραση μέχρις ότου ο ρυθμός καύσης είναι πολύ αργός για να διατηρηθεί η αλυσιδωτή αντίδραση.

Μια φλόγα είναι ένα μίγμα αερίων αντίδρασης και στερεών που εκπέμπουν ορατό, υπέρυθρο και μερικές φορές υπεριώδες φως, το φάσμα συχνοτήτων των οποίων εξαρτάται από τη χημική σύνθεση του καυσίμου υλικού και των ενδιάμεσων προϊόντων αντίδρασης. Σε πολλές περιπτώσεις, όπως η καύση οργανικής ύλης, για παράδειγμα το ξύλο ή η ατελής καύση αερίου, τα στερεά σωματίδια πυρακτώσεως που ονομάζονται αιθάλη παράγουν τη γνωστή κόκκινη-πορτοκαλί λάμψη της "φωτιάς". Αυτό το φως έχει συνεχές φάσμα. Η πλήρης καύση του αερίου έχει ένα αχνό μπλε χρώμα λόγω της εκπομπής ακτινοβολίας μήκους κύματος από διάφορες μεταβολές ηλεκτρονίων στα διεγερμένα μόρια που σχηματίζονται στη φλόγα. Συνήθως εμπλέκεται το οξυγόνο, αλλά το υδρογόνο που καίει σε χλώριο παράγει επίσης φλόγα, παράγοντας υδροχλώριο (HCL). Άλλοι πιθανοί συνδυασμοί που παράγουν φλόγες, μεταξύ πολλών, είναι φθόριο και υδρογόνο, υδραζίνη και τετροξειδίο του αζώτου. Οι φλόγες υδρογόνου και υδραζίνης είναι ομοίως ανοιχτόχρωμες, ενώ η καύση του βορίου και των ενώσεών του, που αξιολογήθηκε στα μέσα του 20ου αιώνα ως καύσιμο υψηλής ενέργειας για κινητήρες αεριωθουμένων και πυραύλων, εκπέμπει έντονη πράσινη φλόγα, οδηγώντας στο ανεπίσημο ψευδώνυμο του "Πράσινου Δράκου".

Η λάμψη μιας φλόγας είναι σύνθετη. Η ακτινοβολία μαύρου σώματος εκπέμπεται από αιθάλη, αέριο και σωματίδια καυσίμου, αν και τα σωματίδια της αιθάλης είναι πολύ μικρά για να συμπεριφέρονται σαν τέλειες μαύρες κηλίδες. Υπάρχει επίσης εκπομπή φωτονίων από τα διεγερμένα άτομα και μόρια στα αέρια. Μεγάλο μέρος της ακτινοβολίας εκπέμπεται στις ορατές και υπέρυθρες ζώνες. Το χρώμα εξαρτάται από τη θερμοκρασία για την ακτινοβολία μαύρου σώματος και από τη χημική σύνθεση των φασμάτων εκπομπής. Το κυρίαρχο χρώμα σε μια φλόγα αλλάζει με τη θερμοκρασία. Η φωτογραφία της δασικής πυρκαγιάς στο ακόλουθο Σχήμα 2.2 είναι ένα εξαιρετικό παράδειγμα αυτής της παραλλαγής. Κοντά στο έδαφος, όπου συμβαίνει το μεγαλύτερο μέρος της καύσης, η φωτιά είναι λευκή, ενώ σιγά-σιγά μετατρέπεται προς το κίτρινο. Πάνω από την κίτρινη περιοχή, το χρώμα αλλάζει σε πορτοκαλί, το οποίο είναι πιο "δροσερό", και στη συνέχεια κόκκινο. Πάνω από την κόκκινη περιοχή, δεν συμβαίνει πλέον καύση και τα αμόλυντα σωματίδια άνθρακα είναι ορατά ως μαύρος καπνός.

Η κοινή κατανομή μιας φλόγας υπό φυσιολογικές συνθήκες βαρύτητας εξαρτάται από τη μεταφορά, καθώς η αιθάλη τείνει να ανέλθει στην κορυφή μιας φλόγας, όπως σε ένα κερί σε κανονικές συνθήκες βαρύτητας, καθιστώντας την κίτρινη. Με τη μικροβαρύτητα ή τη μηδενική βαρύτητα [11], όπως ένα περιβάλλον στο διάστημα, δεν



Σχήμα 2.2: Επίπεδα Χρώματος της Φωτιάς

συμβαίνει πλέον η μεταφορά και η φλόγα γίνεται σφαιρική, με την τάση να γίνεται μπλε και πιο αποτελεσματική (αν και μπορεί να βγαίνει αν δεν μετακινηθεί σταθερά, καθώς το CO₂ από την καύση δεν διασκορπίζεται τόσο εύκολα στη μικροβαρύτητα και τείνει να πνίξει τη φλόγα). Υπάρχουν αρκετές πιθανές εξηγήσεις για αυτή τη διαφορά, από τις οποίες πιθανότατα είναι ότι η θερμοκρασία είναι επαρκώς κατανεμημένη ομοιόμορφα ώστε η αιθάλη δεν σχηματίζεται και εμφανίζεται πλήρης καύση. Τα πειράματα από τη NASA αποκαλύπτουν ότι οι φλόγες διάχυσης με μικροβαρύτητα επιτρέπουν την περισσότερο οξειδωμένη αιθάλη μετά την παραγωγή τους από τις φλόγες διάχυσης στη Γη, λόγω μιας σειράς μηχανισμών που συμπεριφέρονται διαφορετικά στη μικροβαρύτητα σε σύγκριση με τις φυσιολογικές συνθήκες βαρύτητας. Αυτές οι ανακαλύψεις έχουν πιθανές εφαρμογές στην εφαρμοσμένη επιστήμη και τη βιομηχανία, ειδικά όσον αφορά την αποδοτικότητα των καυσίμων.

Σε κινητήρες καύσης, λαμβάνονται διάφορα μέτρα για την εξάλειψη μιας φλόγας. Η μέθοδος εξαρτάται κυρίως από το εάν το καύσιμο είναι πετρέλαιο, ξύλο ή καύσιμο υψηλής ενέργειας όπως αυτό των αεροσκαφών.

2.1.2 Αστικό Περιβάλλον

Οι πυρκαγιές σε ζώνες μίξης δασών-οικισμών (Wildland-Urban Interface fires (WUI)) είναι ένα σχετικά νέο πρόβλημα. Άρχισε να εμφανίζεται από τη δεκαετία του 1960 σε μέρη όπως η Καλιφόρνια, αλλά από τις αρχές της δεκαετίας του 1980 αναδείχθηκε σταδιακά σαν ένα ξεχωριστό πρόβλημα μεγάλης σημασίας σε όλες σχεδόν τις πυρόπληχτες περιοχές του ανεπτυγμένου κόσμου.

Αίτια

Κύρια αίτια του προβλήματος αυτού αποτελούν η δόμηση νέων κατοικιών σε επαφή ή και μίξη με δάση και δασικές εκτάσεις, σε μία προσπάθεια διαφυγής από τα αστικά κέντρα, χωρίς να λαμβάνεται υπόψη ο κίνδυνος πυρκαγιάς, αλλά και οι κάτοικοι που δεν συνειδητοποιούν τις παραμέτρους κινδύνου του νέου τρόπου διαβίωσής τους δίπλα στο δάσος. Επιπλέον, η αυξημένη ποσότητα καύσιμης ύλης και επομένως μεγαλύτερο δυναμικό καταστροφής λόγω εγκατάλειψης της υπαίθρου και η μη λήψη προληπτικών μέτρων ενισχύουν τη δημιουργία φαινομένων πυρκαγιάς.

WUI στην Ελλάδα

Οι κυριότεροι τύποι WUI στην Ελλάδα είναι οι ακόλουθοι:

- Παραδοσιακά χωριά με μειωμένο σήμερα πληθυσμό
- Οικισμοί πρώτης κατοικίας
- Οικισμοί εξοχικής κατοικίας
- Υποδομές τουριστικής ανάπτυξης (ξενοδοχεία, κατασκηνώσεις, πάρκα αναψυχής, αγροτουριστικές επιχειρήσεις, αρχαιολογικοί χώροι, κλπ.)
- Υποδομές βιομηχανικής ανάπτυξης

Ο κάθε ένας από τους παραπάνω τύπους WUI διακρίνεται από ιδιαίτερα χαρακτηριστικά ανάπτυξης, χωροταξίας (δρόμοι, ελεύθεροι χώροι), δόμησης κατοικιών (υλικά δόμησης, ύψος), μεγέθους οικοπέδων, αλλά και από διαφορετικό προφίλ του πληθυσμού. Επομένως υπάρχουν διαφορές και ως προς τα προβλήματα και τους κινδύνους σε περίπτωση πυρκαγιάς. Παρακάτω, ακολουθεί πιο λεπτομερής περιγραφή των παραπάνω τύπων ζωνών μίξης δασών-οικισμών.

1. **Παραδοσιακά χωριά:** Αποτελούν απλές παραδοσιακές κατοικίες από αδρανή υλικά αλλά με αρκετά ευαίσθητα σημεία (π.χ. αποθήκες, πολυκαιρισμένα κουφώματα). Το οδικό δίκτυο συχνά ανεπαρκές εντός του χωριού (πλην του κεντρικού δρόμου) και η βλάστηση συχνά φθάνει μέχρι τις κατοικίες, αλλά υπάρχουν και αρκετές καλλιέργειες. Ο πληθυσμός μειωμένος και ηλικιωμένος, με αρκετή εμπειρία από πυρκαγιές αλλά και ποικίλη διάθεση αντίδρασης σε περίπτωση πυρκαγιάς.
2. **Οικισμοί πρώτης κατοικίας:** Είναι κατοικίες με μέση και άνω ποιότητα κατασκευής από δύσφλεκτα υλικά και με λίγα αδύνατα σημεία. Σχετικά μεγάλα οικόπεδα, συχνά με καλή διαμόρφωση εξωτερικού χώρου και προστατευμένα με τοίχους, καθώς και επαρκές οδικό δίκτυο. Οι κάτοικοι έχουν αρκετό ενδιαφέρον αλλά χωρίς ιδιαίτερες γνώσεις σχετικά με τις πυρκαγιές.
3. **Οικισμοί εξοχικής κατοικίας:** Συχνά παράνομα δομημένοι με χαμηλή έως μέση ποιότητα κατασκευής κατοικιών, με αρκετά εύφλεκτα υλικά και πολλά αδύνατα σημεία. Σχετικά μικρά οικόπεδα, ενώ βλάστηση είναι συχνά σε επαφή με

τις κατοικίες. Το οδικό δίκτυο είναι φτωχό (πλάτος δρόμων, κλίσεις, αδιέξοδοι δρόμοι, αδυναμία στροφής), ενώ οι κάτοικοι χαρακτηρίζονται από μικρή συνειδητοποίηση του κινδύνου πυρκαγιάς και έλλειψη γνώσεων πρόληψης και αντίδρασης.

4. **Περιοχές τουριστικής ανάπτυξης:** Ποικιλία ως προς τα χαρακτηριστικά των κτισμάτων και την κατάσταση της βλάστησης, αλλά και ιδιαίτερα προβλήματα που παρουσιάζονται όσον αφορά τους τουρίστες και επισκέπτες: άγνοια χώρου και οδικού δικτύου, συχνά άγνοια ως προς τις πυρκαγιές και επομένως απρόσμενες αντιδράσεις.
5. **Βιομηχανικές περιοχές:** Συνήθως υπάρχουν επαρκείς χώροι απομόνωσης των κτιρίων από τη βλάστηση, αλλά δεν παύουν να υπάρχουν αδύνατα σημεία. Σε περίπτωση άφιξης πυρκαγιάς, αποτελεί ερώτημα η ύπαρξη γνώσεων αντίδρασης του προσωπικού. Εάν υπάρξει ανάφλεξη, υπάρχουν συχνά ιδιαίτερα εύφλεκτα και επικίνδυνα υλικά που μπορούν να δημιουργήσουν μεγάλο κίνδυνο.

Οι πυρκαγιές σε περιοχές μίξης δασών οικισμών είναι παγκόσμιο φαινόμενο που απαντάται κυρίως στις αναπτυγμένες χώρες.

2.1.3 Δασικό Περιβάλλον

Πρόληψη

Η έγκαιρη πρόληψη της δασικής πυρκαγιάς αναφέρεται στις προληπτικές μεθόδους που στοχεύουν στη μείωση του κινδύνου πυρκαγιών καθώς και στη μείωση της σοβαρότητας και της εξάπλωσής τους. Οι τεχνικές πρόληψης στοχεύουν στη διαχείριση της ποιότητας του αέρα, στη διατήρηση των οικολογικών ισορροπιών, την προστασία των πόρων και την αντιμετώπιση μελλοντικών πυρκαγιών. Οι πολιτικές πυρόσβεσης της Βόρειας Αμερικής επιτρέπουν το φυσικό κάψιμο των πυρκαγιών για να διατηρηθούν τον οικολογικό τους ρόλο, αρκεί οι κίνδυνοι διαφυγής σε περιοχές υψηλής αξίας να μετριάζονται. Ωστόσο, οι πολιτικές πρόληψης πρέπει να εξετάσουν το ρόλο που διαδραματίζουν οι άνθρωποι στις πυρκαγιές, καθώς, για παράδειγμα, το 95% των δασικών πυρκαγιών στην Ευρώπη σχετίζονται με την ανθρώπινη συμμετοχή [12]. Οι εστίες πυρκαγιάς που προκαλούνται από ανθρώπους μπορεί να περιλαμβάνουν εμπρησμό, τυχαία ανάφλεξη ή ανεξέλεγκτη χρήση φωτιάς στην εκκαθάριση της γης και στη γεωργία.

Το 1937, ο Πρόεδρος των ΗΠΑ, Φράνκλιν Ρούσβελτ, ξεκίνησε μια εθνική εκστρατεία πρόληψης πυρκαγιών, επισημαίνοντας τον ρόλο της ανθρώπινης απροσεξίας στις δασικές πυρκαγιές. Χαρακτηριστική είναι η φωτογραφία της επίσημης μασκότ της αμερικανικής δασικής υπηρεσίας Smokey Bear που φαίνεται στο Σχήμα 2.3.

Η μείωση των αναφλέξεων που προκαλούνται από τον άνθρωπο μπορεί να είναι το πιο αποτελεσματικό μέσο για τη μείωση της ανεπιθύμητης πυρκαγιάς. Η μείωση των καυσίμων γίνεται συνήθως όταν επιχειρείται να επηρεάσει τον μελλοντικό κίνδυνο πυρκαγιάς. Τα προγράμματα πρόληψης των πυρκαγιών σε όλο τον κόσμο μπορούν να χρησιμοποιούν τεχνικές όπως η *άγρια φωτιά* και οι *προκαθορισμένες ή ελεγχόμενες εστίες*. Η *άγρια φωτιά* σαν όρος αναφέρεται σε οποιαδήποτε πυρκαγιά φυσικών αιτιών που παρακολουθείται αλλά αφήνεται να καεί. Οι ελεγχόμενες εστίες πυροδοτούνται από κυβερνητικές υπηρεσίες υπό λιγότερο επικίνδυνες καιρικές συνθήκες.



Σχήμα 2.3: Αφίσα της Αμερικάνικης εκστρατείας για την πρόληψη ενάντια στις πυρκαγιές

Η βλάστηση μπορεί να καεί περιοδικά για να διατηρηθεί υψηλή ποικιλία ειδών και η συχνή καύση επιφανειακών καυσίμων περιορίζει τη συσσώρευσή τους. Η χρήση της άγριας φωτιάς από τον άνθρωπο είναι η φθηνότερη και πιο οικολογικά κατάλληλη πολιτική για πολλά δάση. Η ίδια η πυρκαγιά αναφέρεται ως "η πιο αποτελεσματική θεραπεία για τη μείωση της συχνότητας εξάπλωσης πυρκαγιάς, της έντασης της φωτιάς, του μήκους της φλόγας και της θερμότητας ανά μονάδα επιφάνειας", σύμφωνα με τον Jan van Wagtendonk, βιολόγο στο Yellowstone Field Station.

Οι κτιριακοί κώδικες σε περιοχές επιρρεπείς σε πυρκαγιά απαιτούν να κατασκευάζονται εγκαταστάσεις από υλικά ανθεκτικά στη φλόγα και να διατηρείται ένας εύχρηστος χώρος με καθαρισμό εύφλεκτων υλικών εντός μιας προκαθορισμένης απόστασης από τη δομή. Οι κοινότητες στις Φιλιππίνες διατηρούν επίσης γραμμές πυρκαγιάς πλάτους 5 έως 10 μέτρων μεταξύ του δάσους και του χωριού τους και περιπολούν αυτές τις γραμμές κατά τους καλοκαιρινούς μήνες ή τις εποχές του ξηρού καιρού [13]. Η συνεχιζόμενη οικιστική ανάπτυξη σε περιοχές επιρρεπείς σε πυρκαγιά και η αποκατάσταση κατασκευών που καταστράφηκαν από τις πυρκαγιές αντιμετωπίστηκαν με κριτική. Τα οικολογικά οφέλη της φωτιάς συχνά παραβλέπονται από τα οφέλη της οικονομίας και της ασφάλειας από την προστασία των δομών και της ανθρώπινης ζωής.

Αντιμετώπιση

Η γρήγορη και αποτελεσματική ανίχνευση αποτελεί βασικό παράγοντα για την καταπολέμηση των πυρκαγιών. Οι προσπάθειες πρώιμης ανίχνευσης επικεντρώθηκαν στην έγκαιρη ανταπόκριση, τα ακριβή αποτελέσματα τόσο κατά τη διάρκεια της ημέρας όσο

και κατά τη διάρκεια της νύχτας, καθώς και την ικανότητα να δοθεί προτεραιότητα στον κίνδυνο πυρκαγιάς. Οι πυροσβεστικοί πύργοι χρησιμοποιήθηκαν στις Ηνωμένες Πολιτείες στις αρχές του 20ου αιώνα και εντοπίζονταν πυρκαγιές χρησιμοποιώντας τηλέφωνα, περιστέρια μεταφοράς και ηλιογράφοι. Η αεροναυτική και η χερσαία φωτογραφία με τη χρήση στιγμιαίων φωτογραφικών μηχανών χρησιμοποιήθηκαν στη δεκαετία του 1950 μέχρις ότου αναπτύχθηκε η υπέρυθρη σάρωση για ανίχνευση πυρκαγιών στη δεκαετία του 1960. Ωστόσο, η ανάλυση και η παράδοση των πληροφοριών συχνά καθυστέρησαν από τους περιορισμούς στην τεχνολογία επικοινωνιών. Οι πρώτες δορυφορικές αναλύσεις πυρκαγιάς έγιναν με το χέρι σε χάρτες σε απομακρυσμένο χώρο και αποστέλλονταν μέσω αλληλογραφίας στο χειριστή πυρκαγιάς. Κατά τη διάρκεια των πυρκαγιών του Yellowstone του 1988, εγκαταστάθηκε σταθμός δεδομένων στο West Yellowstone, επιτρέποντας την παράδοση πληροφοριών σχετικά με τη φωτιά μέσω δορυφόρου σε περίπου τέσσερις ώρες.

Επί του παρόντος, οι δημόσιες ανοικτές τηλεφωνικές γραμμές, οι επιθεωρήσεις πυρκαγιών σε πύργους και οι επιτόπιες και εναέριας περιπολίες μπορούν να χρησιμοποιηθούν ως μέσο έγκαιρης ανίχνευσης δασικών πυρκαγιών. Ωστόσο, η ακριβής ανθρώπινη παρατήρηση μπορεί να περιορίζεται από την κόπωση του χειριστή, την ώρα της ημέρας, την εποχή του χρόνου και τη γεωγραφική θέση. Τα ηλεκτρονικά συστήματα έχουν κερδίσει τη δημοτικότητα τα τελευταία χρόνια ως μια πιθανή λύση στο σφάλμα του χειριστή του ανθρώπου. Μια κυβερνητική έκθεση σχετικά με μια πρόσφατη δοκιμή τριών αυτοματοποιημένων συστημάτων ανίχνευσης πυρκαγιάς στην Αυστραλία κατέληξε ωστόσο στο συμπέρασμα ότι η ανίχνευση από τα συστήματα κάμερας ήταν πιο αργή και λιγότερο αξιόπιστη από ότι ένας εκπαιδευμένος παρατηρητής ανθρώπων. Αυτά τα συστήματα μπορεί να είναι εν μέρη ή πλήρως αυτοματοποιημένα και χρησιμοποιούν συστήματα βασισμένα στην περιοχή κινδύνου και τον βαθμό ανθρώπινης παρουσίας, όπως προτείνεται από τις αναλύσεις δεδομένων του Geographic Information System (GIS). Μια ολοκληρωμένη προσέγγιση πολλαπλών συστημάτων μπορεί να χρησιμοποιηθεί για τη συγχώνευση δορυφορικών δεδομένων, αεροφωτογραφιών και θέσης προσωπικού μέσω του Global Positioning System (GPS) σε ένα συλλογικό σύνολο για χρήση σχεδόν σε πραγματικό χρόνο από τα ασύρματα κέντρα εντοπισμού συμβάντων.

Μια μικρή, υψηλού κινδύνου περιοχή που χαρακτηρίζεται από πυκνή βλάστηση, ισχυρή ανθρώπινη παρουσία ή είναι κοντά σε μια κρίσιμη αστική περιοχή μπορεί να παρακολουθείται χρησιμοποιώντας ένα τοπικό δίκτυο αισθητήρων. Τα συστήματα ανίχνευσης μπορούν να περιλαμβάνουν ασύρματα δίκτυα αισθητήρων που λειτουργούν ως αυτοματοποιημένα συστήματα καιρού: ανίχνευση θερμοκρασίας, υγρασίας και καπνού [14]. Αυτά μπορεί να είναι με τροφοδοσία από μπαταρία, με ηλιακή ενέργεια ή με επαναφορτιζόμενες δένδεις: μπορούν να επαναφορτίζουν τα συστήματα μπαταρίας τους χρησιμοποιώντας τα μικρά ηλεκτρικά ρεύματα στο φυτικό υλικό [15]. Μεγαλύτερες, μεσαίου κινδύνου περιοχές μπορούν να παρακολουθούνται από πύργους σάρωσης που ενσωματώνουν σταθερές κάμερες και αισθητήρες για την ανίχνευση καπνού ή πρόσθετους παράγοντες όπως η υπέρυθρη υπογραφή του διοξειδίου του άνθρακα που παράγεται από τις πυρκαγιές. Επιπρόσθετες δυνατότητες όπως νυχτερινή όραση, ανίχνευση φωτεινότητας και ανίχνευση αλλαγής χρώματος μπορούν επίσης να ενσωματωθούν στις συστοιχίες αισθητήρων.

Η παρακολούθηση μέσω δορυφόρου και κεραίας, μέσω της χρήσης αεροσκαφών, ελικοπτέρων ή από Unmanned Aerial Vehicle (UAV) μπορεί να προσφέρει μια ευρύτερη εικόνα και μπορεί να επαρκεί για την παρακολούθηση πολύ μεγάλων, χαμηλού κινδύνου

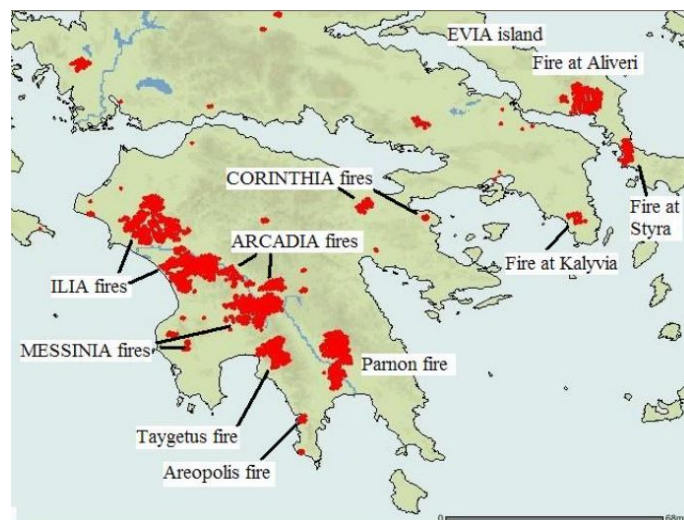
περιοχών. Αυτά τα πιο εξελιγμένα συστήματα χρησιμοποιούν GPS και τα αεροσκάφη έχουν τοποθετημένες υπέρυθρες ή υψηλής ανάλυσης κάμερες για τον εντοπισμό και την εστίαση των πυρκαγιών. Οι δορυφορικοί αισθητήρες, όπως το προηγμένο ραδιόμετρο σάρωσης Envisat (AATSR) και ο ευρωπαϊκός ραδιομετρητής σάρωσης με τηλεανίχνευση (ERS), μπορούν να μετρήσουν την υπέρυθρη ακτινοβολία που εκπέμπεται από τις πυρκαγιές, προσδιορίζοντας θερμά σημεία μεγαλύτερα των 39°C. Το Σύστημα Χαρτογράφησης Κινδύνων της Εθνικής Ωκεανικής και Ατμοσφαιρικής Υπηρεσίας των ΗΠΑ (NOAA) συνδυάζει δεδομένα τηλεπισκόπησης από δορυφορικές πηγές, όπως το Γεωστατικό Δορυφόρο (GOES), το Φασματοραδιόμετρο Μέσης Ανάλυσης (MODIS) και το Προηγμένο Ραδιόμετρο Υψηλής Ανάλυσης (AVHRR). Ωστόσο, η ανίχνευση μέσω δορυφόρου είναι επιρρεπής σε σφάλματα αντιστάθμισης, οπουδήποτε από 2 έως 3 χιλιόμετρα για τα δεδομένα MODIS και AVHRR και μέχρι 12 χιλιόμετρα για τα δεδομένα GOES. Οι δορυφόροι σε γεωστατικές τροχιές ενδέχεται να απενεργοποιηθούν και οι δορυφόροι σε πολιτικές τροχιές συχνά περιορίζονται από το σύντομο παράθυρο παρατήρησής τους. Το κάλυμμα του νέφους και η ανάλυση εικόνας μπορεί επίσης να περιορίσουν την αποτελεσματικότητα των δορυφορικών εικόνων.

Από το 2015 λειτουργεί ένα νέο εργαλείο ανίχνευσης πυρκαγιάς στη Δασική Υπηρεσία του Υπουργείου Γεωργίας των ΗΠΑ (USDSA), το οποίο χρησιμοποιεί δεδομένα από τον δορυφόρο Suomi NPP για την ανίχνευση μικρότερων πυρκαγιών με μεγαλύτερη ακρίβεια σε σχέση με τα προηγούμενα προϊόντα διαστημικής φύσεως. Τα δεδομένα υψηλής ανάλυσης χρησιμοποιούνται με ένα μοντέλο υπολογιστή για να γίνει πρόβλεψη για το πώς η φωτιά θα αλλάξει κατεύθυνση με βάση τις καιρικές συνθήκες και τις συνθήκες του εδάφους. Το ενεργό προϊόν ανίχνευσης πυρκαγιάς που χρησιμοποιεί δεδομένα από το ακτινομετρικό ραδιόμετρο (VIIRS) της Suomi αυξάνει την ανάλυση ανίχνευσης πυρκαγιάς στα 375 μέτρα. Τα προηγούμενα προϊόντα δορυφορικών δεδομένων της NASA που ήταν διαθέσιμα από τις αρχές της δεκαετίας του 2000 παρατηρούσαν πυρκαγιές σε ανάλυση 3.280 ποδιών (1 χιλιόμετρο). Τα δεδομένα είναι ένα από τα εργαλεία πληροφοριών που χρησιμοποιούν οι περισσότερες υπηρεσίες σε όλες τις Ηνωμένες Πολιτείες για να κατευθύνουν την κατανομή των πόρων και τις στρατηγικές αποφάσεις διαχείρισης πυρκαγιάς. Το προϊόν πυρκαγιάς VIIRS επιτρέπει την ανίχνευση κάθε 12 ώρες ή λιγότερες για πολύ μικρότερες πυρκαγιές και παρέχει περισσότερες λεπτομέρειες και συνεπή παρακολούθηση των γραμμών πυρκαγιάς κατά τη διάρκεια πυρκαγιών μεγάλης διάρκειας - ικανότητες κρίσιμες για συστήματα έγκαιρης προειδοποίησης και υποστήριξη της χαρτογράφησης ρουτίνας της εξέλιξης της πυρκαγιάς. Οι ενεργές τοποθεσίες πυρκαγιάς διατίθενται στους χρήστες μέσα σε λίγα λεπτά από τη διέλευση από το δορυφόρο μέσω των εγκαταστάσεων επεξεργασίας δεδομένων στο Κέντρο Εφαρμογών Τηλεπισκόπησης (USFS), το οποίο χρησιμοποιεί τεχνολογίες που αναπτύχθηκαν από το εργαστήριο Direct Readout Laboratory του NASA Goddard Space Flight Center στο Greenbelt του Maryland. Το μοντέλο χρησιμοποιεί δεδομένα σχετικά με τις καιρικές συνθήκες και τη γη που περιβάλλει μια ενεργό φωτιά για να προβλέψει 12-18 ώρες εκ των προτέρων αν μια φωτιά θα αλλάξει κατεύθυνση. Η πολιτεία του Κολοράντο αποφάσισε να ενσωματώσει το μοντέλο της πυρκαγιάς στις πυροσβεστικές της προσπάθειες ξεκινώντας από την πυρκαγιά του 2016 [16].

Από τέτοια είδη δορυφόρων προέρχεται το παρακάτω Σχήμα 2.4, καθώς και το Σχήμα 2.5, τα οποία απεικονίζουν τις καταστροφικές πυρκαγιές που έλαβαν χώρα στην Ελλάδα το 2007.



Σχήμα 2.4: Οι πυρκαγιές στη Νότια Ελλάδα την 15η Αυγούστου 2007 (από δορυφόρο της NASA)



Σχήμα 2.5: Οι καμμένες περιοχές (25 Αυγούστου - 5 Σεπτέμβρη 2007)

2.1.4 Φωτιά και Άνθρωπος: Από τον Προμηθέα στο Σήμερα

Ο μύθος του Προμηθέα και της φωτιάς μάς κάνει να διερωτώμαστε το εξής: Αν ο Προμηθέας δεν είχε κλέψει τη φωτιά από το Δία, τι θα έκανε η ανθρωπότητα; Αλλά ο άτακτος τιτάνας της ελληνικής μυθολογίας την έκλεψε και ενώ γιορτάστηκε από τους θνητούς, τιμωρήθηκε σκληρά από τον Θεό όλων των Θεών.

Η σκέψη για την κλοπή της φωτιάς ήταν εύκολη, αλλά τελικά αποδείχθηκε λίγο πιο περίπλοκη. Ο Προμηθέας, γνωστός για το πνεύμα και τη νοημοσύνη του, είχε ένα άμεσο σχέδιο - να εξαπατήσει τις θεές ρίχνοντάς τους ένα χρυσό αχλάδι (σε κάποια εκδοχή μήλο) στην αυλή με ένα μήνυμα: "Για την ωραιότερη θεά όλων".

Λειτουργούσαν όλα ακριβώς όπως τα είχε σχεδιάσει - οι θεές άρχισαν μια μάχη για το φρούτο, ενώ οι θεοί απολάμβαναν εντελώς τη σκηνή. Όλοι τους αποστασιοποιήθηκαν και ο Προμηθέας δεν είχε κάποια δυσκολία στο να κλέψει τη φωτιά από το εργαστήριο του Ηφαίστου. Ο Ήφαιστος ήταν, μεταξύ άλλων, ο έλληνας Θεός της

φωτιάς. Ο Προμηθέας εγκατέλειψε το γλέντι των Θεών και πήρε μαζί του τη φωτιά είτε σε κοίλο κολοκύθας είτε σε κοίλο από καλάμι (ανάλογα με την ερμηνεία) και την έφερε στη Γη και την έδωσε στον άνθρωπο.

Τότε ο Δίας έγινε τρελός. Μετά από τόσες πολλές φορές που τον αφήφησε ο Προμηθέας, αποφάσισε ότι έπρεπε να τον τιμωρήσει. Έτσι και έγινε, έβαλε τον ίδιο τον Ήφαιστο να αλυσοδέσει τον Προμηθέα πάνω στον Καύκασο, όπου ο αετός θα έτρωγε το συκώτι του για πάντα [17]. Αξίζει επίσης να σημειωθεί ότι ο Προμηθέας αποτελεί και σύμβολο του Εθνικού Μετσόβιου Πολυτεχνείου.

Όσον αφορά στη σημερινή εποχή και τον 21ο αιώνα, όπως αναφέρθηκε και στο Κεφάλαιο 1, οι δασικές πυρκαγιές αποτελούν ένα μείζον θέμα της σύγχρονης κοινωνίας καθώς αρκετοί θάνατοι οφείλονται σε αυτές σχεδόν κάθε χρόνο. Στον Πίνακα 2.1 παρουσιάζονται οι 15 πυρκαγιές με τα περισσότερα θύματα στη νεότερη ιστορία του πλανήτη Γη. Όπως γίνεται αντιληπτό, η Ελλάδα έχει 2 συμμετοχές σε αυτό το στατιστικό με πιο πρόσφατο (και πιο υψηλά στην κατάταξη) γεγονός αυτό του Ιουλίου του 2018, αλλά και του καλοκαιριού του 2007 [18].

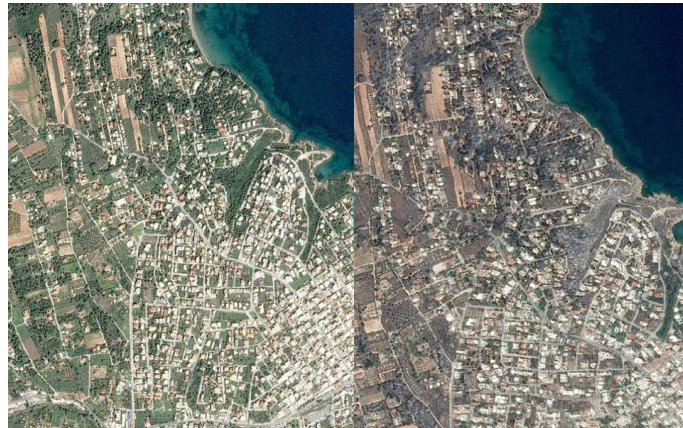
Rank	Death Toll	Event	Date
1	1,200–2,500	Wisconsin, United States	October 8, 1871
2	1,200	Meshchyora, Soviet Union	August 3, 1936
3	453	Minnesota, United States	October 12, 1918
4	418	Minnesota, United States	September 1, 1894
5	282	Michigan, United States	September 5, 1881
6	240	Sumatra, Indonesia	September 1997
7	223	Ontario, Canada	July 29, 1916
8	191	China and Soviet Union	May 1, 1987
9	180	Australia	February 7, 2009
10	160-300	Canada	October 7, 1825
11	102	Attica, Greece	July 23, 2018
12	87	Montana and Idaho, United States	August 20, 1910
13	85	California, United States	November 8, 2018
14	84	Greece	June 28, 2007
15	82	France	August 19, 1949

Πίνακας 2.1: Οι 15 πιο θανατηφόρες πυρκαγιές της σύγχρονης ιστορίας

Μάλιστα, εάν επιλεχθούν γεγονότα που έλαβαν χώρα τον 21ο αιώνα, η Ελλάδα κατέχει τη μερίδα του λέοντος με 2 συμμετοχές από τις 4 συνολικά. Τέλος, παρουσιάζουμε την ακόλουθη φωτογραφία στο Σχήμα 2.6 που απεικονίζει την περιοχή στο Μάτι Αττικής πριν και μετά την καταστροφή.

2.2 Ανάγκη για Εξέλιξη

Όπως παρουσιάστηκε στο Κεφάλαιο αυτό, κατά καιρούς ο άνθρωπος έχει κατασκευάσει πολλά και διάφορα συστήματα για τον έγκαιρο εντοπισμό της πυρκαγιάς, από ταχυδρομικά περιστέρια μέχρι τους δορυφόρους της NASA. Ακόμα και τα πιο σύγχρονα συστήματα όμως, έχουν τα προβλήματά τους, όπως αναφέρθηκε στην Ενότητα 2.1.3.



Σχήμα 2.6: Μάτι Ατικής πριν και μετά την καταστροφική πυρκαγιά του 2018

Με την ραγδαία ανάπτυξη της τεχνολογίας και ειδικότερα της τεχνητής νοημοσύνης, έχει γίνει πλέον ευρέως γνωστός ο όρος "Μηχανική Μάθηση". Είναι εύλογο το ερώτημα κατά πόσον μπορεί αυτός ο διαρκώς αναπτυσσόμενος κλάδος της επιστήμης των υπολογιστών να συμβάλλει στην καταπολέμηση αυτών των φαινομένων. Στα επόμενα Κεφάλαια, αναλύονται ακριβώς οι τρόποι με τους οποίους αυτό το ενδεχόμενο μπορεί να γίνει πραγματικότητα.

Κεφάλαιο 3

Μηχανική Μάθηση

Σε αυτό το κεφάλαιο, παρουσιάζεται η θεωρία του υπόβαθρου της Μηχανικής Μάθησης που χρησιμοποιήθηκε στην παρούσα εργασία, καθώς και των Συνελικτικών Νευρωνικών Δικτύων, αρχιτεκτονικές και εφαρμογές τους με την πρόθεση να αναλυθεί το σύστημα αναγνώρισης πυρκαγιάς σε εικόνες, βίντεο αλλά και σε πραγματικό χρόνο.

3.1 Ορισμός

Μηχανικά μάθηση ονομάζουμε το κομμάτι της επιστήμης υπολογιστών που προσπαθεί να μοντελοποιήσει πολλαπλά επίπεδα αφαιρετικότητας, χρησιμοποιώντας πολλαπλά επίπεδα στις αρχιτεκτονικές της. Τα δίκτυα αυτά είναι εμπνευσμένα από το πώς επεξεργάζεται ο άνθρωπος την πληροφορία και προσπαθούν να προσομοιάσουν την λειτουργία των νευρώνων στο νεοφλοιό του εγκεφάλου (όπου γίνεται περίπου το 80% της ανθρώπινης σκέψης). Η διαδικασία της μάθησης ξεκινά με παρατηρήσεις, που αποτελούν παραδείγματα, ή εμπειρικά αποτελέσματα ή οδηγίες, ούτως ώστε να αναγνωριστούν πρότυπα στα δεδομένα και να ληφθούν καλύτερες αποφάσεις στο μέλλον, με βάση τα παραδείγματα που διαθέτουμε. Ο πρωταρχικός σκοπός είναι να επιτρέψουμε στους υπολογιστές να μαθαίνουν αυτόματα, χωρίς ανθρώπινη παρέμβαση ή βοήθεια, και να προσαρμόζουν τις πράξεις τους κατάλληλα.

Στη μηχανική μάθηση, τα καθήκοντα ταξινομούνται γενικά σε ευρείες κατηγορίες. Οι κατηγορίες αυτές βασίζονται στον τρόπο με τον οποίο λαμβάνεται η μάθηση ή στον τρόπο με τον οποίο δίνεται ανάδραση στην εκμάθηση στο ανεπτυγμένο σύστημα. Δύο από τις πιο ευρέως υιοθετημένες μεθόδους είναι η επιβλεπόμενη μάθηση, η οποία εκπαιδεύει αλγόριθμους που βασίζονται στα δεδομένα εισόδου και εξόδου τα οποία επισημαίνονται (αποκτούν ετικέτες-labels) από τον άνθρωπο και η μη επιβλεπόμενη μάθηση, η οποία παρέχει τον αλγόριθμο χωρίς επισημασμένα δεδομένα, ούτως ώστε να του επιτρέψει να βρει δομή στα δεδομένα εισόδου του.

3.2 Μάθηση με Επίβλεψη

Η Μάθηση με Επίβλεψη είναι ο τομέας μηχανικής μάθησης όπου μια συνάρτηση μαθαίνει να αντιστοιχίζει δεδομένα εισόδου σε δεδομένα εξόδου χρησιμοποιώντας παραδείγματα ζευγών εισόδου-εξόδου. Αυτό το σύνολο ζευγών ονομάζεται σύνολο εκπαίδευσης και η διαδικασία υπολογισμού μιας τέτοιας συνάρτησης από το παραπάνω σύνολο λέγεται εκπαίδευση. Ο σκοπός της μάθησης με επίβλεψη είναι ο υπολογισμός

μιας συνάρτησης που γενικεύει επαρκώς σε δεδομένα εισόδου στα οποία δεν έχει εκπαιδευτεί, αντιστοιχίζοντάς τα σε σωστές εξόδους.

Η μάθηση με επίβλεψη είναι κατάλληλη για εργασίες όπου, ενώ η απεικόνιση εισόδου εξόδου είναι δύσκολο-πολύπλοκο να βρεθεί αναλυτικά, υπάρχει ένα αρκετά μεγάλο σύνολο δεδομένων εκπαίδευσης. Μια τέτοια περίπτωση είναι η ταξινόμηση εικόνων. Από προγραμματιστική άποψη, οι εικόνες αναπαριστώνται από 3D τένσορες με διαστάσεις ύψους, πλάτους και βάθους (καναλιών, πχ RGB). Ένα σύνολο εκπαίδευσης για ταξινόμηση εικόνων περιέχει μια συλλογή εικόνων και αντίστοιχες επισημάνσεις. Ένα παράδειγμα είναι εικόνες με αυτοκίνητα και ποδήλατα με τις αντίστοιχες επισημάνσεις "αυτοκίνητο" ή "ποδήλατο" για κάθε μια από αυτές.

Στη μάθηση με επίβλεψη, οι είσοδοι του συνόλου εκπαίδευσης πρέπει να μετατραπούν σε καταλληλότερες δομές ώστε να γίνει περαιτέρω επεξεργασία. Τέτοια είναι η περίπτωση για σύνολα εκπαίδευσης σχετικά με κείμενο και η κλασική προσέγγιση στην όραση υπολογιστών. Πριν τα συνελικτικά νευρωνικά δίκτυα γίνουν δημοφιλή, η διαδικασία ταξινόμησης εικόνων περιλάμβανε την μετατροπή κάθε εικόνας σε ένα σύνολο διανυσμάτων περιγραφής της εικόνας, τα χαρακτηριστικά, τα οποία χρησιμοποιούνταν σαν είσοδος για την εκπαίδευση μοντέλων. Ένα τυπικό παράδειγμα χαρακτηριστικών είναι οι γωνίες. Η διαδικασία σχεδιασμού τέτοιων χαρακτηριστικών περιλάμβανε κοπιαστική παραμετροποίηση, ενώ τα αποτελέσματα αυτού του πλαισίου εργασίας φάνηκε να μην βελτιώνεται περαιτέρω προς το τέλος της δεκαετίας του 2000. Με την εισαγωγή των ΣΝΔ, η διαδικασία σχεδιασμού χαρακτηριστικών έγινε μέρος της διαδικασίας εκπαίδευσης. Τα ΣΝΔ υπολογίζουν τα δικά τους, ειδικά χαρακτηριστικά, τα οποία δεν είναι διαισθητικά κατανοητά από ανθρώπους, όπως είναι οι γωνίες.

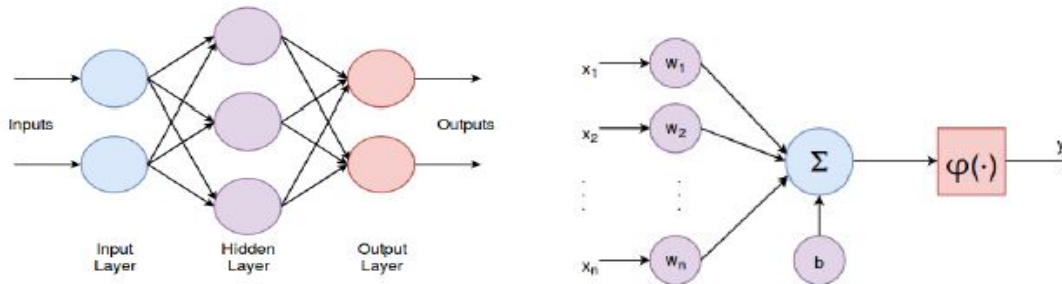
Μια ακόμα εργασία Μάθησης με Επίβλεψη, εκτός της ταξινόμησης, είναι η παλινδρόμηση, όπου, αντί για κατηγορία ως έξοδο, αναμένεται αριθμητική τιμή. Στο πλαίσιο της ανίχνευσης αντικειμένων από εικόνες, ένα παράδειγμα είναι να βρεθούν οι συντεταγμένες που ορίζουν ένα ορθογώνιο περιγράμμα γύρω από το αυτοκίνητο.

3.3 Μάθηση χωρίς Επίβλεψη

Σε άλλα προβλήματα μηχανικής μάθησης, υπάρχει μια διαφορετική τάξη καθηκόντων που αναφέρεται ως μη επιβλεπόμενη μάθηση. Στα προβλήματα αυτής της κατηγορίας, τα δεδομένα εκπαίδευσης είναι διανύσματα x τα οποία δεν έχουν αντίστοιχες ετικέτες. Επομένως, ο στόχος τη μη επιβλεπόμενης μάθησης είναι να βρίσκει μοτίβα όταν δεν υπάρχουν "σωστές απαντήσεις", ή όταν αυτές είναι αδύνατον να υπολογιστούν. Μία μεγάλη υποκατηγορία μη επιβλεπόμενων τεχνικών είναι το πρόβλημα της ομαδοποίησης (clustering). Η μέθοδος αυτή αναφέρεται στην ομαδοποίηση παρατηρήσεων με τέτοιο τρόπο ούτως ώστε τα μέλη μιας κοινής ομάδας να είναι παρόμοια το ένα με το άλλο, και να διαφέρουν σημαντικά από τα μέλη των άλλων ομάδων. Μια άλλη πολύ ενδιαφέρουσα κατηγορία μη επιβλεπόμενων καθηκόντων είναι οι τα γεννητικά μοντέλα (generative models). Τα μοντέλα αυτά μιμούνται τη διαδικασία δημιουργίας των δεδομένων εκπαίδευσης. Ένα καλό γεννητικό μοντέλο θα πρέπει να μπορεί να δημιουργήσει νέα δεδομένα τα οποία, αν και είναι τεχνητά, μοιάζουν με τα αυθεντικά. Αυτός ο τρόπος μάθησης είναι μη επιβλεπόμενος διότι η διαδικασία με την οποία δημιουργούνται ("γεννιούνται") τα δεδομένα δεν είναι άμεσα παρατηρήσιμη - μόνο τα ίδια τα δεδομένα είναι παρατηρήσιμα.

3.4 Νευρωνικά Δίκτυα

Ένα παράδειγμα συνάρτησης Μάθησης με Επίβλεψη είναι τα Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης (ΝΔΠΤ). Τα νευρωνικά αυτά δίκτυα αποτελούνται από ένα επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα και ένα επίπεδο εξόδου, σειριακά συνδεδεμένα. Το επίπεδο εισόδου παρέχει ένα διάνυσμα εισόδου στο δίκτυο, ενώ το επίπεδο εξόδου παρέχει την πρόβλεψη σαν μια τιμή ή διάνυσμα. Τα κρυφά επίπεδα ορίζουν την πολύπλοκη εσωτερική λειτουργία του δικτύου. Η γενική όψη τους παρουσιάζεται στο Σχήμα 3.1.



(α) Απλό Νευρωνικό Δίκτυο Πρόσθιας Τροφοδότησης

(β') Αρχιτεκτονική ενός Νευρώνα

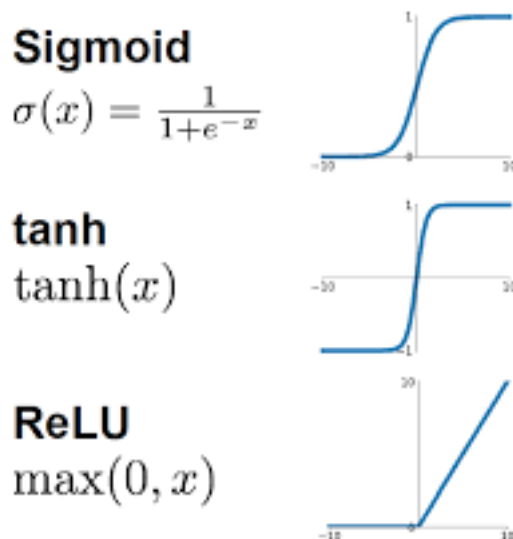
Σχήμα 3.1: Κάτοψη Νευρωνικού Δικτύου Πρόσθιας Τροφοδότησης

Η δομή των ΝΔΠΤ είναι ένας κατευθυνόμενος ακυκλικός γράφος, όπου κάθε κόμβος, ο Νευρώνας, είναι μια μη γραμμική συνάρτηση $R^n \rightarrow R$. Ο Νευρώνας υπολογίζει την έξοδο σε δύο βήματα - υπολογίζει το σταθμισμένο άθροισμα των n εισόδων του συν έναν όρο πόλωσης και στη συνέχεια εφαρμόζει μια συνάρτηση κανονικοποίησης $\varphi(\cdot)$ στο άθροισμα. Η τελευταία λέγεται συνάρτηση ενεργοποίησης. Υπάρχουν $n + 1$ παράμετροι σχετικές του κάθε Νευρώνα, τα n βάρη w_1, \dots, w_n και η πόλωση b , όπως φαίνεται κι από την εξίσωση 4.3. Η έξοδος κάθε Νευρώνα γίνεται είσοδος σε κάθε Νευρώνα στο επόμενο επίπεδο. Το ΝΔΠΤ μπορεί να ρυθμιστεί με εκπαίδευση με προσαρμογή του συνόλου των παραμέτρων των Νευρώνων.

$$y = \varphi \left(\sum_{j=0}^n w_j x_j + b \right) \quad (3.1)$$

Χρησιμοποιώντας μια μη γραμμική συνάρτηση ενεργοποίησης, το ΝΔΠΤ που προκύπτει γίνεται επίσης μη γραμμικό. Παραδείγματα συναρτήσεων ενεργοποίησης είναι η Σιγμοειδής (Sigmoid), η Υπερβολική Εφαπτομένη (Tanh) και η ReLU. Οι μη γραμμικές συναρτήσεις ενεργοποίησης κάνουν τα ΝΔΠΤ καθολικές προσεγγίσεις, δηλαδή μπορούν να προσεγγίσουν επαρκώς συνεχείς συναρτήσεις ορισμένες σε συμπαγή υποσύνολα του R^n επιλέγοντας κατάλληλες παραμέτρους και έναν επαρκή, αλλά πεπερασμένο, αριθμό νευρώνων. Αυτή η ιδιότητα προσθέτει μια θεωρητική δικαιολόγηση χρήσης ΝΔΠΤ για εργασίες Μάθησης με Επίβλεψη.

Η εκπαίδευση ΝΔΠΤ γίνεται με τον δημοφιλή αλγόριθμο οπισθοδιάδοσης. Η οπισθοδιάδοση είναι ένας αλγόριθμος βελτιστοποίησης που ψάχνει για παραμέτρους (βάρη



Σχήμα 3.2: Συνήθεις Συναρτήσεις Ενεργοποίησης

και πολώσεις) που ελαχιστοποιούν την συνάρτηση σφάλματος. Το σφάλμα ορίζεται ως μια συνάρτηση που μετράει την απόσταση (σφάλμα) μεταξύ των εξόδων του συνόλου εκπαίδευσης και των αντίστοιχων εξόδων υπολογισμένων από το ΝΔΠΤ χρησιμοποιώντας τις εισόδους του συνόλου εκπαίδευσης. Η διαισθητική επεξήγηση της οπισθοδιάδοσης είναι ότι το σφάλμα υπολογίζεται στην έξοδο και διανέμεται προς τα προηγούμενα επίπεδα του δικτύου, ανανεώνοντας τις παραμέτρους ανάλογα με την ευαισθησία τους σε αλλαγές του σφάλματος, χρησιμοποιώντας τον κανόνα παραγωγίσιμης αλυσίδας.

3.5 Γραμμική Ταξινόμηση

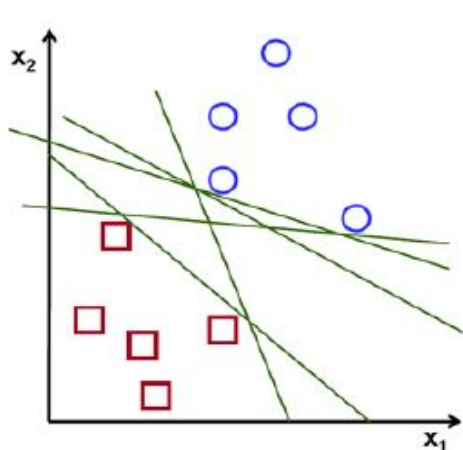
3.5.1 Μηχανές Διανυσμάτων Υποστήριξης

Μια μηχανή διανυσμάτων υποστήριξης (SVM) είναι ένας διακριτικός ταξινομητής που ορίζεται τυπικά από ένα διαχωριστικό υπερεπίπεδο. Με άλλα λόγια, δεδομένου ότι τα δεδομένα κατάρτισης είναι επισημασμένα (επιβλεπόμενη μάθηση), ο αλγόριθμος εξάγει ένα βέλτιστο υπερεπίπεδο που κατηγοριοποιεί νέα παραδείγματα. Σε δύο διαστάσεις, αυτό το υπερεπίπεδο είναι μια γραμμή που διαιρεί ένα επίπεδο σε δύο μέρη όπου κάθε τάξη βρίσκεται σε κάθε πλευρά. Ας υποθέσουμε ότι έχουμε ένα δυαδικό πρόβλημα ταξινόμησης με γραμμικά μοντέλα της μορφής:

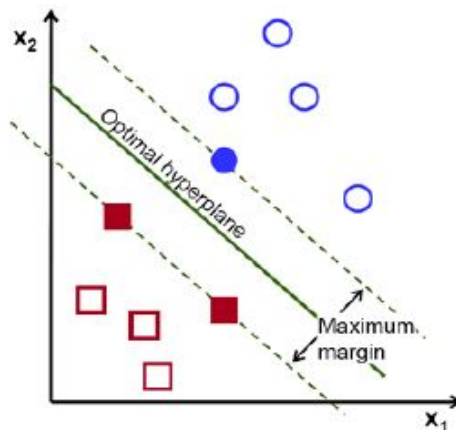
$$f(x) = w^T \cdot \varphi(x) + b \quad (3.2)$$

όπου $\varphi(x)$ είναι ένα μετασχηματισμός στο χώρο χαρακτηριστικών και η παράμετρος b (bias) έχει οριστεί. Το σύνολο δεδομένων εκπαίδευσης αποτελείται από N διανύσματα εισόδου x_1, \dots, x_N με αντίστοιχες τιμές εξόδου y_1, \dots, y_N όπου $y_i \in \{-1, 1\}$, και νέα δείγματα x ταξινομούνται ανάλογα με το πρόσημο της $f(x)$.

Υποθέτουμε ότι τα δεδομένα εκπαίδευσης είναι γραμμικά διαχωρίσιμα στο χώρο χαρακτηριστικών, ούτως ώστε εξ ορισμού υπάρχει τουλάχιστον μία επιλογή παραμέτρων w και b τέτοια ώστε μία συνάρτηση της μορφής της παραπάνω εξίσωσης ικανοποιεί την ανισότητα $f(x_i) > 0$ για δεδομένα που έχουν $y_i = +1$ και $f(x_i) < 0$ για δεδομένα που έχουν $y_i = -1$, έτσι ώστε $y_i f(x_i) > 0$ για όλα τα δεδομένα εκπαίδευσης. Ένα παράδειγμα δεδομένων εκπαίδευσης φαίνεται στο Σχήμα 3.3 αριστερά, όπου τα δείγματα που ανήκουν στην πρώτη κατηγορία είναι τα κόκκινα τετράγωνα, ενώ τα δείγματα που ανήκουν στην δεύτερη κατηγορία είναι οι μπλε κύκλοι.



(α) Πιθανά υπερεπίπεδα διαχωρισμού.



(β) Ένας γραμμικός ταξινομητής μέγιστου περιθωρίου (maximum margin).

Σχήμα 3.3: Παράδειγμα δυαδικού διαχωρισμού δύο γραμμικά διαχωρίσιμων κλάσεων.

Υπάρχει άπειρο πλήθος πιθανών ευθειών που διαχωρίζουν τις δύο κλάσεις. Ο στόχος του αλγορίθμου SVM είναι να βρει τον πιο γενικό ταξινομητή. Με άλλα λόγια, ο αλγόριθμος SVM προσπαθεί να βρει το υπερεπίπεδο για το οποίο η ελάχιστη απόσταση μεταξύ των δύο κλάσεων (margin) έχει την μέγιστη δυνατή τιμή. Το υπερεπίπεδο που ικανοποιεί την παραπάνω απαίτηση είναι το βέλτιστο υπερεπίπεδο και μπορεί να παρατηρηθεί στο Σχήμα 3.3 δεξιά.

Αν η $f(x)$ διαχωρίζει τα δείγματα, η γεωμετρική απόσταση μεταξύ ενός σημείου x_i και του υπερεπιπέδου $f(x) = 0$ είναι ίση με $\frac{|f(x_i)|}{\|w\|}$. Επιπλέον, ενδιαφερόμαστε μόνο για λύσεις για τις οποίες όλα τα δείγματα ταξινομούνται σωστά, ούτως ώστε $y_i f(x_i) > 0$ για κάθε i . Έπειτα, η απόσταση μεταξύ ενός σημείου x_i και του βέλτιστου υπερεπιπέδου δίνεται από:

$$\frac{y_i f(x_i)}{\|w\|} = \frac{y_i (w^T \varphi(x_i) + b)}{\|w\|} \quad (3.3)$$

Το περιθώριο δίνεται από την κάθετη απόσταση στο κοντινότερο σημείο x_i από το σύνολο δεδομένων, και επιθυμούμε να βελτιστοποιήσουμε τις παραμέτρους w και b για να μεγιστοποιήσουμε αυτή την απόσταση. Οπότε, το μέγιστο περιθώριο βρίσκεται λύνοντας την:

$$L(w, \mathbf{b}) = \arg_{w, \mathbf{b}} \max \left\{ \frac{1}{\|w\|} \min y_i (w^T \varphi(x_i) + \mathbf{b}) \right\} \quad (3.4)$$

Η μεγιστοποίηση $\frac{1}{\|w\|}$ ισοδυναμεί με ελαχιστοποίηση της $\frac{1}{2}\|w\|^2$. Το πρόβλημα τώρα μετασχηματίζεται ως εξής:

$$L(w) = \arg_{w, \mathbf{b}} \min \left\{ \frac{1}{2}\|w\|^2 \right\} \quad (3.5)$$

$$y_i (w^T \varphi(x_i) + \mathbf{b}) \geq 1, \quad i = 1, \dots, N. \quad (3.6)$$

Η λύση της Εξίσωσης 3.5 δίνεται από τους πολλαπλασιαστές Lagrange [19].

Παράμετρος C

Η παράμετρος Regularization (συχνά ονομαζόμενη παράμετρος C) αναφέρει στη βελτιστοποίηση SVM πόσο θέλετε να αποφύγετε την εσφαλμένη ταξινόμηση κάθε παραδείγματος εκπαίδευσης. Για τις μεγάλες τιμές του C, η βελτιστοποίηση θα επιλέξει ένα αεροπλάνο μικρότερου περιθωρίου αν αυτό το υπερπληρωμένο κάνει καλύτερη δουλειά να πάρει όλα τα σημεία εκπαίδευσης ταξινομημένα σωστά. Αντίστροφα, μια πολύ μικρή τιμή του C θα αναγκάσει το βελτιστοποιητή να αναζητήσει ένα μεγαλύτερο περιθώριο που διαχωρίζει το υπερπύρην, ακόμα και αν αυτό το υπερπληθυσμό ταξινομεί εσφαλμένα περισσότερα σημεία.

Παράμετρος γ

Ορίζει πόσο μακριά φτάνει η επιρροή ενός μόνο εκπαιδευτικού παραδείγματος. Εάν έχει χαμηλή τιμή, σημαίνει ότι κάθε σημείο έχει μια μεγάλη απόσταση και αντίστροφα υψηλή τιμή γ σημαίνει ότι κάθε σημείο έχει κοντά. Αν το γ έχει πολύ μεγάλη αξία, τότε το όριο απόφασης θα εξαρτηθεί απλώς από τα σημεία που βρίσκονται πολύ κοντά στη γραμμή, πράγμα που οδηγεί στην παραβίαση ορισμένων σημείων που βρίσκονται πολύ μακριά από το όριο απόφασης. Αυτό οφείλεται στο γεγονός ότι τα πιο κοντινά σημεία παίρνουν περισσότερο βάρος και οδηγούν σε μια κυματοειδή καμπύλη όπως φαίνεται στο προηγούμενο γράφημα. Από την άλλη πλευρά, αν η τιμή γ είναι χαμηλή, ακόμη και τα μακρινά σημεία παίρνουν σημαντικό βάρος και έχουμε μια πιο γραμμική καμπύλη.

3.5.2 Λογιστική Παλινδρόμηση

Σε προβλήματα ταξινόμησης, θέλουμε να καθορίσουμε την πιθανότητα μια παρατήρηση να ανήκει ή όχι σε μια συγκεκριμένη κλάση. Επομένως, επιθυμούμε να εκφράσουμε την πιθανότητα με μια τιμή μεταξύ του 0 και του 1. Ένας απλός αλγόριθμος ταξινόμησης που δημιουργεί τιμές αυτής της μορφής είναι ο ταξινομητής λογιστικής παλινδρόμησης.

Ας υποθέσουμε ότι έχουμε ένα απλό πρόβλημα δυαδικής ταξινόμησης, όπως αυτό που περιγράφηκε νωρίτερα στο ίδιο Κεφάλαιο. Έστω $x = x_1, \dots, x_N$ τα διανύσματα εισόδου όπου $y_i \in \{0, 1\}$. Η συνάρτηση ενεργοποίησης του LR ταξινομητή καθορίζεται από την εφαρμογή μιας σιγμοειδούς συνάρτησης πάνω στην γραμμική παλινδρόμηση ούτως ώστε να λάβουμε την τελική απόφαση ταξινόμησης. Όπως περιγράφηκε στα SVMs:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (3.7)$$

Η συνάρτηση ενεργοποίησης της LR για ένα δοσμένο διάνυσμα x ορίζεται ως εξής:

$$h_w(x) = \sigma(w^T x) = \frac{1}{1 + e^{-w^T x}} \quad (3.8)$$

Η συνάρτηση κόστους που θέλουμε να ελαχιστοποιηθεί κατά τη διάρκεια της εκπαίδευσης είναι η εξής:

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \log(\exp(-y_i(w^T x_i + b)) + 1) \quad (3.9)$$

όπου $C > 0$ και b είναι οι συντελεστές που αναπαριστούν την τιμωρία (penalty) των λανθασμένων αποτελεσμάτων ταξινόμησης και την τομή του υπερεπιπέδου αντίστοιχα.

3.6 Ταξινόμηση Εικόνων

Οι αρχιτεκτονικές που θα χρησιμοποιήσουμε στην παρούσα εργασία ονομάζονται βαθιά συνελικτικά νευρωνικά δίκτυα (Deep Convolutional Neural Networks - CNNs). Αυτά σχηματίζονται από μια σειριακή ακολουθία μη γραμμικών επεξεργαστικών μονάδων με σκοπό την εξαγωγή χαρακτηριστικών. Οι μονάδες αυτές ονομάζονται επίπεδα και όσο περισσότερα επίπεδα έχει ένα δίκτυο στη σειρά, τόσο πιο βαθύ είναι. Μερικές βασικές έννοιες στα δίκτυα αυτά είναι η αραιή συνδεσιμότητα μεταξύ των νευρώνων των διαφόρων επιπέδων και τα η χρήση κοινών βαρών σε γειτονικούς νευρώνες. Αυτό βοηθάει τα δίκτυα με δύο τρόπους: Αφενός μειώνει πολύ τον αριθμό των παραμέτρων, πράγμα που κάνει την εκπαίδευση εφικτή από υπολογιστικής άποψης και αφετέρου επιτρέπει στα δίκτυα να εξάγουν τοπικά χαρακτηριστικά. Τα Συνελικτικά Νευρωνικά Δίκτυα είναι ένα είδος Νευρωνικών Δικτύων Πρόσθιας Τροφοδότησης. Η είσοδος, για ταξινόμηση εικόνων, είναι ένας τένσορας εικόνας, ενώ τα κρυφά επίπεδα αποτελούνται από δύο μέρη - το δίκτυο εξαγωγής χαρακτηριστικών και το δίκτυο ταξινόμησης. Το δίκτυο εξαγωγής χαρακτηριστικών περιέχει επίπεδα συνέλιξης και υποδειγματοληψίας. Το δίκτυο ταξινόμησης περιέχει επίπεδα νευρώνων όπως στην

περίπτωση του απλού ΝΔΠΤ, που αναφέρονται ως πλήρως συνδεδεμένα επίπεδα. Η έξοδος είναι ένα n -διάστατο διάνυσμα πιθανοτήτων για τις n προκαθορισμένες κλάσεις.

3.6.1 Επίπεδα Συνελικτικών Νευρωνικών Δικτύων

Επίπεδο Συνέλιξης

Το Επίπεδο Συνέλιξης συνελίσσει τον τένσορα εισόδου με ένα πυρήνα (φίλτρο) ίδιου βάρους. Για παράδειγμα, για μια $32 \times 32 \times 3$ εικόνα RGB, κατάλληλοι πυρήνες είναι οι $3 \times 3 \times 3$, $5 \times 5 \times 3$ κτλ. Η έξοδος κάθε τέτοιας πράξης ονομάζεται χάρτης ενεργοποίησης. Σε όρους όρασης υπολογιστών, η έξοδος είναι μια φιλτραρισμένη εικόνα. Νευρώνες σε αυτή την περίπτωση θεωρούνται όλες οι θέσεις της εικόνας όπου το φίλτρο μπορεί να εφαρμοστεί. Οι παράμετροι ενός επιπέδου συνέλιξης που μπορούν να εκπαιδευτούν είναι τα στοιχεία του πυρήνα και μια παράμετρος πόλωσης, που είναι κοινά για όλους τους νευρώνες. Σύμφωνα με τον τυπικό ορισμό της διακριτής συνέλιξης, υπολογίζεται ολισθαίνοντας έναν πυρήνα ένα χωρικό βήμα τη φορά. Τα συνελικτικά επίπεδα μπορούν να χρησιμοποιήσουν ένα βήμα ολίσθησης s μεγαλύτερο του ενός όπου το s αναφέρεται ως *stride* του επιπέδου. Υπάρχουν συνήθως περισσότεροι από έναν πυρήνες ανά συνελικτικό επίπεδο, δίνοντας ως έξοδο πολλαπλούς χάρτες ενεργοποίησης. Έτσι, τυπικά, η έξοδος ενός συνελικτικού επιπέδου είναι ένας τρισδιάστατος τένσορας μεγάλου βάρους.

Επίπεδο Υποδειγματοληψίας

Τα επίπεδα υποδειγματοληψίας τοποθετούνται συνήθως μετά από επίπεδα συνέλιξης και μειώνουν το πλάτος και ύψος των χαρτών ενεργοποίησης. Διαχωρίζουν τον τένσορα εισόδου σε ένα σύνολο μη επικαλυπτόμενων ορθογωνίων, και για κάθε υποπεριοχή, δίνουν ως έξοδο μια αντιπροσωπευτική τιμή. Αυτή η τιμή μπορεί να υπολογιστεί ως το μέγιστο (υποδειγματοληψία μεγίστου), ο μέσος όρος (υποδειγματοληψία μέσου όρου) ή ακόμα και με τυχαίο τρόπο (στοχαστική υποδειγματοληψία). Τα επίπεδα υποδειγματοληψίας χρησιμοποιούνται για μείωση του μεγέθους των χαρτών ενεργοποίησης, κι έτσι μειώνεται ο αριθμός των παραμέτρων και οι υπολογισμοί στο δίκτυο. Η μείωση των παραμέτρων μειώνει και τις πιθανότητες εμφάνισης υπερπροσαρμογής. Χρησιμοποιώντας επίπεδα υποδειγματοληψίας, η αναπαράσταση των χαρακτηριστικών της εικόνας συνοψίζεται σε μικρότερες διαστάσεις. Αυτό δίνει τη δυνατότητα σε επόμενους συνελικτικούς πυρήνες να συσχετίσουν μακρινές περιοχές της εικόνας εισόδου ενώ διατηρείται ένα μικρό μέγεθος πυρήνα. Σε ορολογία ΣΝΔ, οι πράξεις υποδειγματοληψίας αυξάνουν το υποδεκτικό πεδίο των πυρήνων συνέλιξης.

Επίπεδα Ενεργοποίησης

Τα επίπεδα ενεργοποίησης δρουν πάνω σε χάρτες ενεργοποίησης εφαρμόζοντας μια μη γραμμική απεικόνιση στα στοιχεία τους. Η ReLU είναι η συνήθης επιλογή για την απεικόνιση, η οποία, εμπειρικά, μειώνει το χρόνο εκπαίδευσης.

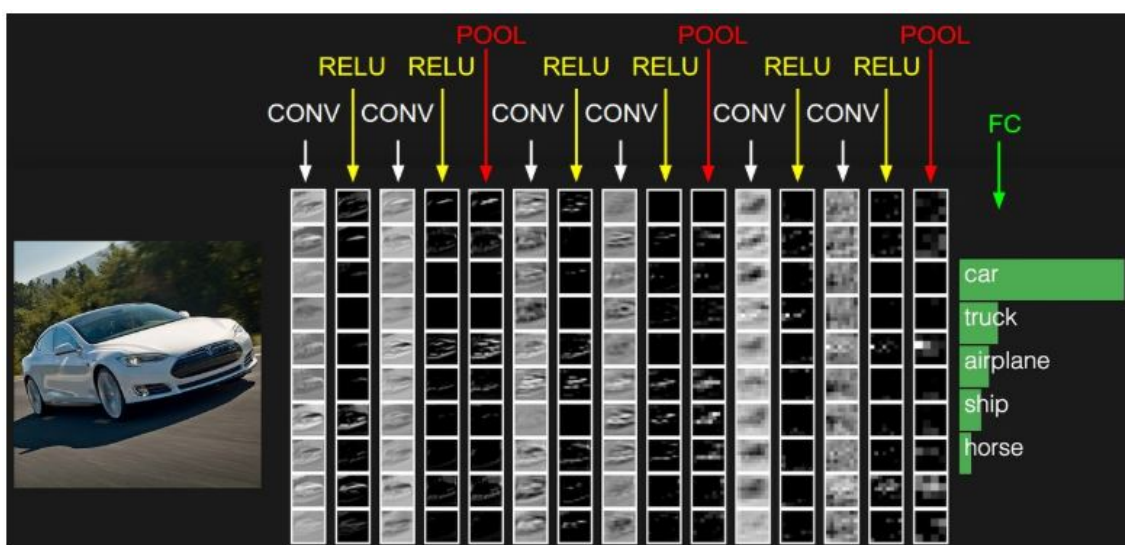
Πλήρως Συνδεδεμένα Επίπεδα

Η έξοδος του δικτύου εξαγωγής χαρακτηριστικών ενός ΣΝΔ είναι ένας τρισδιάστατος τένσορας μεγάλου βάρους, π.χ. $6 \times 6 \times 256$. Αναδιατάσσοντας τον τένσορα

σε διάνυσμα, μπορεί να χρησιμοποιηθεί σαν είσοδος σε ένα απλό, βαθύ ΝΔΠΤ. Τα αντίστοιχα κρυφά επίπεδα του ΝΔΠΤ ονομάζονται Πλήρως Συνδεδεμένα Επίπεδα σε αυτήν την περίπτωση.

3.6.2 Εξαγωγή Χαρακτηριστικών

Τα ΣΝΔ έχουν δείξει ενθαρρυντικές επιδόσεις σε πολλά προβλήματα και εφαρμογές στον τομέα της όρασης του υπολογιστή, όπως η ανίχνευση και εντοπισμός αντικειμένων [20], [21], ο διαχωρισμός των εικόνων, η ταξινόμηση [22] και η ανάκτηση [23]. Αυτή η ευρεία επιτυχία οφείλεται στην ιεραρχική δομή τους, η οποία αποκτά αυτόματα πολύ δυνατά χαρακτηριστικά από τα πρωτογενή δεδομένα (raw data). Το ακόλουθο Σχήμα 3.4 απεικονίζει τη διαδικασία με την οποία το δίκτυο "μαθαίνει" τα χαρακτηριστικά:



Σχήμα 3.4: Εξαγωγή χαρακτηριστικών από CNN

3.6.3 Συνήθεις Αρχιτεκτονικές Συνελικτικών Νευρωνικών Δικτύων

Παρακάτω παρουσιάζονται οι πιο γνωστές και ευρέως χρησιμοποιούμενες αρχιτεκτονικές βαθιών συνελικτικών νευρωνικών δικτύων.

LeNet

Οι πρώτες επιτυχημένες εφαρμογές των Συνελικτικών Δικτύων αναπτύχθηκαν από τον Yann LeCun το 1990. Από αυτές, η πιο γνωστή είναι η αρχιτεκτονική LeNet που χρησιμοποιείται για την ανάγνωση ταχυδρομικών κωδικών, ψηφίων κλπ [24].

AlexNet

Το ΣΝΔ AlexNet [25] συμμετείχε στον διαγωνισμό ImageNet Large Scale Visual Recognition Challenge (ILSVRC) το 2012, επιτυγχάνοντας ένα top-5 σφάλμα

15.3%, ενώ η δεύτερη καλύτερη συμμετοχή που χρησιμοποιούσε μη βαθιά αρχιτεκτονική, έδωσε 26.2% σφάλμα. Το AlexNet είναι προσέγγιση ορόσημο που όρισε εκ νέου την εφαρμογή της βαθιάς μάθησης στην όραση υπολογιστών. Αυτοματοποιώντας την διαδικασία σχεδίασης χαρακτηριστικών, έκανε δημοφιλή την ταξινόμηση εικόνας και άλλων πεδίων όρασης υπολογιστών σε μη ειδικούς. Το AlexNet περιέχει πέντε επίπεδα συνέλιξης, τρία επίπεδα υποδειγματοληψίας μεγίστου, τρία πλήρως συνδεδεμένα επίπεδα και δύο επίπεδα κανονικοποίησης. Αυτή η βαθιά αρχιτεκτονική των 60 εκατομμυρίων παραμέτρων εκπαιδεύτηκε χρησιμοποιώντας την Compute Unified Device Architecture (CUDA) [26] σε δύο GTX 580 GPU.

ZFNet

Ο νικητής του ILSVRC 2013 ήταν ένα συνελικτικό νευρωνικό δίκτυο από τους Matthew Zeiler και Rob Fergus. Έγινε γνωστός ως ZFNet (σύντομη για Zeiler & Fergus Net). Ήταν μια βελτίωση για το AlexNet, χάρη στην εξισορρόπηση των υπερπαραμέτρων της αρχιτεκτονικής, ιδίως με τη διεύρυνση του μεγέθους των μέσω συνελικτικών στρωμάτων και τη μείωση του μεγέθους του φίλτρου και του φίλτρου στο πρώτο στρώμα [27].

GoogLeNet/Inception

Μετά το AlexNet, οι αρχιτεκτονικές ΣΝΔ κυριάρχησαν στην έρευνα, σε εφαρμογές και σε διαγωνισμούς αναγνώρισης εικόνας, χρησιμοποιώντας σειριακά συνδεδεμένα επίπεδα. Το 2015, το GoogLeNet [28], νικήτρια αρχιτεκτονική του ILSVRC 2014 (6.7% top-5 σφάλμα), εισήγαγε τη μονάδα inception. Οι μονάδες inception διαδίδουν τις ενεργοποιήσεις του προηγούμενου επιπέδου σε συνελικτικά επίπεδα πολλαπλών μεγεθών πυρήνα, παράλληλα, και στη συνέχεια συνδέουν τις ενεργοποιήσεις. Η διαισθητική εξήγηση πίσω από αυτήν την αρχιτεκτονική είναι ότι η σχεδιαστική επιλογή του μεγέθους πυρήνα συνέλιξης μεταφέρεται στην διαδικασία εκπαίδευσης. Οι μονάδες inception περιέχουν συνελικτικά επίπεδα 1×1 για μείωση της διαστατικότητας εξόδου, και δε χρησιμοποιούνται πλήρως συνδεδεμένα επίπεδα ώστε να μειωθούν οι παράμετροι του μοντέλου. Έτσι, αν και περιέχει 22 επίπεδα, το GoogLeNet έχει 12 φορές λιγότερες παραμέτρους από το AlexNet. Η πιο σύγχρονη εφαρμογή του GoogLeNet είναι το Inception-v4 [29].

VGGNet

Ο φιναλίστ στο ILSVRC 2014 ήταν το δίκτυο από την Karen Simonyan και τον Andrew Zisserman που έγινε γνωστό ως VGGNet. Η κύρια συμβολή του ήταν να δείξει ότι το βάθος του δικτύου είναι ένα κρίσιμο στοιχείο για καλές επιδόσεις. Το τελικό τους καλύτερο δίκτυο περιέχει 16 στρώματα CONV/FC και, ελκυστικά, διαθέτει μια εξαιρετικά ομοιογενή αρχιτεκτονική που εκτελεί μόνο συσπειρώσεις 3×3 και συγκέντρωση 2×2 από την αρχή μέχρι το τέλος. Το πρόδρομο μοντέλο τους είναι διαθέσιμο για χρήση με plug and play στο Caffe. Ένα μειονέκτημα του VGGNet είναι ότι είναι ακριβότερο να αξιολογείται και χρησιμοποιεί πολύ περισσότερη μνήμη και παραμέτρους (140M). Οι περισσότερες από αυτές τις παραμέτρους είναι στην πρώτη πλήρως συνδεδεμένη στρώση και από τότε διαπιστώθηκε ότι αυτά τα στρώματα FC μπορούν να αφαιρεθούν χωρίς υποβάθμιση της απόδοσης, μειώνοντας σημαντικά τον αριθμό των απαραίτητων παραμέτρων [30].

ResNet

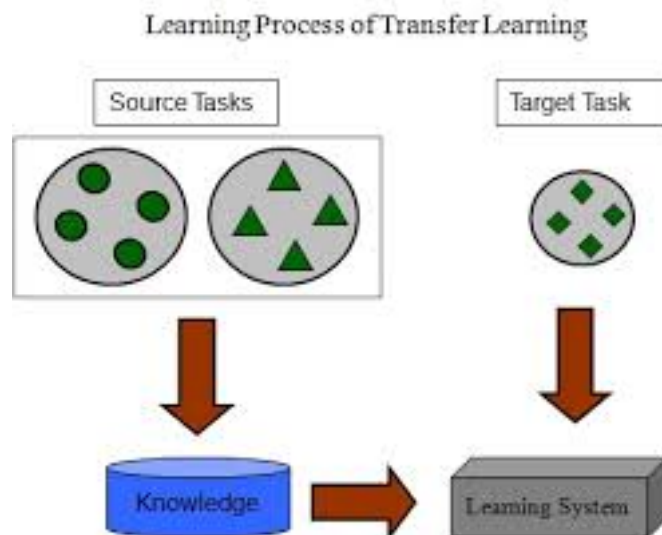
Η αρχιτεκτονική ResNet (Residual Network) [31] ήταν η νικήτρια του ILSVRC 2015 χρησιμοποιώντας 122 επίπεδα. Οι συγγραφείς ισχυρίστηκαν ότι αν απλοϊκά αυξηθεί ο αριθμός των επιπέδων, η επίδοση του δικτύου θα μειωθεί λόγω της δυσκολίας βελτιστοποίησης με οπισθοδιάδοση. Αυτό το τεχνικό πρόβλημα είναι γνωστό σαν εξαφανιζόμενη κλίση (vanishing gradient). Αντιμετώπισαν αυτό το ζήτημα με τη μονάδα residual block. Το residual block περιέχει δύο σειριακά συνδεδεμένα επίπεδα συνέλιξης και μια σύνδεση (skip connection) που αθροίζει την είσοδο στην έξοδο πριν το επίπεδο ενεργοποίησης. Αυτή η σύνδεση θεωρείται ότι δρα σαν μια απεικόνιση αναφοράς που διευκολύνει την διαδικασία μάθησης, ενώ προσφέρει καλύτερη κατανομή της κλίσης σφάλματος προς τα πρώτα επίπεδα. Κατά την περίοδο συγγραφής αυτής της διπλωματικής εργασίας, τα ResNet εξακολουθούν να είναι η σύγχρονη προσέγγιση σε εργασίες βαθιάς μάθησης στην όραση υπολογιστών, όπως φαίνεται και σε πιο πρόσφατες εργασίες [32].

3.7 Μεταφορά Μάθησης

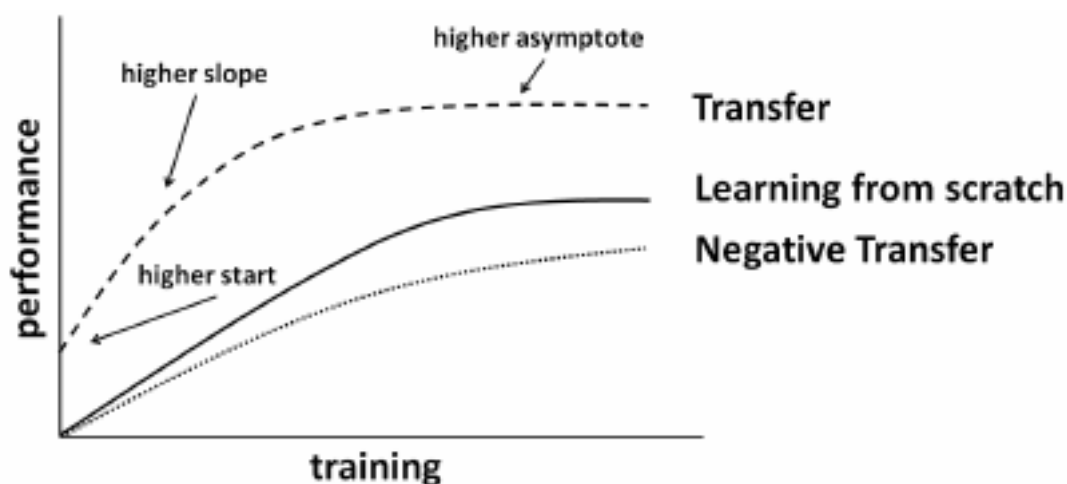
Μεταφορά μάθησης (Transfer Learning) ονομάζουμε το πρόβλημα της μηχανικής μάθησης, στο οποίο προσπαθούμε να αξιοποιήσουμε την γνώση που απέκτησε ένα σύστημα σε ένα πρόβλημα, σε ένα διαφορετικό αλλά σχετικό πρόβλημα. Η μέθοδος αυτή αξιοποιείται πολύ από τα βαθιά νευρωνικά δίκτυα, γιατί αυτά απαιτούν μεγάλο αριθμό δεδομένων για να εκπαιδευτούν. Θεωρητικά, άμα δεν έχουμε αρκετά δεδομένα για ένα πρόβλημα, μπορούμε να εκπαιδύσουμε ένα δίκτυο σε ένα σχετικό πρόβλημα, στο οποίο έχουμε περισσότερα δεδομένα και να χρησιμοποιήσουμε τη γνώση που απέκτησε στο αρχικό μας πρόβλημα. Στη μηχανική μάθηση (και ειδικά στη βαθιά μάθηση), αντιμετωπίζουμε ένα πολύ σημαντικό πρόβλημα. Αυτό είναι το γεγονός ότι τα δίκτυα που επιλύουν περίπλοκα προβλήματα απαιτούν τεράστιες ποσότητες δεδομένων. Ωστόσο, η απόκτηση αυτών των δεδομένων για τα επιβλεπόμενα μοντέλα είναι συχνά ανέφικτη λόγω χρονικών ή υπολογιστικών περιορισμών. Επιπλέον, τα μοντέλα που έχουν εκπαιδευτεί σε μικρά, ειδικά σύνολα δεδομένων έχουν χειρότερη απόδοση όταν χρησιμοποιούνται για να αντιμετωπίσουν ένα διαφορετικό πρόβλημα, το οποίο μπορεί να είναι σχετικά παρεμφερές με το πρόβλημα στο οποίο έχουν εκπαιδευτεί.

Ο στόχος της μεταφοράς μάθησης είναι να βελτιώσει την εκμάθηση του προβλήματος-στόχου (target task) αξιοποιώντας γνώση από το πρόβλημα-πηγή (source task), όπως φαίνεται στο Σχήμα 3.5.

Υπάρχουν τρεις τρόποι με τους οποίους συνήθως η μεταφορά μάθησης βελτιώνει τη διαδικασία εκπαίδευσης, οι οποίοι φαίνονται στο Σχήμα 3.6. Πρώτον, η αρχική απόδοση που επιτυγχάνεται στο target task χρησιμοποιώντας μόνο τη γνώση που έχει μεταφερθεί από το source task, προτού εκπαιδευτεί παραπάνω, σε σχέση με την αρχική απόδοση ενός τυχαία αρχικοποιημένου μοντέλου. Δεύτερον, ο χρόνος που χρειάζεται για να εκπαιδευτεί πλήρως το μοντέλο στο target task δεδομένης της γνώσης που έχει μεταφερθεί, σε σχέση με το χρόνο που χρειάζεται για να το μάθει εξ αρχής. Τρίτον, το τελικό επίπεδο απόδοσης που επιτυγχάνεται στο target task σε σχέση με το τελικό επίπεδο χωρίς μεταφορά μάθησης [33].



Σχήμα 3.5: Μέθοδος Μεταφοράς Μάθησης



Σχήμα 3.6: Οι 3 τρόποι με τους οποίους η μεταφορά μάθησης βελτιώνει την μάθηση

Η περίπτωση της μη-επιβλεπόμενης προεκπαίδευσης (unsupervised pretraining)

Μία συγκεκριμένη περίπτωση μεταφοράς μάθησης είναι όταν το source task είναι μη επιβλεπόμενο και το target task είναι επιβλεπόμενο. Αυτή η περίπτωση έχει ιδιαίτερο ενδιαφέρον, καθώς πολύ συχνά έχουμε διαθέσιμες μεγάλες ποσότητες μη επιβλεπόμενων δεδομένων εκπαίδευσης, αλλά πολύ λίγα δεδομένα εκπαίδευσης με ετικέτες. Η εκπαίδευση με επιβλεπόμενες τεχνικές στο επισήμασμένο υποσύνολο πολλές φορές οδηγεί σε overfitting. Αποκτώντας ποιοτικές αναπαραστάσεις από τα μη επιβλεπόμενα δεδομένα, το μοντέλο μας μπορεί να έχει καλύτερη απόδοση στο πρόβλημα επιβλεπόμενης μάθησης που αντιμετωπίζουμε [34].

Αυτή η περίπτωση μεταφοράς μάθησης ονομάζεται μη-επιβλεπόμενη προεκπαίδευση (unsupervised pretraining). Αυτή η διαδικασία αποτελεί παράδειγμα του πώς μια αναπαράσταση που έχει δημιουργηθεί από το μοντέλο, όταν αυτό αντιμετωπίζει ένα συγκεκριμένο πρόβλημα (μη επιβλεπόμενο) μπορεί κάποιες φορές να είναι χρήσιμη για ένα άλλο πρόβλημα (επιβλεπόμενο). Ονομάζεται προεκπαίδευση (pretraining), επειδή

αποτελεί μόνο το πρώτο βήμα προτού ένας αλγόριθμος εκπαίδευσης εφαρμοστεί για να προσαρμόσει (fine-tune) όλα τα επίπεδα μαζί. Ως προς το πρόβλημα επιβλεπόμενης μάθησης, μπορεί να θεωρηθεί ένας όρος κανονικοποίησης και αρχικοποίησης των παραμέτρων.

Κανονικοποίηση

Είναι πιθανό ότι η προεκπαίδευση αρχικοποιεί ένα βαθύ νευρωνικό δίκτυο σε μία περιοχή που θα ήταν αλλιώς απροσπέλαστη - για παράδειγμα, μια περιοχή που περιτριγυρίζεται από περιοχές όπου η συνάρτηση κόστους εναλλάσσεται τόσο πολύ από το ένα παράδειγμα στο άλλο που μπορεί να υπολογιστεί μόνο μια εκτίμηση του gradient που περιέχει πολύ θόρυβο.

Αρχικοποίηση παραμέτρων

Η προεκπαίδευση, στις περισσότερες περιπτώσεις, βελτιώνει την απόδοση στο επιβλεπόμενο πρόβλημα. Η βασική ιδέα είναι πως κάποια χαρακτηριστικά που είναι χρήσιμα για την επίλυση του μη επιβλεπόμενου προβλήματος είναι επίσης χρήσιμα και για τη ν επίλυση του επιβλεπόμενου προβλήματος.

Κεφάλαιο 4

Σχετικές Εργασίες

Στο παρόν κεφάλαιο αναλύονται state-of-the-art τεχνικές ανίχνευσης φωτιάς από μικρά συστήματα, μέχρι πολυδιάστατα συνελικτικά νευρωνικά δίκτυα. Πιο συγκεκριμένα, περιγράφονται τεχνικές οι οποίες έχουν ως σκοπό την αναγνώριση φωτιάς, καπνού αλλά και συνδυασμού τους, βασισμένες σε εικόνες και βίντεο. Επίσης, παρουσιάζονται και συστήματα που δουλεύουν σε πραγματικό χρόνο (real-time).

4.1 Ανίχνευση Φωτιάς

Η ανίχνευση της πυρκαγιάς σε πρώιμο στάδιο προτού να μετατραπεί σε καταστροφικό φαινόμενο είναι ζωτικής σημασίας για την πρόληψη της καταστροφής και για τη διάσωση της ζωής και των ιδιοκτησιών των ανθρώπων. Παρόλο που οι αισθητήρες ανίχνευσης πυρκαγιάς και καπνού είναι ευρέως εγκατεστημένοι σε εσωτερικούς χώρους, απαιτούν γενικά η φωτιά να καεί για λίγο για να παράγει μεγάλη ποσότητα καπνού και στη συνέχεια να ενεργοποιήσει τον συναγερμό. Εκτός αυτού, αυτές οι συσκευές δεν μπορούν να αναπτυχθούν σε υπαίθρια περιβάλλοντα μεγάλης κλίμακας, π.χ. δάσος, άγρια περιοχή. Αντίθετα, το σύστημα ανίχνευσης πυρκαγιάς βασισμένο σε όραση υπολογιστών συλλαμβάνει εικόνες από κάμερες και ανιχνεύει άμεσα φωτιά, επομένως είναι κατάλληλο για έγκαιρη ανίχνευση πυρκαγιάς. Ένα τέτοιο σύστημα είναι επίσης φθηνό και εύκολο στην εγκατάσταση. Λόγω των ταχείων εξελίξεων στην τεχνολογία των ψηφιακών φωτογραφικών μηχανών και των τεχνικών επεξεργασίας βίντεο, υπάρχει μια σημαντική τάση να αντικατασταθούν οι συνήθεις μέθοδοι ανίχνευσης πυρκαγιάς με συστήματα βασισμένα σε όραση υπολογιστή. Οι βιντεοσκοπημένες τεχνικές ανίχνευσης πυρκαγιάς είναι κατάλληλες για την ανίχνευση φωτιάς σε μεγάλους και ανοιχτούς χώρους. Επιπλέον, χάρη σε αυτά τα συστήματα μπορεί κανείς να αναλύσει τη συμπεριφορά της φωτιάς και να εκτελέσει έναν τρισδιάστατο εντοπισμό της φωτιάς. Επιπλέον, τα συστήματα επιτήρησης τηλεόρασης κλειστού κυκλώματος (CCTV) εγκαθίστανται σήμερα σε διάφορα σημεία παρακολούθησης σε εσωτερικούς και εξωτερικούς χώρους. Υπό αυτήν την περίπτωση, θα ήταν επωφελές να αναπτυχθεί ένα σύστημα ανίχνευσης πυρκαγιάς το οποίο θα μπορούσε να χρησιμοποιήσει αυτόν τον υπάρχοντα εξοπλισμό χωρίς να εισάγει επιπλέον κόστος.

4.1.1 Κλασικές Προσεγγίσεις

Στην εργασία των Zhang et al. [35], προτείνεται μια μέθοδο ανίχνευσης πυρκαγιάς που χρησιμοποιεί αλγόριθμους μηχανικής μάθησης και όρασης υπολογιστών, η οποία

μπορεί να λειτουργήσει με μη στατική κάμερα. Το σύστημα ανίχνευσης πυρκαγιάς μπορεί να τοποθετηθεί σε μη επανδρωμένα εναέρια οχήματα (UAV) για ανίχνευση πυρκαγιάς μεγάλης κλίμακας.

Υπήρξαν αρκετές μέθοδοι στις βιβλιογραφίες για ανίχνευση πυρκαγιάς. Χρωματικό μοντέλο, κίνηση, χωρικά και χρονικά χαρακτηριστικά χρησιμοποιούνται κυρίως επειδή η φωτιά έχει πολύ συγκεκριμένα χαρακτηριστικά σε σύγκριση με άλλα αντικείμενα. Σχεδόν όλες οι προτεινόμενες μέθοδοι ακολουθούν παρόμοια διαδικασία ανίχνευσης, δηλαδή βρίσκουν πρώτα κινούμενα pixel χρησιμοποιώντας αφαίρεση φόντου και στη συνέχεια εφαρμόζουν έγχρωμο μοντέλο για να βρουν περιοχές χρωμάτων πυρκαγιάς. Αυτές οι περιοχές αναλύονται περαιτέρω χωρικά και χρονικά για να ανιχνευθούν τα ακανόνιστα και "τρεμάμενα" χαρακτηριστικά της φωτιάς. Δεδομένου ότι η κίνηση είναι το κυρίαρχο χαρακτηριστικό, οι μέθοδοι αυτές λειτουργούν μόνο με σταθερή κάμερα, δηλαδή σε σενάρια επιτήρησης. Σε αυτή την εργασία, η μέθοδος δεν περιορίζεται από τέτοιους περιορισμούς. Χρησιμοποιείται η σύγχρονη ισχυρή μέθοδος βαθιάς εκμάθησης για να γίνουν γνωστές οι αναπαραστάσεις χαρακτηριστικών από τους ταξινομητές διάκρισης δεδομένων για ανίχνευση περιοχής πυρκαγιάς. Συγκεκριμένα, χρησιμοποιούνται νευρωνικά δίκτυα βαθιάς συνέλιξης (CNN) ως μηχανή εκμάθησης. Μία από τις συνεισφορές αυτής της εργασίας είναι ότι προτείνεται ένα benchmark πυρανίχνευσης. Από όσο γνωρίζουμε, δεν υπάρχει πρότυπο σύνολο δεδομένων και πρωτόκολλο αξιολόγησης στις προηγούμενες βιβλιογραφίες, πράγμα το οποίο καθιστά δύσκολη τη σύγκριση διαφόρων μεθόδων ανίχνευσης πυρκαγιάς.

Έγινε συλλογή από ένα σύνολο δεδομένων πυρανίχνευσης από διάφορους διαδικτυακούς πόρους. Τα δεδομένα βίντεο αποκτώνται κυρίως από το [36], το οποίο όμως είναι σχετικά παλιό dataset, όμως χρησιμοποιείται συνήθως στις πρόσφατες βιβλιογραφίες ανίχνευσης πυρκαγιάς. Το συγκεκριμένο σύνολο δεδομένων περιέχει 25 βίντεο που έχουν ληφθεί σε δασικό περιβάλλον, συμπεριλαμβανομένων 21 θετικών (φωτιάς) ακολουθιών και 4 αρνητικών (μη πυρκαγιών) ακολουθιών. Για να εκπαιδευτεί ο ταξινομητής πυρκαγιάς με βάση την εικόνα, εξάγονται περαιτέρω εικόνες από τα βίντεο με ρυθμό δειγματοληψίας 5, δηλαδή λαμβάνεται δείγμα μιας εικόνας κάθε πέντε καρέ. Αυτές οι εικόνες μετατρέπονται σε κανονικό μέγεθος 240×320 . Προκειμένου να υποστηριχτεί η αξιολόγηση του εντοπισμού των πλαισίων (patches) πυρκαγιάς, επισημαίνονται με μη αυτόματο τρόπο τα εξερχόμενα πλαίσια όπου υπάρχει πυρκαγιά με πλαίσια 32×32 . Ως απόδειξη της έννοιας, σε αυτή την εργασία, χρησιμοποιείται ένα μικρό υποσύνολο για την ανάπτυξη και την αξιολόγηση των αλγορίθμων. Τα στατιστικά στοιχεία των χρησιμοποιούμενων εικόνων για εκπαίδευση (train) και αξιολόγηση (test), αλλά και ο αριθμός των επισημασμένων patches παρουσιάζονται στον Πίνακα 4.1.

	Images	Patches	Positive	Negative
Train	178	12460	1307	11153
Test	59	4130	539	3591

Πίνακας 4.1: Περιγραφή του dataset των Zhang et al.

Στο παρακάτω Σχήμα 4.1 παρουσιάζεται ένα δείγμα της εξόδου του συστήματος που περιγράφεται σε αυτή την εργασία.



Σχήμα 4.1: Παραδείγματα ανίχνευσης φωτιάς σε patches

Οι περισσότερες από τις βιβλιογραφίες αναφέρουν την αξιολόγηση επιπέδου εικόνας (image-level evaluation), δηλαδή την ταξινόμηση της πυρκαγιάς και της μη πυρκαγιάς. Μερικές από αυτές αναφέρουν επίσης την ακρίβεια ανίχνευσης επιπέδου πλαισίου (patch-level detection accuracy). Με βάση τους σχολιασμούς σε επίπεδο πλαισίου, θεωρούμε κριτήριο αξιολόγησης και των δύο την ακρίβεια της ταξινόμησης σε επίπεδο εικόνας και την επιδιόρθωση.

Μετρικές Αξιολόγησης

Χρησιμοποιείται γενικά το ακόλουθο κριτήριο για την αξιολόγηση του συστήματος ανίχνευσης πυρκαγιάς:

$$accuracy = \frac{TP + TN}{N_{pos} + N_{neg}} \quad (4.1)$$

$$detection\ rate = \frac{TP}{N_{pos}} \quad (4.2)$$

$$false\ alarm\ rate = \frac{FN}{N_{neg}} \quad (4.3)$$

όπου το TP είναι ο αριθμός των true positives στοιχείων, δηλαδή ο αριθμός των πλαισίων πυρκαγιάς που ταξινομούνται ως πυρκαγιά, TN είναι ο αριθμός των true negatives, δηλαδή ο αριθμός των πλαισίων που δεν περιέχουν φωτιά και που ταξινομούνται ως μη πυρκαγιά, ο FN είναι ο αριθμός των false negatives, δηλαδή ο αριθμός των πλαισίων που δεν περιέχουν φωτιά και ταξινομούνται ως φωτιά. N_{pos} και N_{neg} είναι ο αριθμός των πραγματικών πλαισίων που περιλαμβάνουν και δεν περιλαμβάνουν φωτιά αντίστοιχα.

Προτεινόμενη Μέθοδος

Στόχος είναι η ανίχνευση πυρκαγιάς σε επίπεδο πλαισίου. Χρησιμοποιώντας τις μεθόδους μηχανικής μάθησης, ο απλός τρόπος είναι να εκπαιδευτεί ένας ταξινομητής

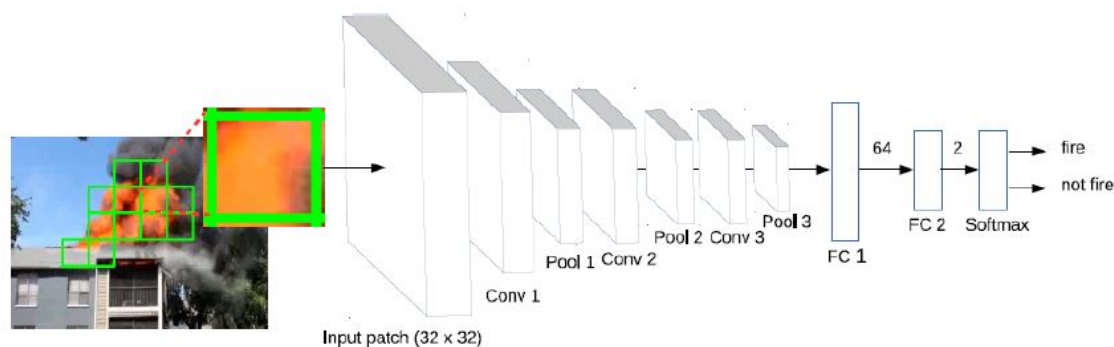
πυρκαγιάς χρησιμοποιώντας επισημασμένα θετικά και αρνητικά δείγματα, δηλαδή ένα δυαδικό μοντέλο ταξινόμησης. Μετά από αυτή τη βασική ιδέα, έχουμε δύο προτάσεις,

- να εκπαιδύσουμε έναν δυαδικό ταξινομητή χρησιμοποιώντας δικές μας δεδομένα, τα οποία πρώτα τα έχουμε επισημάνει (learning from scratch), ή
- πρώτα να γίνει μάθηση ενός ταξινομητή πυρκαγιάς πλήρους εικόνας και στη συνέχεια εφαρμογή ενός ταξινομητή πλαισίου εάν η εικόνα ταξινομηθεί ότι περιέχει φωτιά (cascade CNN fire detector),

Λεπτομέρειες σχετικά με τις δύο αυτές προτάσεις παρουσιάζονται στις παρακάτω ενότητες.

Learning fire patch classifier from scratch

Η εκμάθηση ενός δυαδικού ταξινομητή χρησιμοποιώντας τα επισημασμένα (annotated) patches είναι απλή. Ωστόσο, η εκπαίδευση αυτή είναι σημαντική επειδή τα θετικά και τα αρνητικά δείγματα είναι ασύμμετρα, καθώς υπάρχει μεγάλη ποσότητα αρνητικών patches και μόνο λίγα θετικά πλαίσια (βλ. Πίνακα 4.1). Σε αυτή την εργασία, έγινε διερεύνηση δύο τύπων ταξινομητών: γραμμικού και μη γραμμικού ταξινομητή. Για το γραμμικό ταξινομητή, χρησιμοποιείται γραμμικό SVM και για τον μη γραμμικό ταξινομητή, χρησιμοποιείται CNN από το Caffe work-frame [37]. Λόγω του μικρού μεγέθους του annotated συνόλου δεδομένων, υιοθετήθηκε το CIFAR 10 δίκτυο. Η δομή του δικτύου CIFAR 10 (Σχήμα 4.2) διατηρήθηκε ως έχει, έγινε όμως αλλαγή ως προς τον αριθμό εξόδου στο τελευταίο πλήρως συνδεδεμένο στρώμα σε 2 για το δυαδικό πρόβλημα ταξινόμησης. Σε αυτό μικρό δίκτυο, συμπεριλήφθηκε επίσης ένα στρώμα απόρριψης (drop-out layer) για να μειωθεί η υπερφόρτωση (overfitting).



Σχήμα 4.2: Δομή του δικτύου για patches

Cascade CNN fire detector

Ένα μειονέκτημα του παραπάνω ταξινομητή πατση είναι ότι εκπαιδύεται με τοπικά patches τα οποία δεν περιέχουν γενικές πληροφορίες. Επίσης, ο υπολογισμός αυτών των ταξινομητών patch είναι δύσκολος λόγω του ότι πρέπει να γίνει δοκιμή πολλαπλών patches για κάθε πλαίσιο. Για να αντιμετωπιστεί το πρόβλημα της έλλειψης γενικών πληροφοριών, μπορούμε να εκπαιδύσουμε έναν ταξινομητή πλήρους εικόνας με βάση τις θετικές και αρνητικές εικόνες. Μια εικόνα ονομάζεται αρνητική εικόνα εάν δεν

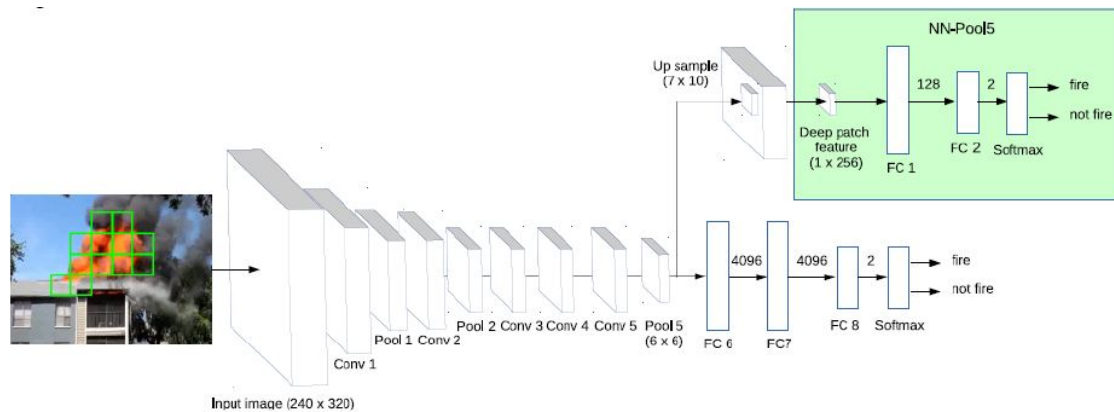
υπάρχει ίχνος πυρκαγιάς σε αυτήν και αντίστοιχα, εάν η εικόνα είναι θετική, πρέπει να περιέχει τουλάχιστον ένα κομμάτι πυρκαγιάς. Δεδομένου ότι έχουμε ένα πολύ μικρό σύνολο δεδομένων εκμάθησης, η εκπαίδευση ενός ταξινομητή πλήρους εικόνας από το μηδέν μπορεί να υποφέρει από overfitting. Ένας αποτελεσματικός τρόπος για να εκπαιδευτεί ένας ταξινομητής CNN σε μικρό σύνολο δεδομένων είναι η λεγόμενη τεχνική fine-tuning, η οποία μπορεί να μεταφέρει τις ήδη γνωστές παραμέτρους του δικτύου στο νέο μοντέλο για το νέο σύνολο δεδομένων. Χρησιμοποιούνται τα 8 επίπεδα του AlexNet που είναι διαθέσιμο στο Caffe ως προ-εκπαιδευμένο μοντέλο του αλγορίθμου.

Μόλις εκπαιδευτεί ο ταξινομητής πλήρους εικόνας, μπορεί να γίνει ταξινόμηση για το αν μια εικόνα περιέχει σημεία πυρκαγιάς ή όχι. Αν υπάρχουν, εφαρμόζεται ένας ταξινομητής "πυροπροστασίας" για να βρεθεί η ακριβής τοποθεσία της εστίας. Ένας τρόπος είναι να εφαρμόσουμε άμεσα τον ταξινομητή patch. Ωστόσο, η απλή εφαρμογή του ταξινομητή αυτού εκπαιδευμένο από το μηδέν δεν είναι βέλτιστη. Ο λόγος είναι ότι, πρώτον, η καθολικοί (πλήρης εικόνα) και οι τοπικοί (patch) ταξινομητές είναι ανεξάρτητοι, δεν υπάρχει αλληλεπίδραση ή ανταλλαγή πληροφοριών μεταξύ τους. Δεύτερον, ο υπολογισμός είναι το άθροισμα τόσο των ταξινομητών πλήρους εικόνας όσο και των τοπικών ταξινομητών, πράγμα το οποίο είναι περιττό. Μία καλύτερη λύση είναι να μοιραστούν οι δυνατότητες των δύο ταξινομητών.

Οι συγγραφείς της παρούσας εργασίας προτείνουν να γίνει εκπαίδευση του τοπικού ταξινομητή απευθείας από το πλήρης εικόνας CNN, δηλαδή να γίνει εξαγωγή των τοπικών χαρακτηριστικών από το CNN. Στην πράξη, εξάγουμε τοπικά χαρακτηριστικά από το στρώμα Pool-5. Τα χαρακτηριστικά από το στρώμα Pool-5 έχουν διάσταση $6 \times 6 \times 256$, το οποίο μπορεί να θεωρηθεί ως ένα πλέγμα 6×6 με 256-διάστατα χαρακτηριστικά για κάθε πλέγμα. Σημειώστε ότι το μέγεθος του patch είναι 32×32 , πράγμα που σημαίνει ότι έχουμε 7×10 πλέγματα για την εικόνα μεγέθους 240×320 . Επομένως, τα χαρακτηριστικά του τμήματος Pool-5 δεν ταιριάζουν με το μέγεθος του πλέγματος. Υπάρχουν αρκετές στρατηγικές για την αντιμετώπιση αυτού του προβλήματος, π.χ. μετατροπή του πλήρως συνδεδεμένου στρώματος σε στρώμα συνέλιξης [38]. Ωστόσο, βρέθηκε ότι μια απλή γραμμική παρεμβολή λειτουργεί καλά. Επομένως, απλά προστέθηκε ένα δείγμα 7×10 πλέγματος από τα χαρακτηριστικά 6×6 του Pool-5. Με αυτόν τον τρόπο αποκτήθηκαν 256-διάστατα χαρακτηριστικά για κάθε patch. Με αυτά τα εξαγόμενα βαθιά χαρακτηριστικά, είμαστε σε θέση να εκπαιδευσουμε έναν ταξινομητή patch. Στην πράξη, βρήκαμε ότι ένα μη γραμμικό μικρό νευρωνικό δίκτυο λειτουργεί καλύτερα από το γραμμικό SVM. Η δομή αυτού του μικρού μη γραμμικού δικτύου που προστέθηκε απεικονίζεται στο Σχήμα 4.3.

Το πλεονέκτημα αυτής της μεθόδου είναι ότι:

1. είναι γνωστό το γεγονός ότι οι αναπαραστάσεις χαρακτηριστικών υψηλού επιπέδου από το βαθύ δίκτυο είναι πιο διακριτικές από τα πρωτογενή χαρακτηριστικά από pixel. Έτσι, μπορούμε να εκπαιδευσουμε έναν ταξινομητή patch με μεγαλύτερη ακρίβεια.
2. δεν υπάρχει υπολογισμός πλεονάζοντος χαρακτηριστικού δεδομένου ότι τόσο οι καθολικοί όσο και οι τοπικοί ταξινομητές μοιράζονται το ίδιο δίκτυο για υπολογιστικές λειτουργίες.
3. ο υπολογισμός του βαθμιδωτού ταξινομητή είναι ακόμα χαμηλότερος λόγω της χαμηλής διάστασης των χαρακτηριστικών υψηλού επιπέδου στο CNN.



Σχήμα 4.3: Η συνολική δομή του δικτύου: Ο τοπικός ταξινομητής εκπαιδεύεται με τα χαρακτηριστικά που εξάγονται από το μικρό μη γραμμικό νευρωνικό δίκτυο (πράσινο πλαίσιο)

Για παράδειγμα, χρησιμοποιώντας την προέκταση Pool-5, η διάσταση για κάθε patch είναι 256 ενώ όταν χρησιμοποιούνται στοιχεία pixel, είναι $32 \times 32 \times 3$ δηλαδή 3072.

Τα αποτελέσματα είναι αρκετά ενθαρρυντικά, καθώς πετυχαίνεται accuracy της τάξης του **92%** και **93%** για SVM και CNN αντίστοιχα, ενώ χρησιμοποιώντας την επέκταση με το Pool-5 το νούμερο αυτό εκτοξεύεται στο **96%** και **97%**.

Το χρώμα και η υφή είναι τα πιο ευρέως χρησιμοποιούμενα χαρακτηριστικά για ανίχνευση και αναγνώριση αντικειμένων. Και τα δύο χαρακτηριστικά μπορούν να υπολογιστούν αποτελεσματικά, επομένως είναι κατάλληλα για συστήματα που δουλεύουν σε πραγματικό χρόνο. Σχεδόν όλες οι μέθοδοι ανίχνευσης πυρκαγιάς χρησιμοποιούν το χρώμα ως ξεχωριστό χαρακτηριστικό της φωτιάς, αλλά σε διαφορετικό χρωματικό χώρο. Υπάρχουν επίσης και άλλοι χρωματικοί χώροι όπως το YUV, το HSI και το LAB. Ωστόσο, δεν υπάρχει έρευνα για τη θεωρητική ανάλυση και την εμπειρική σύγκριση της επιλογής χρωμάτων.

Πολύ ενδιαφέρουσα είναι η εργασία των Jiang et al. [39], η οποία χρησιμοποιεί τέτοιου είδους τεχνικές, όπως θα φανούν και στη συνέχεια. Επιπλέον, αποτέλεσε σημείο αναφοράς στην υλοποίηση της βάσης του συστήματος της παρούσας διπλωματικής, το οποίο παρουσιάζεται στην Ενότητα 6.2. Πιο συγκεκριμένα, στη δουλειά αυτή ανακτήθηκαν καρέ από βίντεο και έγινε εξαγωγή εικόνων με το χέρι, οι οποίες αποτέλεσαν τα δείγματα για ανάλυση. Συλλέχθηκαν 726 δείγματα πυρκαγιάς που ποικίλλουν στο χρώμα, το σχήμα και την έντασή τους και 637 δείγματα μη πυρκαγιάς που περιλαμβάνουν μια ποικιλία αντικειμένων, συμπεριλαμβανομένων ανθρώπων και αυτοκινήτων, τα οποία είναι γεμάτα οπτικές λεπτομέρειες και απλά αντικείμενα όπως οροφές, τοίχοι και ουρανοί, αντικείμενα τα οποία έχουν απλό χρώμα και υφή. Επιλέχθηκαν τυχαία τα 4/5 από τα δείγματα για εκπαίδευση και 1/5 για αξιολόγηση.

Καθολικά Χαρακτηριστικά (Global Features)

Πειραματιζόμαστε με τέσσερις χώρους χρώματος, συμπεριλαμβανομένων των RGB, YUV, HSV και LAB. Το καθολικό χρωματικό ιστόγραμμα διαφορετικού χώρου χρώματος αποτελείται από 96 bins, που είναι η συνένωση τριών ιστογραμμάτων 32-bin για κάθε κανάλι. Όπως προκύπτει από το αποτέλεσμα, το ιστόγραμμα RGB δίνει τα

φτωχότερα αποτελέσματα, το YUV και το HSV είναι καλύτερα και το LAB παράγει το καλύτερο σε κάθε ρύθμιση πυρήνα SVM. Η υψηλότερη απόδοση επιτυγχάνεται με ιστόγραμμα LAB με πυρήνα RBF.

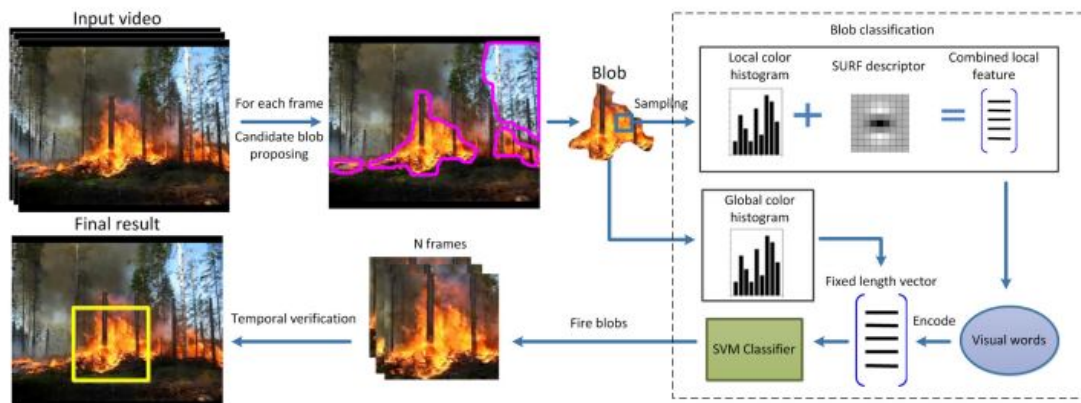
Τοπικά Χαρακτηριστικά (Local Features)

Η υφή είναι ένα ακόμη σημαντικό χαρακτηριστικό εκτός από το χρώμα. Πολλοί περιγραφητές της υφής όπως οι SIFT [40] και HOG [41] έχουν χρησιμοποιηθεί ευρέως και επιτυχώς σε διάφορα είδη αναγνώρισης αντικειμένων και σκηνών. Ωστόσο, η υψηλή υπολογιστική πολυπλοκότητα περιορίζει την εφαρμογή τους στην παρακολούθηση και την ανάκτηση δεδομένων. Για να διασφαλιστεί ότι το σύστημα μπορεί να λειτουργεί σε πραγματικό χρόνο, υιοθετείται ο περιγραφητής SURF [42]. Έχει παρόμοια απόδοση σε σύγκριση με το SIFT, αλλά πολύ πιο γρήγορος ως προς τον υπολογισμό. Επιπλέον, ο υπολογισμός SURF είναι παράλληλος, έτσι ώστε να μπορεί να επωφεληθεί από συστήματα με επεξεργαστές πολλαπλών πυρήνων και GPUs.

Οι περιγραφητές SURF υπολογίζονται πάνω από τα βασικά σημεία που ανιχνεύονται από τον Fast Hessian Detector, αποδίδοντας διανύσματα 64 διαστάσεων. Επειδή η ποιότητα και η ευκρίνεια των πλαισίων βίντεο είναι σχετικά χαμηλή, το κατώφλι του ανιχνευτή έχει οριστεί σε 100 για να ανιχνεύσει περισσότερα σημεία-κλειδιά. Τα τοπικά χρωματικά ιστογράμματα υπολογίζονται εντός του πλαισίου κάθε πυρήνα SURF. Το μέγεθος του είναι 24 bins, με 8 bins για κάθε κανάλι. Συνδυάζονται για να σχηματίσουν ένα διάνυσμα χαρακτηριστικών 88 διαστάσεων.

Όπως φαίνεται στο Σχήμα 4.4, το σύστημα ανίχνευσης πυρκαγιάς αποτελείται από τρία τμήματα:

1. υποψήφιες περιοχές πυρκαγιάς που προτείνονται με μοντέλο υποβάθρου/φόντου (candidate fire region proposing),
2. ταξινόμηση περιοχή πυρκαγιάς με χαρακτηριστικά χρώματος υφής και λεξικού οπτικών λέξεων (feature extraction and region classification), και
3. χρονική επαλήθευση (temporal verification)



Σχήμα 4.4: Το σύστημα ανίχνευσης φωτιάς των Jiang et al.

Θα αναλύσουμε λεπτομερώς κάθε μέρος στις ακόλουθες τρεις ενότητες αντίστοιχα.

Candidate fire region proposing

Τα περισσότερα συστήματα ανίχνευσης πυρκαγιάς εντοπίζουν περιοχές φωτιάς ή pixels άμεσα. Με βάση τις παρατηρήσεις και τα στατιστικά πειράματα, διαπιστώνουμε ότι οι περιοχές φωτιάς είναι σχετικά φωτεινές, δηλαδή οι εντάσεις των εικονοστοιχείων τους είναι πάνω από κάποια ορισμένη τιμή. Μπορούμε να χρησιμοποιήσουμε αυτή τη λειτουργία για να βρούμε υποψήφιες περιοχές πυρκαγιάς και να οικοδομήσουμε την περαιτέρω διαδικασία πυρανίχνευσης. Σύμφωνα με την έρευνά μας, μόνο ο Toreyin et al. [1] προσπάθησε να εξαλείψει το φόντο χωρίς φωτιά. Ωστόσο, χρησιμοποίησαν ένα σταθερό ευρετικό κατώφλι (threshold), το οποίο μπορεί να αποτύχει όταν το φόντο είναι φωτεινό ή όταν υπάρχουν αντανακλάσεις φωτιάς σε λευκό τοίχο. Για να αντιμετωπίσουμε τέτοιες δύσκολες περιστάσεις, υιοθετούμε ένα κατώφλι πολλαπλών επιπέδων που βασίζεται σε εμπειρική μελέτη. Το σύστημα μπορεί να ρυθμίσει αυτόματα το κατώτατο όριο ανάλογα με τις στατιστικές έντασης του φόντου. Η υψηλότερη ένταση φόντου οδηγεί σε χαμηλότερο κατώφλι και αντίστροφα. Για ένα βίντεο κλιπ παρακολούθησης, το οποίο συνήθως συλλαμβάνεται από μια σταθερή κάμερα, το φόντο θα είναι σχετικά στατικό. Εφαρμόστηκε υποβάθμιση φόντου για να μειωθούν περαιτέρω οι υποψήφιες περιοχές πυρκαγιάς.

Feature extraction and region classification

Στο στάδιο της αρχικοποίησης της ταξινόμησης, χρησιμοποιείται ένα βιβλίο κωδικών (codebook) για την κατασκευή ενός δέντρου $K - D$, το οποίο επιτρέπει την γρήγορη εύρεση και αναζήτηση χαρακτηριστικών. Για κάθε εισερχόμενη κηλίδα (blob), γίνεται πυκνή δειγματοληψία με LAB + SURF περιγραφητές από την περιοχή της με διάστημα 9 pixel και πυρήνες SURF μεγέθους 9×9 . Αυτό το βήμα παράγει N περιγραφείς.

Για κάθε περιγραφητή $d_j, j = 1, 2, \dots, N$, αναζητούμε την απόσταση $D(d_j, u_i)$ σε m πλησιέστερους γείτονες $u_i = 1, 2, \dots, m$ από το δέντρο $K - D$. Προκειμένου να αποφευχθούν τα μειονεκτήματα της ενιαίας αντιστοίχισης, το d_j μπορεί να συνδυαστεί με όλους τους πλησιέστερους γείτονες. Θεωρητικά, όσο μεγαλύτερο το m , τόσο καλύτερο το αποτέλεσμα ταξινόμησης, και τόσο υψηλότερο υπολογιστικό κόστος. Σύμφωνα με τα πειράματα, το $m = 10$ μπορεί να επιτύχει τον καλύτερο λόγο κόστους/απόδοσης. Το βάρος της αντιστοίχισης d_j με u_i υπολογίζεται ως εξής:

$$w_{ij} = \frac{K_\sigma(D(d_j, u_i))}{\sum_{k=1}^m K_\sigma(D(d_j, u_k))}, \quad i \in [1, m] \quad (4.4)$$

όπου το $K_\sigma(x)$ είναι η γκαουσιανή συνάρτηση πυρήνα:

$$K_\sigma(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp -\frac{1}{2} \frac{x^2}{\sigma^2} \quad (4.5)$$

Ενώ χρησιμοποιούμε αυτή τη μέθοδο αντιστοίχισης, υποθέτουμε ότι η απόσταση μεταξύ d_j και u_i ικανοποιεί τη γκαουσιανή κατανομή λόγω της επίδρασης της ομαδοποίησης (clustering). Αυτός ο πυρήνας μετατρέπει την μικρότερη απόσταση σε υψηλότερη πιθανότητα. Χρησιμοποιούμε ένα ιστογράμμο h των 500-bin για να κωδικοποιήσουμε όλες τις τοπικές λειτουργίες σε μια υποψήφια εστία πυρκαγιάς. Οι τιμές κάθε bin υπολογίζονται ως εξής:

$$h = \sum_{j=1}^N w_{ij}, \quad i \in [1, 500] \quad (4.6)$$

Μετά την κανονικοποίηση, κάθε bin αντιπροσωπεύει την πιθανότητα εμφάνισης του αντίστοιχου στοιχείου στο codebook και στο ιστογράμμο. Τέλος, υπολογίζεται το ιστογράμμο LAB με κανονικοποιημένο LAB 96-bin για το υποψήφιο σημείο φωτιάς. Ο συνδυασμός αμφοτέρων των ιστογραμμμάτων είναι η τελική αναπαράσταση της κηλίδας (blob) και χρησιμοποιείται ως είσοδος στον ταξινομητή SVM.

Temporal verification

Αντικείμενα όπως λαμπτήρες δρόμου και φώτα αυτοκινήτων έχουν μεγάλη ομοιότητα με την εμφάνιση φωτιάς και είναι δύσκολο να αποκλειστούν με βάση ένα μόνο πλαίσιο. Επομένως, οι περιοχές/κηλίδες που ταξινομούνται ως πυρκαγιά στο προηγούμενο βήμα θα πρέπει να επαληθεύονται περαιτέρω με στατιστική χρονική διακύμανση.

Η πυρκαγιά εμφανίζει κυρίως διακύμανση σχήματος αλλά όχι χρώματος ή υψής. Όταν εμφανίζεται ένα νέο blob, χρησιμοποιούμε τρεις απλές παραμέτρους: περίμετρο, εμβαδόν και χωρική κατανομή για να υπολογίσουμε τη σταθερότητα της περιοχής σε διαδοχικά 25 καρέ, χωρίς να χρειάζεται να υπολογίσουμε την ακριβή αναπαράστασή της, όπως οι περιγραφείς Fourier που χρησιμοποιούνται στο [7]. Η περίμετρος και το εμβαδόν είναι πολύ διαισθητικές, και τα μ_p , μ_a , σ_p , σ_a δηλώνουν τη μέση και τυπική απόκλιση αντίστοιχα. Για να υπολογίσουμε τη χωρική κατανομή, διαιρούμε το ορθογώνιο οριοθέτησης ενός μπλοκ σε 4 ίσα υπο-ορθογώνια και υπολογίζουμε τον αριθμό των pixel μέσα σε κάθε υπο-ορθογώνιο που έχουν επισημανθεί ως "1" στη δυαδική μάσκα, δηλαδή d_1 , d_2 , d_3 , d_4 . Η τυπική απόκλιση της χωρικής κατανομής ορίζεται ως εξής:

$$\sigma_d = \sigma_{d_1} + \sigma_{d_2} + \sigma_{d_3} + \sigma_{d_4} \quad (4.7)$$

Ορίζουμε ότι ένα blob είναι σταθερό εάν οι παράμετροι πληρούν την ακόλουθη προϋπόθεση:

$$\sigma_p < t_1 \mu_p \quad \text{and} \quad \sigma_a < t_1 \mu_a \quad \text{and} \quad \sigma_d < t_1 \mu_a \quad (4.8)$$

και ένα blob είναι ασταθές αν ισχύει η ακόλουθη συνθήκη:

$$\sigma_p > t_2\mu_p \quad \text{and} \quad \sigma_a < t_2\mu_a \quad \text{and} \quad \sigma_d > t_2\mu_d \quad (4.9)$$

όπου t_1 και t_2 είναι κατώτατα όρια βάσει περιβαλλοντικών ρυθμίσεων. Η παρακολούθηση σε εσωτερικό χώρο μπορεί να χρησιμοποιεί χαμηλότερες τιμές. Σε υπαίθρια σενάρια, όπου η φωτιά μπορεί να επηρεαστεί σε μεγάλο βαθμό από τη ροή του αέρα, τα όρια είναι υψηλότερα. Ο ψευδής ρυθμός συναγερμού (false alarm rate) μπορεί να μειωθεί περαιτέρω από αυτές τις δύο συνθήκες, δηλαδή η ικανοποίηση της Εξίσωσης (4.8) και (4.9) δεν θα πρέπει να προταθεί ως πυρκαγιά. Για παράδειγμα, εάν χρησιμοποιήσουμε την παρακολούθηση κίνησης και διαπιστώνουμε ότι ένα νεοδημιουργηθέν blob εξαφανίζεται από την αρχική του θέση μετά από αρκετά πλαίσια. Μπορούμε να ισχυριστούμε ότι δεν είναι φωτιά επειδή η φωτιά δεν κινείται από μόνη της. Μπορεί να υπάρχουν άλλες προηγούμενες γνώσεις τις οποίες μπορούμε να εκμεταλλευτούμε, αλλά είναι συγκεκριμένες για την εφαρμογή και εκτός του πλαισίου αυτής της εργασίας. Καταλήγοντας, με τη μέθοδο αυτή επιτυγχάνεται accuracy της τάξης του **92%**.

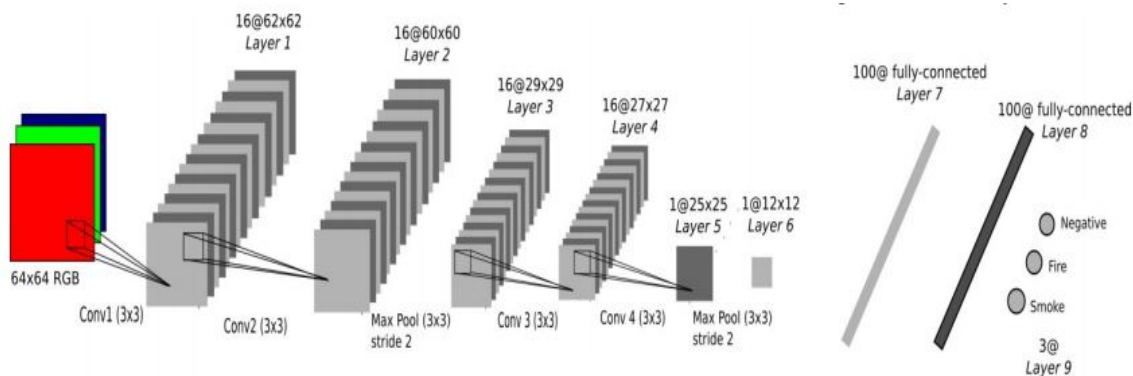
4.1.2 Τεχνικές με Βαθιά Μάθηση

Όσον αφορά τη βιβλιογραφία, όπως φάνηκε και παραπάνω, στόχος είναι να υλοποιηθεί ένας αλγόριθμος βασισμένος σε κανόνες ή πολυδιάστατο διάνυσμα χαρακτηριστικών που χρησιμοποιείται ως είσοδος σε ένα συμβατικό αλγόριθμο για ταξινόμηση: SVM, νευρωνικό δίκτυο, Adaboost κ.λπ. Επομένως, οι συμβατικές μέθοδοι ανίχνευσης πυρκαγιάς με βίντεο αντιμετωπίζουν το πρόβλημα βασιζόμενο στην εμπειρογνωμοσύνη για την οικοδόμηση σχετικών εξαγωγέων χαρακτηριστικών (feature extractors). Απαιτούνται εμπειρογνώμονες για τη δημιουργία μοντέλων βασισμένων σε κανόνες και διακριτικών χαρακτηριστικών. Μια διαφορετική προσέγγιση για αυτό το πρόβλημα είναι να χρησιμοποιηθεί ένας αλγόριθμος εκμάθησης που να εξάγει τα χρήσιμα χαρακτηριστικά αντί να χρησιμοποιηθεί ένα σύστημα να τα δημιουργήσει. Οι βαθιές τεχνικές εκμάθησης μπορούν να μάθουν τέτοια χρήσιμα χαρακτηριστικά για την ανίχνευση φωτιάς και καπνού στο βίντεο. Τα συνελκτικα νευρωνικά δίκτυα είναι μια παραλλαγή της βαθιάς μάθησης που μπορεί να εξαγάγει τοπολογικές ιδιότητες από μια εικόνα.

Σε αντίθεση, η προσέγγιση του [43] είναι εννοιολογικά πολύ πιο απλή. Χρησιμοποιείται ένα CNN ως ανιχνευτής πλαισίου πυρκαγιάς/καπνού. Το CNN λειτουργεί απευθείας σε ακατέργαστο πλαίσιο RGB χωρίς την ανάγκη της βαθμίδας εξαγωγής χαρακτηριστικών. Το CNN μαθαίνει αυτόματα ένα σύνολο οπτικών χαρακτηριστικών από την εκπαίδευση δεδομένων.

Η αρχιτεκτονική ταξινόμησης είναι κλασική για το συνελκτικό νευρωνικό δίκτυο, συνδυάζοντας συνέλιξη (convolution) και Max Pooling. Ωστόσο, για να επιτευχθεί μια γρήγορη ταξινόμηση, επιλέγεται ένα μικρό δίκτυο. Στο Σχήμα 4.5 φαίνεται η αρχιτεκτονική του δικτύου που χρησιμοποιείται, το οποίο αποτελείται από 9 στρώματα (layers).

Μια έγχρωμη εικόνα RGB περνά διαδοχικά δύο συνελκτικές διαδικασίες με πυρήνα μεγέθους 3×3 . Η ίδια δομή εφαρμόζεται μετά το τρίτο επίπεδο. Ένα max pooling 3×3



Σχήμα 4.5: Αρχιτεκτονική του δικτύου: Το πρώτο μέρος αποτελείται από 6 στρώματα: 1 έως 6 (στρώματα συνένωσης και max pooling), το δεύτερο μέρος είναι τα πλήρως συνδεδεμένα στρώματα

με βήμα (stride) 2 ακολουθεί το στρώμα συνέλιξης δύο και πέντε. Τα στρώματα από ένα έως τέσσερα έχουν 16 χάρτες χαρακτηριστικών. Τα στρώματα πέντε και έξι έχουν μόνο ένα χάρτη χαρακτηριστικών. Τα στρώματα επτά και οκτώ είναι πλήρως συνδεδεμένα. Η έξοδος του τελευταίου πλήρως συνδεδεμένου στρώματος τροφοδοτείται σε έναν Softmax 3-δρόμων που παράγει την ταξινόμηση σε 3 ετικέτες κλάσης, δηλαδή σε 3 διαφορετικές κατηγορίες. Παρόμοια με τα [44] και [45] επιλέγεται για τις συνελικτικές και τις πλήρως συνδεδεμένες στρώσεις μια λειτουργία ενεργοποίησης Leaky ReLU με συντελεστή $a = 1/3$. Ο στόχος της ταξινόμησης είναι να αποφασίσουμε αν μια εικόνα περιέχει φωτιά ή/και καπνό. Για την επίλυση αυτού του προβλήματος, ο ταξινομητής εκπαιδεύεται από μια συλλογή επισημασμένων εικόνων. Επιπλέον, θέλουμε να εντοπίσουμε τη θέση της φωτιάς και του καπνού σε μια εικόνα ή σε ένα βίντεο. Το σύνολο εκπαίδευσης αποτελείται από 27919 δείγματα εικόνων RGB μεγέθους 64×64 pixel, 8915 για καπνό, 7257 για πυρκαγιά και 11752 αρνητικά (δεν υπάρχει φωτιά ή καπνός). Δημιουργούμε 3 υποσύνολα: 60% των εικόνων για εκπαίδευση, 20% για επικύρωση (validation) και 20% για αξιολόγηση. Τα δεδομένα εκπαίδευσης έχουν πραγματοποιηθεί με έναν υπολογιστή που αποτελείται από ένα μικροεπεξεργαστή Intel Xeon (CPU συχνότητας 3.1Ghz, RAM 16Go) και μια κάρτα γραφικών GTX 980 Ti (2816 πυρήνες, 6 GB μνήμες). Χρησιμοποιήθηκε στοχαστική κλίση (SGD) με mini-batches μεγέθους 100. Τα βάρη στο δίκτυο αρχικοποιούνται τυχαία. Ο αρχικός ρυθμός εκμάθησης (learning rate) είναι 0.01 και ορμή (momentum) 0.9. Ο ρυθμός εκμάθησης μειώνεται κατά ένα συντελεστή 0.95 κάθε 5 εποχές. Αντιθέτως, η ορμή αυξάνεται για να φθάσει στο 0.9999. Για να επιτευχθεί η καλύτερη ακρίβεια για αυτές τις παραμέτρους, έγιναν αρκετές δοκιμές. Το CNN υλοποιήθηκε με Theano [46], [47] και Lasagne. Χρησιμοποιήθηκε drop-out 0.5 στα δύο πλήρως συνδεδεμένα στρώματα για να αποφευχθεί περίπτωση overfitting. Το δίκτυο εκπαιδεύτηκε για περίπου 100 εποχές.

Αποτελέσματα

Η ακρίβεια ταξινόμησης στο σετ αξιολόγησης είναι της τάξης του **97.9%**. Το τέστ σύνολο αποτελείται από 1427 εικόνες πυρκαγιάς, 1758 εικόνες καπνού και 2399 αρνητικές εικόνες, σύνολο 5584 εικόνες. Μπορούμε να συμπεράνουμε ότι οι παράμετροι του μοντέλου μας CNN επιτρέπουν την καλή διάκριση ανάμεσα στη φωτιά και τον καπνό.

Βίντεο

Στόχος μας είναι να εντοπίσουμε την εστία της φωτιάς ή να εντοπίσουμε πιθανή πυρκαγιά σε ένα βίντεο. Ο χρόνος επεξεργασίας για ανίχνευση αποτελεί βασικό παράγοντα για την ακρίβεια. Επομένως, αποφασίζουμε να χρησιμοποιήσουμε την "ελαφριά δομή" που περιγράφεται στο Σχήμα 4.5. Στην πραγματικότητα, οι συνήθεις μέθοδοι χρησιμοποιούν κυλιόμενα παράθυρα για να ανιχνεύουν και να ταξινομούν το αντικείμενο σε πρωτότυπες ή αναμορφωμένες εικόνες. Αυτά τα παράθυρα περνούν από το συνελικτικό νευρωνικό δίκτυο και τα πλήρως συνδεδεμένα στρώματα και στη συνέχεια πρέπει να ταξινομηθούν. Για να γίνει ανάλυση ολόκληρης της εικόνας ενός πλαισίου βίντεο, η θέση του παραθύρου πρέπει να αλλάξει και να εισαχθεί και πάλι στο συνελικτικό νευρωνικό δίκτυο. Η προσέγγιση στην παρούσα εργασία είναι αρκετά διαφορετική, αντί να ολισθαίνει ένα παράθυρο των 64×64 εικονοστοιχείων στην εικόνα για να εντοπίσει τη φωτιά και τον καπνό, οι Frizzi et al. αποφάσισαν να εργαστούν μόνο για τον τελευταίο χάρτη χαρακτηριστικών. Το δίκτυο διαχωρίζεται σε δύο μέρη. Το πρώτο μέρος αποτελείται από 6 στρώματα: 1 έως 6 (συνελικτικά και Max Pooling στρώματα), και στο δεύτερο μέρος είναι τα πλήρως συνδεδεμένα στρώματα. Χρησιμοποιώντας το πρώτο μέρος του δικτύου, αξιολογείται ο τελευταίος χάρτης χαρακτηριστικών (επίπεδο 6) ολόκληρης της εικόνας. Γνωρίζουμε ότι με βάση τη δομή δικτύου του CNN, ένα συρόμενο παράθυρο με μέγεθος 64×64 pixels στην εικόνα RGB αντιστοιχεί σε ένα μέγεθος παραθύρου 12×12 pixels στον τελευταίο χάρτη χαρακτηριστικών.

Για να ανιχνεύσουμε τη φωτιά και τον καπνό σε ένα βίντεο πλαισίου, εφαρμόζουμε ένα συρόμενο παράθυρο μεγέθους 12×12 pixels στον τελευταίο χάρτη χαρακτηριστικών. Για να επιταχύνουμε την πρόβλεψη για κάθε παράθυρο 12×12 , κατασκευάζουμε έναν τανυστή $12 \times 12 \times 1 \times N$ (N : αριθμός παραθύρων) από τον τελευταίο χάρτη χαρακτηριστικών και χρησιμοποιούμε τη GPU της κάρτας γραφικών. Με αυτή τη μέθοδο, φαίνεται η ακρίβεια να παραμένει αμετάβλητη και η ταχύτητα ανίχνευσης και πρόβλεψης να αυξάνεται ανάλογα με την εικόνα του αρχικού μεγέθους και τον προβλεπόμενο αριθμό παραθύρων. Τέλος, δείχθηκε ότι κατά τη διάρκεια της δοκιμής ανίχνευσης, η σάρωση απευθείας από τον χάρτη χαρακτηριστικών αντί της σάρωσης του πλήρους αρχικού πλαισίου θα μπορούσε να μειώσει το κόστος χρόνου σε αναλογία από 6 έως 60.

Προκλήσεις

Οι τεχνικές ανίχνευσης πυρκαγιών που βασίζονται στο βίντεο έχουν διερευνηθεί ευρέως κατά τη διάρκεια της τελευταίας δεκαετίας. Η κύρια πρόκληση στην ανίχνευση φωτιάς με βάση το βίντεο έγκειται στη μοντελοποίηση της χαοτικής και πολύπλοκης φύσης του φαινομένου πυρκαγιάς και των μεγάλων διακυμάνσεων της εμφάνισης της φλόγας στο βίντεο. Για να αντιμετωπιστεί αυτό το πρόβλημα πολλοί ερευνητές παίρνουν τα χαρακτηριστικά κίνησης της φλόγας καθώς και τη χωρική κατανομή των χρωμάτων φωτιάς στη σκηνή ή προσπαθούν να συνδυάσουν τόσο χρονικά όσο και χωρικά χαρακτηριστικά. Ωστόσο, πολλά φυσικά αντικείμενα έχουν παρόμοια συμπεριφορά με τη φωτιά, π.χ. τον ήλιο, διάφορα τεχνητά φώτα ή αντανάκλασεις φωτός σε διάφορες επιφάνειες, σωματίδια σκόνης, τα οποία συχνά μπορούν να ανιχνευθούν λανθασμένα ως φλόγες. Επιπλέον, η πολυπλοκότητα της σκηνής και η χαμηλή ποιότητα βίντεο επηρεάζουν επίσης την ευρωστία των αλγόριθμων ανίχνευσης φλόγας που βασίζονται σε όραση υπολογιστών, αυξάνοντας έτσι το ποσοστό ψευδούς

συναγερμού.

Στο παρελθόν έχουν προταθεί πολλές καινοτόμες τεχνικές για την κατασκευή ενός ισχυρού συστήματος ανίχνευσης πυρκαγιάς το οποίο βασίζεται σε μεγάλο βαθμό στις τεχνικές επεξεργασίας εικόνες από υπολογιστή. Οι σύγχρονες τεχνικές βασισμένες σε όραση υπολογιστών για την ανίχνευση πυρκαγιάς και καπνού έχουν αξιολογηθεί και συγκριθεί συνολικά στο [48]. Η τεχνική ανάλυσης χρώματος έχει χρησιμοποιηθεί ευρέως στη βιβλιογραφία για την ανίχνευση και ανάλυση της πυρκαγιάς σε εικόνες και βίντεο [49], [50], [51], [52]. Στην κορυφή της ανάλυσης χρωμάτων χρησιμοποιήθηκαν πολλές καινοτόμες μέθοδοι για την εξαγωγή χαρακτηριστικών υψηλού επιπέδου από εικόνες πυρκαγιάς όπως ανάλυση υψής [49], δυναμική χρονική ανάλυση με φιλτράρισμα σε επίπεδο pixel και χωρική ανάλυση με αποσύνθεση και επισήμανση αντικειμένων [53] και ακανόνιστη ανίχνευση σχήματος πυρκαγιάς με μετασχηματισμό wavelet [52] κλπ. Αυτές οι τεχνικές δίνουν επαρκή απόδοση δεν υπερέρχουν των τεχνικών μηχανικής μάθησης. Μια συγκριτική ανάλυση μεταξύ μοντέλων βασισμένων σε χρώμα για την εξαγωγή κανόνων και ενός αλγορίθμου μηχανικής μάθησης γίνεται για το πρόβλημα της ανίχνευσης πυρκαγιάς στο [54]. Η τεχνική μηχανικής μάθησης που χρησιμοποιείται εδώ είναι η λογιστική παλινδρόμηση που είναι μία από τις απλούστερες τεχνικές στη μηχανική μάθηση (βλ. Ενότητα 3.5.2) και εξακολουθεί να υπερνικά τους αλγόριθμους με βάση το χρώμα σε όλα σχεδόν τα σενάρια. Αυτά τα σενάρια αποτελούνται από εικόνες που περιέχουν διαφορετικά χρώματα pixel φωτιάς διαφορετικών εντάσεων, με ή χωρίς καπνό.

Αντί να χρησιμοποιούμε πολλούς διαφορετικούς αλγόριθμους τον έναν πάνω στον άλλον για να εξαγάγουμε τα σχετικά χαρακτηριστικά, μπορούμε να χρησιμοποιήσουμε ένα δίκτυο που μαθαίνει τα χαρακτηριστικά αυτά από μόνο του. Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί με επιτυχία σε πολλούς διαφορετικούς τομείς όπως η επεξεργασία φυσικής γλώσσας, η αναγνώριση ομιλίας, η ανάλυση κειμένου και ιδιαίτερα η ταξινόμηση εικόνων. Η εξαγωγή συναφών χαρακτηριστικών από τις εικόνες είναι το κλειδί για την ακριβή ταξινόμηση και ανάλυση και γι'αυτό το πρόβλημα της ανίχνευσης πυρκαγιάς είναι ιδανικά κατάλληλο για βαθιά εκμάθηση. Τα Βαθιά Νευρωνικά Δίκτυα (Deep Neural Networks - DNNs) χρησιμοποιούνται για την αυτόματη "εκμάθηση" ιεραρχίας των σχετικών χαρακτηριστικών από τα δεδομένα χωρίς ανθρώπινη παρέμβαση και ο τύπος του νευρωνικού δικτύου ιδανικά για την ταξινόμηση των εικόνων είναι τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks - CNNs).

Έχουν υπάρξει πολλές σημαντικές συνεισφορές από διάφορους ερευνητές για την ανάπτυξη ενός συστήματος που να ανιχνεύει με ακρίβεια τη φωτιά στο περιβάλλον. Όμως, η πιο αξιοσημείωτη έρευνα στον τομέα αυτό περιλαμβάνει Deep Convolutional Neural Networks (DCNN). Τα μοντέλα DCNN αποτελούν σήμερα τα πιο επιτυχημένα μοντέλα ταξινόμησης εικόνων τα οποία τα καθιστούν ιδανικά για μια εργασία όπως η ανίχνευση πυρκαγιάς στις εικόνες. Αυτό έχει αποδειχθεί από προηγούμενες έρευνες που δημοσιεύθηκαν στον τομέα αυτό. Στην εργασία που περιγράφηκε προηγουμένως, οι συγγραφείς χρησιμοποιούν CNN για ανίχνευση πυρκαγιάς και καπνού σε βίντεο. Μια απλή διαδοχική αρχιτεκτονική του CNN, παρόμοια με τη LeNet-5 [24], χρησιμοποιήθηκε για ταξινόμηση. Οι συγγραφείς αναφέρουν μια ακρίβεια στο **97.9%** με ικανοποιητικό false alarm rate.

Ενώ στο [35] χρησιμοποιείται μια πολύ καινοτόμος τεχνική CNN για την ανίχνευση πυρκαγιάς σε μια εικόνα. Το cascaded CNN αποτελείται από την αρχιτεκτονική

AlexNet CNN με προ-εκπαιδευμένα βάρη ImageNet [55] και ένα άλλο μικρό δίκτυο μετά το τελικό pooling layer, το οποίο εξάγει χαρακτηριστικά patch και επισημαίνει τα κομμάτια που περιέχουν φωτιά. Η αρχιτεκτονική AlexNet χρησιμοποιείται επίσης στο [56] που υλοποιείται για την ανίχνευση του καπνού στις εικόνες. Εκπαιδύεται σε ένα αρκετά μεγάλο σύνολο δεδομένων που περιέχει εικόνες καπνού και μη καπνού για πολύ μεγάλο χρονικό διάστημα. Τα αναφερόμενα σκορ για τα μεγάλα και μικρά σύνολα δεδομένων είναι **96.88%** και **99.4%** αντίστοιχα, με σχετικά χαμηλά ποσοστά false alarm.

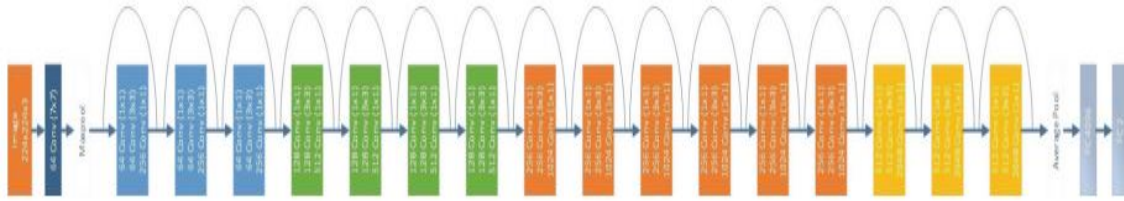
Σε αντίθεση με τα CNN που εξάγουν χαρακτηριστικά απευθείας από τις εικόνες, σε μερικές μεθόδους τα χαρακτηριστικά εικόνας/βίντεο εξάγονται χρησιμοποιώντας τεχνικές επεξεργασίας εικόνας και στη συνέχεια δίνονται ως είσοδος σε ένα νευρωνικό δίκτυο. Μια τέτοια προσέγγιση χρησιμοποιήθηκε στο [57]. Οι περιοχές φωτιάς από τα πλαίσια βίντεο λαμβάνονται με τιμές κατωφλίου στο χώρο χρώματος HSV. Τα γενικά χαρακτηριστικά της πυρκαγιάς υπολογίζονται χρησιμοποιώντας αυτές τις τιμές από πέντε συνεχή πλαίσια και η μέση και τυπική απόκλιση τους δίδεται ως είσοδος σε ένα νευρωνικό δίκτυο το οποίο εκπαιδύεται χρησιμοποιώντας αλγόριθμο οπισθοδιάδοσης (backpropagation) για να εντοπίσει περιοχές δασικής πυρκαγιάς. Αυτή η μέθοδος βασίζεται στην κατάτμηση των εικόνων με μεγάλη ακρίβεια και τα αποτελέσματα δείχνουν υψηλή ακρίβεια. Στο [58] χρησιμοποιείται ένα νευρωνικό δίκτυο για την εξαγωγή χαρακτηριστικών πυρκαγιάς με βάση το μοντέλο χρώματος HSI, το οποίο δίνει την περιοχή φωτιάς στην εικόνα ως έξοδο. Το επόμενο βήμα είναι η τμηματοποίηση της περιοχής πυρκαγιάς, όπου οι περιοχές πυρκαγιάς είναι κατά προσέγγιση τμηματοποιημένες και οι ψευδείς περιοχές φωτιάς όπως σκιά πυρκαγιάς και αντικείμενα που θυμίζουν φωτιά αφαιρούνται από τη διαφορά εικόνας. Μετά από αυτό, η μεταβολή του σχήματος φωτιάς εκτιμάται με τη λήψη διαφοράς εικόνας περιγράμματος και αναλογίας λευκού pixel για την εκτίμηση του βαθμού πυρκαγιάς, δηλαδή χωρίς πυρκαγιά, μικρού, μεσαίου και μεγάλου. Τα πειραματικά αποτελέσματα δείχνουν ότι η μέθοδος είναι σε θέση να εντοπίσει διαφορετικά σενάρια πυρκαγιάς με σχετικά καλή ακρίβεια.

Deep Convolutional Neural Networks

Στην εργασία των Sharma et al. [59] χρησιμοποιούνται παραλλαγές γνωστών βαθιών αρχιτεκτονικών CNN και συγκεκριμένα των VGG16 και ResNet50, οι αρχιτεκτονικές των οποίων αναλύονται στο Κεφάλαιο 3. Ειδικότερα, χρησιμοποιήθηκαν και δύο παραλλαγές αυτών των διάσημων δικτύων οι οποίες μάλιστα έδωσαν αποτελέσματα καλύτερα της τάξης του 1% και περιγράφονται στη συνέχεια.

VGG16

Σε αυτή την εργασία, δοκιμάζεται επίσης μια τροποποιημένη έκδοση του VGG16, η οποία αποτελείται από 4 πλήρως συνδεδεμένα στρώματα, με λεπτομέρεια στα δεδομένα εκπαίδευσης, τα οποία ήταν ικανά να αυξήσουν την ακρίβεια της ταξινόμησης. Πραγματοποιήθηκε επίσης δοκιμή και με πιο πλήρως συνδεδεμένα στρώματα, αλλά η αύξηση της ακρίβειας επισκιάστηκε από την αύξηση του χρόνου εκπαίδευσης. Τα Σχήματα 4.6 και 4.7 δείχνουν τις αρχικές και τροποποιημένες αρχιτεκτονικές VGG16 αντίστοιχα.



Σχήμα 4.9: Modified ResNet50 Architecture

να δοκιμάζονται οι δυνατότητες γενίκευσης και την αποτελεσματικότητα και ακρίβεια των μοντέλων για την εξαγωγή σχετικών χαρακτηριστικών από εικόνες όταν τα δεδομένα εκπαίδευσης είναι σπάνια. Το σύνολο δεδομένων χωρίζεται σε σύνολα εκπαίδευσης και αξιολόγησης. Το σύνολο εκπαίδευσης αποτελείται από 549 εικόνες: 59 εικόνες πυρκαγιάς και 490 εικόνες χωρίς πυρκαγιά. Η ανισορροπία είναι σχόπιμη για την αναπαραγωγή πραγματικών καταστάσεων, καθώς η πιθανότητα εμφάνισης πυρκαγιών είναι πολύ μικρή. Τα σύνολα δεδομένων που χρησιμοποιήθηκαν σε προηγούμενες εργασίες ήταν ισορροπημένα και δεν μιμούνταν το περιβάλλον του πραγματικού κόσμου. Το σετ δοκιμών περιέχει 102 εικόνες: 51 εικόνες από τις κατηγορίες φωτιάς και μη πυρκαγιάς αντίστοιχα. Καθώς το σύνολο εκπαίδευσης είναι εξαιρετικά μη ισορροπημένο και το σετ δοκιμών είναι ακριβώς ισορροπημένο, αυτό αποτελεί μια καλή δοκιμή για να δούμε αν τα μοντέλα είναι σε θέση να γενικεύουν καλά ή όχι. Για ένα μοντέλο με καλή ακρίβεια, πρέπει να είναι σε θέση να εξαγάγει τα διακριτικά χαρακτηριστικά από τη μικρή ποσότητα φωτιάς. Για να εξαγάγει τέτοια χαρακτηριστικά από μικρή ποσότητα δεδομένων, το μοντέλο πρέπει να είναι αρκετά βαθύ. Ένα κακό μοντέλο θα επισήμαινε απλώς όλες τις εικόνες ως μη πυρκαγιά.

Εκτός από την έλλειψη ισορροπίας, υπάρχουν μερικές εικόνες που είναι πολύ δύσκολο να ταξινομηθούν. Το σύνολο δεδομένων περιέχει εικόνες από όλα τα σενάρια όπως πυρκαγιά σε σπίτι, δωμάτιο, γραφείο, δασική πυρκαγιά, με διαφορετική ένταση φωτισμού και διαφορετικές αποχρώσεις κόκκινου, κίτρινου και πορτοκαλί, μικρές και μεγάλες φωτιές, φωτιά τη νύχτα, φωτιά το πρωί. Οι εικόνες χωρίς φωτιά περιέχουν μερικές εικόνες που είναι δύσκολο να διακριθούν από εικόνες φωτιάς όπως ένα φωτεινό κόκκινο δωμάτιο με υψηλή φωτισμό, ηλιοβασίλεμα, κόκκινα σπίτια και οχήματα, έντονα φώτα με διαφορετικές αποχρώσεις του κίτρινου και του κόκκινου κ.λπ.

Το Σχήμα 4.10 δείχνει τις εικόνες πυρκαγιάς με διαφορετικά περιβάλλοντα: εσωτερικό και εξωτερικό περιβάλλον, ημέρα, νύχτα, μεγάλη και μικρή πυρκαγιά. Και το Σχήμα 4.11 δείχνει τις εικόνες μη πυρκαγιάς που είναι δύσκολο να ταξινομηθούν. Λαμβάνοντας υπόψη αυτά τα χαρακτηριστικά του συνόλου δεδομένων, η ανίχνευση πυρκαγιάς μπορεί να είναι ένα δύσκολο έργο.

Ο Πίνακας 4.2 δείχνει τα αποτελέσματα των πειραμάτων της συγκεκριμένης εργασίας. Τα βαθιά μοντέλα επιτυγχάνουν ακρίβεια μεγαλύτερη από **90%**. Τα τροποποιημένα μοντέλα VGG16 και Resnet50 ξεπερνούν τα μοντέλα βάσης με μικρό περιθώριο με ελαφρώς υψηλότερο χρόνο εκπαίδευσης.

Φαίνεται προφανές ότι η προσθήκη πλήρως συνδεδεμένων στρωμάτων σε ένα δίκτυο θα αυξήσει την ακρίβεια. Αλλά σε ένα τόσο μικρό σύνολο δεδομένων, η αντιστάθμιση (trade-off) μεταξύ της ακρίβειας και του χρόνου εκπαίδευσης είναι αρκετά κακή, οπότε σταματάμε αφού προσθέσουμε μόνο ένα πλήρως συνδεδεμένο στρώμα.



(α')



(β')

Σχήμα 4.10: Παραδείγματα εικόνων με φωτιά



(α')



(β')

Σχήμα 4.11: Παραδείγματα εικόνων χωρίς φωτιά δύσκολες να κατηγοριοποιηθούν

Model	Train Accuracy %	Test Accuracy %
VGG16	100	90.19
VGG16 modified	100	91.18
ResNet50	100	91.18
ResNet50 modified	100	92.15

Πίνακας 4.2: Σύγκριση των αποτελεσμάτων των τεσσάρων αρχιτεκτονικών σύμφωνα με τους Sharma et al.

4.2 Ανίχνευση Καπνού

Η πυρκαγιά είναι συχνά ξαφνική και ισχυρή, επηρεάζει ένα ευρύ φάσμα και είναι δύσκολο να περιοριστεί. Επομένως, η παρακολούθηση της πυρκαγιάς σε πραγματικό χρόνο έχει καταστεί ιδιαίτερα σημαντική. Η έγκαιρη ανίχνευση πυρκαγιάς είναι το κλειδί για τη μείωση της απώλειας, διότι μόλις εξαπλωθεί η φωτιά θα είναι δύσκολο να ελεγχθεί. Γενικά, η φλόγα είναι μικρή στις πρώτες στιγμές, αλλά ο καπνός είναι πολύ προφανής, επομένως η ανίχνευση του καπνού αποτελεί σημαντική βάση για να καθοριστεί εάν η φωτιά εντοπίστηκε έγκαιρα.

Τα παραδοσιακά συστήματα ανίχνευσης καπνού πυρκαγιάς βασίζονται σε

αισθητήρες που λειτουργούν μόνο όταν ο καπνός είναι κοντά στον αισθητήρα, επομένως δεν είναι κατάλληλοι για ανοιχτό χώρο. Επιπλέον, επειδή οι αισθητήρες είναι επιρρεπείς σε παρεμβολές από τη σκόνη, τη ροή του αέρα και τους ανθρώπινους παράγοντες, αυτά τα συστήματα ανίχνευσης έχουν συχνά υψηλά ποσοστά λανθασμένων συναγεργιών. Με την ανάπτυξη της τεχνολογίας επεξεργασίας βίντεο υψηλής απόδοσης [61], ο αλγόριθμος ανίχνευσης καπνού πυρκαγιάς που βασίζεται στον βίντεο έχει μεγάλες προοπτικές εφαρμογής.

Ο καπνός έχει πλούσια χαρακτηριστικά, όπως χρώμα, υφή και παραμόρφωση, τρεμούλιασμα, τρεμόπαιγμα, συχνότητα. Σήμερα, ο αλγόριθμος ανίχνευσης καπνού που βασίζεται στο βίντεο βασίζεται συνήθως σε ένα ή περισσότερα χαρακτηριστικά καπνού και λαμβάνει αποφάσεις απευθείας ή από ταξινομητές. Οι Toreyin et al. [62] χρησιμοποίησαν κίνηση, τρεμοπαίζει, θόλωση άκρων και χρώμα για ανίχνευση καπνού. Επίσης, διακύμανση των μεγεθών των άκρων εξήχθη για ανίχνευση καπνού, ενώ το πεδίο εφαρμογής του αλγορίθμου είναι περιορισμένο λόγω της ανάγκης να αναλυθεί το φόντο της ενσωματωμένης σκηνής. Οι Chen et al. [63] χρησιμοποίησαν ένα μοντέλο χρώματος για να εξαγάγει τα εικονοστοιχεία πυρκαγιάς και καπνού, για το λόγο αυτό τα δυναμικά μέτρα ανάπτυξης και διαταραχής εξήχθησαν για επαλήθευση. Ο ψευδής ρυθμός ανίχνευσης του αλγορίθμου θα γίνει πολύ υψηλός εάν η σκηνή έχει κινούμενα αντικείμενα πυρκαγιάς. Στους Fujiwara et al. [64] προτείνεται μια τεχνική για την εξαγωγή των περιοχών καπνού από μια εικόνα που χρησιμοποιεί την έννοια της κωδικοποίησης fractal. Για την εικόνα χαμηλής αντίθεσης ή ασαφούς καπνού, τα εξαγόμενα χαρακτηριστικά του fractal δεν είναι αρκετά σταθερά. Οι Ko et al. [65] ανέπτυξαν μια μέθοδο εξαγωγής χαρακτηριστικών με βάση το μοντέλο χωρικών και χρονικών χαρακτηριστικών (Bag of Features - BoF), το οποίο συνδυάζει στατικά και δυναμικά χαρακτηριστικά και στη συνέχεια τα ταξινομεί μέσω random forests. Οι Chenebert et al. [66] πρότειναν μια μη χρονική μέθοδο με γνώμονα την υφή για την ανίχνευση εικονοστοιχείων φωτιάς σε βίντεο ή στατικές εικόνες. Αλλά η μέθοδος δεν χρησιμοποιεί καμία χρονική πληροφορία. Ο Yuan [67] παρουσίασε μια μέθοδο ανίχνευσης καπνού μέσω βίντεο χρησιμοποιώντας ένα μοντέλο συσσωρευμένης κίνησης και υιοθετήθηκε ένας αλγόριθμος βασισμένος σε μπλοκ για τη βελτίωση της αποτελεσματικότητας της ανίχνευσης, αλλά δεν μπορεί να εντοπίσει κινούμενο καπνό προς οποιαδήποτε κατεύθυνση. Ο Tian [68] εξήγαγε τις υποψήφιες περιοχές με το μοντέλο Gaussian Mixture Model (GMM) και η υφή του καπνού περιγράφηκε από το χαρακτηριστικό NR-LBP (Non-Redundant Local Binary Pattern) και στη συνέχεια ταξινομήθηκε με μηχανή διανυσμάτων υποστήριξης (SVM). Οι Genovese et al. [69] πρότειναν μια μέθοδο ανίχνευσης καπνού με βάση τις τεχνικές υπολογιστικής νοημοσύνης. Η διαδικασία ανίχνευσης επικεντρώνεται στην εξαγωγή χαρακτηριστικών όπως η κίνηση, το χρώμα και το άκρο του καπνού. Στη συνέχεια, τα νευρωνικά δίκτυα τροφοδοσίας δύο στρωμάτων χρησιμοποιούνται για την εκτίμηση των περιοχών που περιγράφουν τις περιοχές καπνού στα διάφορα πλαίσια. Ο Yuan [70] εξήγαγε έναν αποτελεσματικό φορέα χαρακτηριστικών, συνδυάζοντας τις αλληλουχίες ιστόγραμμα των πυραμίδων τοπικής δυαδικής διαμόρφωσης (LBP) και τοπικών δυαδικών παραλλαγών (LBPV) και ένα νευρωνικό δίκτυο BP χρησιμοποιήθηκε για ανίχνευση καπνού. Αυτή η μέθοδος έχει υψηλή ακρίβεια στα σύνολα δοκιμών. Όσο για τους Yu et al. [71], υλοποίησαν μια μέθοδο ανίχνευσης καπνού βίντεο χρησιμοποιώντας χαρακτηριστικά χρώματος και κίνησης. Η μέθοδος δεν μπορεί να επιτύχει ποσοστά ανίχνευσης σε πραγματικό χρόνο λόγω της υπολογιστικά δαπανηρής οπτικής ροής. Ο Wang [72] χρησιμοποίησε ένα χρονικό παράθυρο ολίσθησης για να ανιχνεύσει τις περιοχές

"τρεμουλιάσματος", εξήγαγε αυτά τα χαρακτηριστικά, όπως την κατεύθυνση πλεύσης, το περιοδικό μέγεθός του, το θετικό περιοδικό μέγεθος και το αντίστροφο περιοδικό μέγεθος πλεύσης και τελικά αποφάσισε αν ήταν καπνός μέσω ενός εκπαιδευμένου ασαφούς νευρωνικού δικτύου (fuzzy neural network). Με το συμπυκνωμένο βίντεο, ο Luo [73] διαπίστωσε ότι οι τροχιές καπνού έχουν κάποια ειδικά χαρακτηριστικά, όπως η ευθεία γραμμή, η ομαλή ροή, η χαμηλή συχνότητα, η σταθερή πηγή και η κατακόρυφη οριζόντια αναλογία. Αυτή η μέθοδος είναι πολύ νέα, αλλά έχει κακή επίδραση ανίχνευσης στον αργά διαδιδόμενο καπνό. Στα [56] και [43] χρησιμοποίησαν CNN για να επιτύχουν την ανίχνευση του καπνού ή της φλόγας, αλλά η αξιολόγηση είναι επιπέδου εικόνας και δεν λαμβάνει υπόψη την πολυπλοκότητα του χρονικού περιθωρίου του συρόμενου παραθύρου. Μέχρι τώρα, εξακολουθεί να είναι δύσκολο να ανιχνευθεί ο καπνός σε βίντεο λόγω της μεταβλητότητας των χαρακτηριστικών του.

Συμπερασματικά, όλοι αυτοί οι αλγόριθμοι βασίζονται σε χαρακτηριστικά που έχουν επιλεγεί και επεξεργαστεί χειροκίνητα, τα οποία απαιτούν συγκεκριμένη εμπειρία και επιλεγμένες λειτουργίες έχουν καλή απόδοση σε ορισμένες σκηνές, αλλά ενδέχεται να μην είναι έγκυρες σε άλλες σκηνές λόγω της μεγάλης ποικιλίας καπνού. Στην πρόσφατη εργασία των Luo et al. [74], εφαρμόζεται CNN υψηλής δυναμικότητας στην ανίχνευση καπνού βίντεο. Ο αλγόριθμος μπορεί όχι μόνο να εξαγάγει χαρακτηριστικά καπνού αυτόματα, αλλά και να βελτιώσει την ακρίβεια ανίχνευσης και άλλα κριτήρια. Οι συνεισφορές αυτής της εργασίας συνοψίζονται στα εξής:

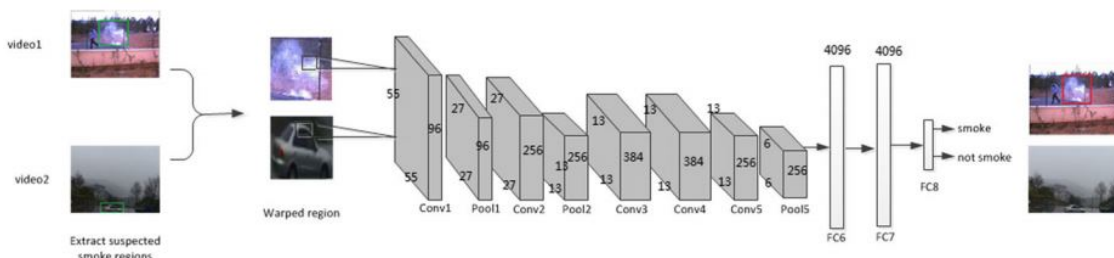
1. Προτείνεται αλγόριθμος δύο βημάτων για την ανίχνευση πυρκαγιάς βάσει βίντεο, στον οποίο χρησιμοποιείται το χαρακτηριστικό κίνησης για την εξαγωγή υποψιών καπνού που όχι μόνο μειώνει τον αριθμό των υποψήφιων περιοχών για να ικανοποιήσει τις απαιτήσεις σε πραγματικό χρόνο αλλά επίσης μειώνει την πιθανότητα ψευδούς ένδειξης.
2. Επί του παρόντος, τα σύνολα δεδομένων των εικόνων καπνού είναι ελλιπή και δύσκολο να ληφθούν, με αποτέλεσμα το ανεπαρκές μοντέλο εκπαίδευσης να καταλήγει σε υπερβολική τοποθέτηση (overfitting). Η επιτάχυνση της εκπαίδευσης επιτυγχάνεται καθώς και ένα υψηλότερο ποσοστό ακρίβειας κάτω από τον ίδιο αριθμό επαναλήψεων.
3. Η άμεση αναμόρφωση της υποψήφιας περιοχής μπορεί να μειώσει τα βασικά χαρακτηριστικά της εικόνας όταν η περιοχή είναι ιδιαίτερα μικρή. Προτείνεται μία σιωπηρή στρατηγική, δηλαδή να διευρυνθούν οι υπόνοιες περιφερειών ώστε να βελτιωθεί η ακρίβεια και η αποτελεσματικότητα της ανίχνευσης.

Μια ποικιλία πρόσφατων ερευνών προσφέρει τεχνικές για τη δημιουργία κατηγορίας ανεξάρτητης περιοχής στο πεδίο της ανίχνευσης στόχων. Για παράδειγμα, επιλεκτική αναζήτηση [75], περιορισμένες παραμετρικές μικρο-διασπάσεις (CPMC) [76] κλπ. Αλλά ο καπνός δεν είναι ο ίδιος με τον γενικό στόχο, ο οποίος δεν έχει εμφανή χαρακτηριστικά χρώματος και περιγράμματος και δεν χρειαζόμαστε την εξαγωγή των υποψήφιων περιοχών για κάθε πλαίσιο λόγω της πιθανότητας πυρκαγιάς είναι πολύ μικρή στην πραγματική σκηνή, οπότε αυτοί οι αλγόριθμοι δεν είναι κατάλληλοι για τις υποψίες περιοχές εξαγωγή καπνού. Λαμβάνοντας υπόψη τα χαρακτηριστικά κίνησης του καπνού, προτείνεται μια μέθοδος ανίχνευσης κίνησης για την εξαγωγή πιθανών περιοχών καπνού.

Υπό τις συνθήκες του στατικού εξοπλισμού παρακολούθησης, οι κοινές μέθοδοι ανίχνευσης κίνησης είναι η μέθοδος διαφοράς πλαισίου (frame difference) [77], η μέθοδος Gaussian mixture [78] και η μέθοδος οπτικής ροής (optical flow) [79]. Αυτές οι μέθοδοι έχουν καλή επίδραση στην ανίχνευση άκαμπτων αντικειμένων, αλλά για την ανίχνευση καπνού που δεν είναι άκαμπτος και το μοντέλο κίνησης παρουσιάζει διάχυση, δηλαδή είναι εύκολο να παράγει κοίλο φαινόμενο, επομένως είναι δύσκολο να εξαχθεί ολόκληρη η περιοχή του καπνού. Επιπλέον, τα δέντρα, οι πεζοί, τα αυτοκίνητα και άλλα κινούμενα αντικείμενα είναι πολύ συχνά στη φυσική σκηνή. Τα εικονοστοιχεία του προσκηνίου που παράγονται από αυτά τα κινούμενα αντικείμενα μπορούν να εξάγονται ως περιοχές ύποπτων καπνών, αν δεν υπάρχει ειδική επεξεργασία, η οποία όχι μόνο επηρεάζει την αποτελεσματικότητα του αλγορίθμου αλλά και βελτιώνει την πιθανότητα ψευδούς συναγερμού σε κάποιο βαθμό.

Δίκτυο για Αναγνώριση Καπνού

Υπάρχουν πολλές παραλλαγές των αρχιτεκτονικών του CNN στη βιβλιογραφία. Ωστόσο, τα βασικά συστατικά τους είναι πολύ παρόμοια. Η προτεινόμενη αρχιτεκτονική δικτύου χρησιμοποιείται για την εξαγωγή χαρακτηριστικών και την ταξινόμηση του καπνού. Όπως φαίνεται στο Σχήμα 4.12, το δίκτυο έχει οκτώ στρώματα εκτός από το το στρώμα εισόδου και χωρίς να λαμβάνεται υπόψη η αυστηρή διάκριση μεταξύ των συνθετικών στρώσεων και των pooling στρώσεων, αποτελείται από 5 στρώματα συνέλιξης και 3 πλήρως συνδεδεμένα στρώματα. Η επιλογή για σχεδιασμό τέτοιου είδους δικτύου είναι κίνητρο τόσο από την επιθυμία να μειωθεί τον κίνδυνο υπερφόρτωσης καθώς και την επίτευξη υψηλής ακρίβειας μόνο από το απλό δίκτυο.



Σχήμα 4.12: Συνελικτικό νευρωνικό δίκτυο για αναγνώριση καπνού

Για να εκτελεστεί η λειτουργία πολλαπλών επιπέδων για την εικόνα εισόδου, το μέγεθος της εικόνας εισόδου για την αρχιτεκτονική δικτύου είναι σταθερό στα 227×227 pixels. Για την έγχρωμη εικόνα που περιέχει συνήθως τρία κανάλια χρώματος, το συνολικό μέγεθος της εικόνας εισόδου είναι $227 \times 227 \times 3$. Το στρώμα συνέλιξης στοχεύει στο να μάθει τις αναπαραστάσεις χαρακτηριστικών των εισόδων. Αυτό το στρώμα αποτελείται από αρκετούς πυρήνες συνέλιξης, οι οποίοι χρησιμοποιούνται για τον υπολογισμό των διαφορετικών χαρτογραφικών χαρακτηριστικών. Το μέγεθος του πυρήνα συνέλιξης καθορίζει το μέγεθος του χάρτη εξόδου.

Το πρώτο πλήρως συνδεδεμένο στρώμα λαμβάνει την έξοδο του πέμπτου συνελικτικού στρώματος και περιέχει 4096 νευρώνες, ακολουθούμενη από μία Relu και ένα στρώμα απόρριψης. Το δεύτερο πλήρως συνδεδεμένο στρώμα που λαμβάνει την έξοδο 4096 διαστάσεων του πρώτου πλήρως συνδεδεμένου στρώματος και περιέχει πάλι

4096 νευρώνες, ακολουθούμενη από μία Relu και ένα στρώμα απόρριψης. Το τελευταίο πλήρως συνδεδεμένο στρώμα περιέχει δύο νευρώνες οι οποίοι χαρτογραφούν τις τελευταίες κατηγορίες για καπνό και μη καπνό.

Δεδομένα και Πειράματα

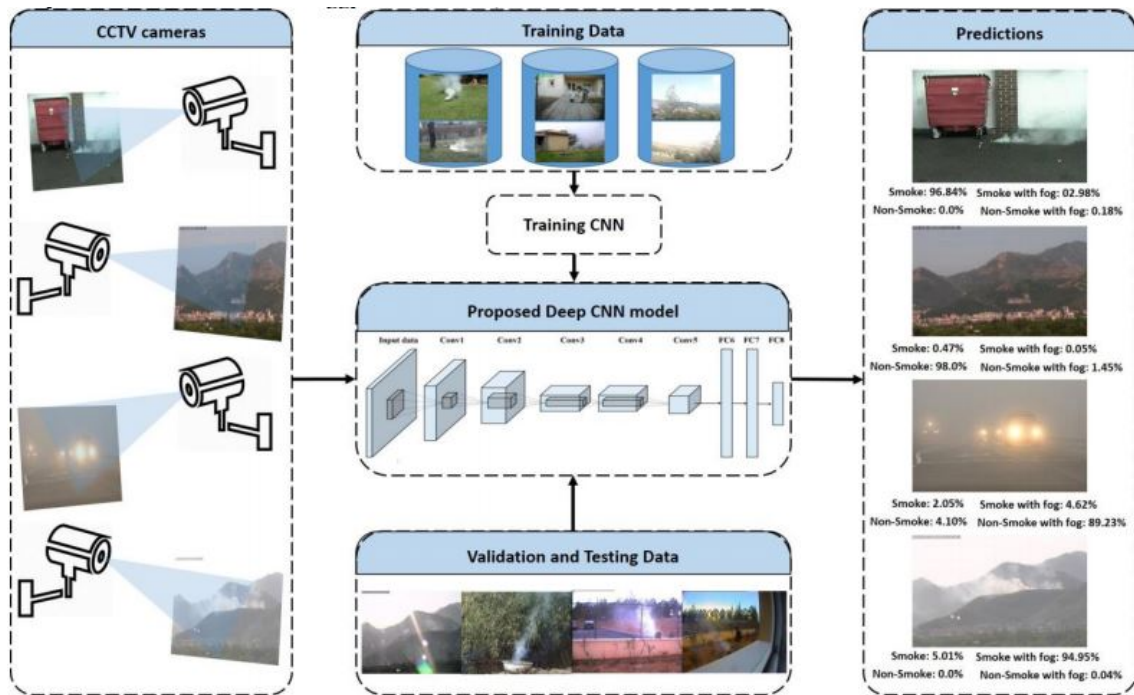
Το σύνολο εκπαίδευσης περιέχει 6776 εικόνες καπνού και 12.352 εικόνες χωρίς καπνό, συνολικού ύψους 19.128 θετικών και αρνητικών δειγμάτων. Το πρώτο σύνολο αξιολόγησης (test1) περιέχει 1548 εικόνες καπνού και 1880 εικόνες χωρίς καπνό, συνολικά 3428 θετικά και αρνητικά δείγματα. Το test2 περιέχει 1872 εικόνες καπνού και 3962 εικόνες χωρίς καπνό, συνολικού αριθμού 5834 θετικών και αρνητικών δειγμάτων.

Τα πειράματα διεξήχθησαν με τυχαία αρχικοποίηση και ρύθμιση των παραμέτρων για την εκπαίδευση και τη δοκιμή του μοντέλου. Αντίστοιχα, η τυχαία μέθοδος αρχικοποίησης συγκλίνει μετά από 15.000 επαναλήψεις, ενώ μπορεί να επιτευχθεί υψηλή ακρίβεια με μικρότερο αριθμό επαναλήψεων με λεπτό συντονισμό που μπορεί να υπερβεί το **98%** μετά από περίπου 1000 επαναλήψεις. Αυτό ελαττώνει σημαντικά την έλλειψη δεδομένων καπνού κατά την εκπαίδευση του δικτύου. Επίσης, η ακρίβεια της μεθόδου fine-tuning είναι υψηλότερη από την τυχαία αρχικοποίηση στις ίδιες επαναλήψεις.

Ακόμα πιο πρόσφατες εργασίες που υπάρχουν όσον αφορά την αναγνώριση του καπνού της φωτιάς σε εικόνα και βίντεο χρησιμοποιούν κι αυτές μεγάλα υπάρχοντα δίκτυα, όπως αυτή των Khan et al. [80] που συγκρίνει τα VGG16, GoogleNet και AlexNet.

Πιο συγκεκριμένα, ερευνήθηκαν διαφορετικά μοντέλα CNN για ανίχνευση καπνού σε περιβάλλον ομίχλης και Internet of Things (IoT) και προτείνεται ένα ενεργειακά αποδοτικό πλαίσιο CNN για έγκαιρη ανίχνευση καπνού τόσο σε κανονικό όσο και σε ομιχλώδες περιβάλλον. Η λογική υπολογιστική πολυπλοκότητα της αρχιτεκτονικής αυτής, μαζί με την ακρίβεια και το μέγεθος του μοντέλου της, το καθιστούν κατάλληλο σύστημα για τον εντοπισμό καπνού στις έξυπνες πόλεις που υποστηρίζονται από το IoT σε σχέση με τις τελευταίες τεχνολογίες. Η ανίχνευση καπνού με τη χρήση χειροποίητων (hand-crafted) χαρακτηριστικών στην επιτήρηση έξυπνων σημείων είναι μια κουραστική δουλειά, ιδιαίτερα όταν ο καπνός βρίσκεται στα αρχικά του στάδια ή σε μεγάλη απόσταση ή σε αβέβαιο περιβάλλον. Τα υπάρχοντα συστήματα που χρησιμοποιούν την προσέγγιση παραδοσιακής εξαγωγής χαρακτηριστικών, παράγουν υψηλότερο αριθμό ψευδών συναγερμών με περιορισμένη ακρίβεια ανίχνευσης καπνού. Πρόσφατα διερευνώνται αρκετές μέθοδοι ανίχνευσης καπνού των οποίων η απόδοση είναι λογική σε σύγκριση με τις παραδοσιακές μεθόδους. Ωστόσο, το κύριο ζήτημα με αυτές τις μεθόδους είναι η περιορισμένη επίδοσή τους για αβέβαιο και ομιχλώδες περιβάλλον. Για να ξεπεραστούν αυτά τα ζητήματα, ερευνήθηκαν διαφορετικές αρχιτεκτονικές CNN και προτείνεται μια ενεργειακά αποδοτική τεχνολογία βασισμένη στο CNN για ανίχνευση καπνού σε βίντεο παρακολούθησης που συλλαμβάνονται σε περιβάλλον φυσιολογικού και ομιχλώδους περιβάλλοντος IoT. Στο ακόλουθο Σχήμα 4.13 παρουσιάζεται το framework που χρησιμοποιήθηκε.

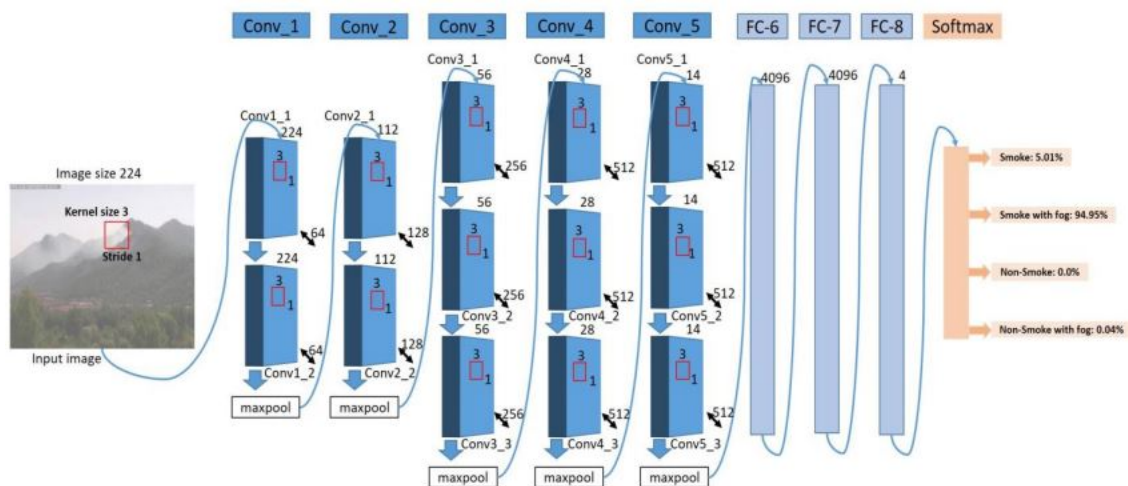
Στην προτεινόμενη μέθοδο διερευνήθηκαν και συγκρίθηκαν διάφορα μοντέλα CNN με διαφορετικές ρυθμίσεις παραμέτρων για την ανίχνευση καπνού σε κανονικά και ομιχλώδη περιβάλλοντα. Μετά τη σύγκριση των διαφόρων CNN, διαπιστώνουμε ότι



Σχήμα 4.13: Framework για αναγνώριση καπνού σε περιβάλλον IoT

το VGG-16 είναι καλύτερο από άλλα μοντέλα, όπως τα AlexNet και GoogleNet. Για την ανίχνευση καπνού σε περιβάλλον αβέβαιου IoT, τροποποιήθηκε η αρχιτεκτονική του μοντέλου VGG-16 σύμφωνα με το πρόβλημά μας. Παρόμοια με άλλα CNN, το τρέχον VGG-16 είναι προ-εκπαιδευμένο στο σύνολο δεδομένων ImageNet για την ταξινόμηση των εικόνων σε 1000 διαφορετικές κατηγορίες. Επεξεργάστηκε λεπτομερώς αυτό το μοντέλο με το σύνολο δεδομένων που δημιουργήθηκε προσαρμόζοντας το τελευταίο πλήρως συνδεδεμένο στρώμα από 1000 σε τέσσερις κατηγορίες για να εκτελεστεί η προβλεπόμενη ταξινόμηση του καπνού και του μη καπνού σε κανονικά και ομιχλώδη περιβάλλοντα IoT. Έτσι, η τροποποίηση του τελευταίου στρώματος της προαναφερθείσας αρχιτεκτονικής επέτρεψε στο σύστημα να εντοπίζει αποτελεσματικά τα πρότυπα καπνού. Η αρχιτεκτονική του απασχολούμενου μοντέλου CNN δίνεται στο Σχήμα 4.14. Το μέγεθος της εικόνας εισόδου στην αρχιτεκτονική είναι σταθερό στα $224 \times 224 \times 3$ pixels. Κάθε εικόνα περνάει μέσα από πέντε διαφορετικές συνελίξεις της αρχιτεκτονικής. Η πρώτη συνέλιξη αποτελείται από δύο συνελικτικά στρώματα με μέγεθος εισόδου 224×224 , όπου εφαρμόζονται 64 πυρήνες μεγέθους 3×3 με βήμα 1. Το αποτέλεσμα μεταδίδεται προς τα εμπρός στο στρώμα συγκέντρωσης (max pooling) με πυρήνα 2×2 και βήμα 2 για να ληφθούν οι μέγιστες ενεργοποιήσεις από χάρτες χαρακτηριστικών. Η δεύτερη συνέλιξη είναι ίδια με την προηγούμενη, η οποία αποτελείται από δύο στρώματα συνέλιξης με μέγεθος εισόδου 112×112 ακολουθούμενη από max pooling. Η τρίτη συνέλιξη αποτελείται από τρία συνεργατικά στρώματα με 256 πυρήνες μεγέθους 3×3 με βήμα 1 στην είσοδο μεγέθους 56×56 που ακολουθείται και αυτή από max pooling. Οι επόμενες δύο συνέλιξεις είναι ίδιες με την τρίτη με μέγεθος εισόδου 28×28 και 14×14 αντίστοιχα. Μια στοίβα από αυτά τα συνελικτικά στρώματα ακολουθείται από τρεις πλήρως συνδεδεμένα στρώματα. Το πρώτο και το δεύτερο πλήρως συνδεδεμένο στρώμα έχουν 4096 κανάλια ανά τεμάχιο, ενώ το τρίτο πλήρως συνδεδεμένο στρώμα τροποποιείται από 1000 έως 4 κανάλια για την ταξινόμηση του καπνού και του μη καπνού σε ένα ορισμένο και αβέβαιο περιβάλλον.

Τέλος, ο ταξινομητής Softmax χρησιμοποιείται για την πρόβλεψη των πιθανοτήτων τεσσάρων κατηγοριών στόχων (καπνός, καπνός με ομίχλη, μη καπνός και μη καπνός με ομίχλη).



Σχήμα 4.14: Αρχιτεκτονική του CNN. Ο ταξινομητής Softmax παρέχει τις τελικές προβλέψεις για τις τέσσερις προβλεπόμενες κατηγορίες, όπως "Smoke", "Smoke with fog", "Non-Smoke" και "Non-Smoke with fog".

Το σύνολο δεδομένων που χρησιμοποιείται για τους πειραματισμούς αποτελείται από εικόνες με τέσσερις διαφορετικές κατηγορίες: 'Μη καπνός', 'Καπνός', 'Μη καπνός με ομίχλη' και 'Καπνός με ομίχλη'. Σε αυτό το σύνολο δεδομένων, οι κλάσεις 'μη καπνός' και 'καπνός' ανήκουν σε τρία σύνολα δεδομένων αναφοράς ενώ οι άλλες δύο κλάσεις 'Μη καπνός με ομίχλη' και 'Καπνός με ομίχλη' δημιουργούνται συνθετικά προσθέτοντας ομίχλη σε κάθε εικόνα των δύο πρώτων τάξεων. Το σύνολο δεδομένων αποτελείται από συνολικά 72012 εικόνες, που εξάγονται από βίντεο και είναι το λεγόμενο MIVIA Dataset που περιγράφεται στο Κεφάλαιο 5. Για την αξιολόγηση του συστήματος, το 20% των δεδομένων χρησιμοποιήθηκε για την εκπαίδευση (training) και το υπόλοιπο 30% και 50% για την επικύρωση (validation) και τον έλεγχο (testing), αντίστοιχα. Χρησιμοποιώντας αυτή τη διανομή δεδομένων, το μοντέλο εκπαιδεύεται με 3495 εικόνες μη καπνού, 3706 εικόνες καπνού, 3495 μη καπνού με εικόνες ομίχλης και 3706 καπνού με εικόνες ομίχλης.

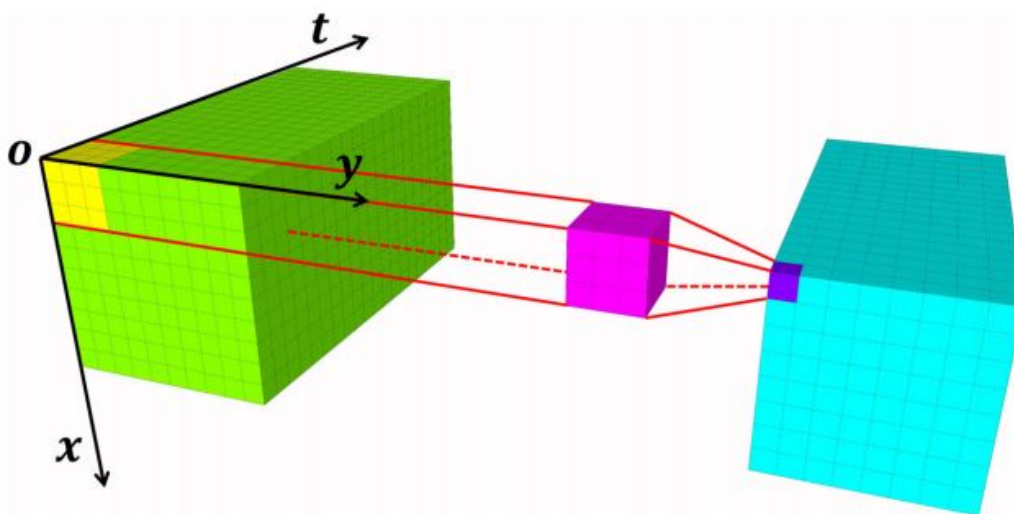
Στον ακόλουθο Πίνακα 4.3 παρουσιάζονται τα accuracies τόσο της προτεινόμενης μεθόδου των Khan et al. όσο και των άλλων δύο state-of-the-art δικτύων.

Model	Accuracy
AlexNet	95.87%
GoogleNet	96.11%
VGG-16	97.72%

Πίνακας 4.3: Σύγκριση των αποτελεσμάτων των τριών αρχιτεκτονικών σύμφωνα με τους Khan et al.

Μία διαφορετική προσέγγιση είναι αυτή των Lin et al. [81], όπου γίνεται χρήση 3D CNN δικτύου το οποίο αναγνωνρίζει καπνό συνδυάζοντας δυναμική και χωροχρονική πληροφορία.

Συγκεντρωτικά δίκτυα μπορούν να εφαρμοστούν σε δεδομένα αυθαίρετων διαστάσεων στη θεωρία. Αυτό έχει εμπνεύσει τους ερευνητές να επεκτείνουν το 2D χωρικό CNN σε 3D χωροχρονικό CNN για ανάλυση βίντεο. Γενικά, υπάρχουν τρία αρχιτεκτονικά σχεδιαστικά χωροχρονικά CNN: 2-διάστατο CNN, 3D CNN [82] και LSTM [83]. Τα τρισδιάστατα CNN είναι τα πιο απλά χωροχρονικά δίκτυα που χρησιμοποιούν 3D συνέλιξη για να λειτουργούν σε ακολουθίες βίντεο. Το Σχήμα 4.15 απεικονίζει τη διαφορά στο 2D φίλτρο συνέλιξης και το φίλτρο 3D.



Σχήμα 4.15: Διαφορά μεταξύ 2D και 3D συνελικτικού φίλτρου

Το 3D CNN για χωρική και χρονική εξαγωγή χαρακτηριστικών που χρησιμοποιείται στο πλαίσιο αυτής της εργασίας βασίζεται στο C3D-v1.0. Το σύνολο δεδομένων Sports-1M [84], το οποίο αποτελεί σήμερα το μεγαλύτερο σημείο αναφοράς για την ταξινόμηση βίντεο, αποτελείται από 1.1 εκατομμύρια αθλητικά βίντεο και 487 αθλητικές κατηγορίες. Ακόμη, ένα μικρότερο σύνολο δεδομένων είναι το UCF101 και αποτελείται από 13.320 βίντεο μερικών δευτερολέπτων και 101 κατηγορίες ανθρώπινης δράσης. Η υπερφόρτωση είναι πολύ δυνατή, καθώς το σύνολο δεδομένων που χρησιμοποιείται είναι πολύ μικρότερο από το Sports-1M και το UCF101. Εφαρμόζεται ένα μικρότερο δίκτυο για να αποτραπεί η υπερφόρτωση. Το δίκτυο αποτελείται από πέντε συνελικτικά 3D στρώματα, πέντε στρώματα 3D max pooling και τρία πλήρως συνδεδεμένα στρώματα, καθώς και ένα softmax στρώμα για την πρόβλεψη ετικετών καπνού. Ως συνάρτηση ενεργοποίησης, τοποθετείται μια ReLU στην έξοδο κάθε τρισδιάστατου συνελικτικού στρώματος και των δύο πρώτων πλήρως συνδεδεμένων στρώματων. Τα πρώτα δύο πλήρως συνδεδεμένα στρώματα ακολουθούνται από τα στρώματα dropout [85] για να αποφευχθεί το overfitting.

Στις παραδοσιακές χρήσεις του C3D, η είσοδος είναι ένα σύντομο βίντεο κλιπ 16 καρέ. Ωστόσο, η διάχυση του καπνού είναι πολύ αργή, ιδιαίτερα σε περιβάλλον παρακολούθησης μεγάλων αποστάσεων, όπως τα δάση. Η κίνηση είναι ελάχιστα παρατηρήσιμη σε χρονικά πλαίσια μικρότερα του δευτερολέπτου. Εκτός από τα κλιπ 16

πλαισίων, χρησιμοποιούνται κλιπ 32 και 60 καρέ ως είσοδοι στο 3D CNN. Όλα τα κλιπ έχουν μετασχηματιστεί σε 128×171 . Οι διαστάσεις εισόδου είναι $3 \times length \times 128 \times 171$ και οι τυχαίες διασπάσεις είναι μεγέθους $3 \times length \times 112 \times 112$ και χρησιμοποιούνται κατά τη διάρκεια της εκπαίδευσης. Το σύνολο δεδομένων αποτελείται από 38 βίντεο καπνού και 20 βίντεο χωρίς καπνό. Παρόλο που ο καπνός είναι σχετικά ελαφρός στην αρχή κάποιων βίντεο καπνού, τα πλαίσια εξακολουθούν να επισημαίνονται ως πλαίσια καπνού με την ετικέτα 1. Τα βίντεο καπνού χωρίζονται σε ένα σύνολο δεδομένων εκπαίδευσης και ένα σύνολο δεδομένων δοκιμών, που περιέχει 20 και 18 βίντεο, αντίστοιχα.

Όπως προκύπτει, το framework του τρισδιάστατου 3D CNN αποδίδει το καλύτερο αποτέλεσμα, με ένα ποσοστό ανίχνευσης **95.23%** και χαμηλό ποσοστό ψευδούς συναγερμού **0.39%** για ακολουθίες βίντεο καπνού, νούμερα τα οποία είναι πολύ καλύτερα από το κλασικό RCNN.

4.3 Real-Time Ανίχνευση

Μια έρευνα της υπάρχουσας βιβλιογραφίας, όπως περιγράφηκε, δείχνει ότι οι υπολογιστικά ακριβείς μέθοδοι έχουν μεγαλύτερη ακρίβεια, ενώ απλούστερες μέθοδοι συμβιβάζονται στην ακρίβεια και τον ρυθμό των ψευδών θετικών. Ως εκ τούτου, υπάρχει ανάγκη να βρεθεί ένας καλύτερος συνδυασμός μεταξύ αυτών των μετρήσεων για διάφορα σενάρια εφαρμογών πρακτικού ενδιαφέροντος, για τα οποία οι υπάρχουσες υπολογιστικές μέθοδοι δεν αποδίδουν καλά.

Για να επιλύσει τα παραπάνω ζητήματα, η εργασία των Muhammad et al. [86] διερευνάει τα βαθιά χαρακτηριστικά των συνελκτικών νευρωνικών δικτύων (CNN) για έγκαιρη ανίχνευση πυρκαγιάς στα δίκτυα επιτήρησης. Οι πρωταρχικές συνεισφορές μπορούν να συνοψιστούν ως εξής:

1. Αποφεύγονται οι χρονοβόρες προσπάθειες συμβατικών χειροποίητων χαρακτηριστικών για ανίχνευση πυρκαγιάς και διερευνώνται αρχιτεκτονικές βαθιάς μάθησης για έγκαιρη ανίχνευση πυρκαγιάς σε δίκτυα επιτήρησης κλειστού κυκλώματος τηλεόρασης (CCTV) τόσο για εσωτερικά όσο και για υπαίθρια περιβάλλοντα. Το προτεινόμενο πλαίσιο ανίχνευσης πυρκαγιάς βελτιώνει την ακρίβεια ανίχνευσης και μειώνει τους ψευδείς συναγερμούς, σε σύγκριση με τις πλέον σύγχρονες μεθόδους. Έτσι, ο αλγόριθμος μπορεί να διαδραματίσει ζωτικό ρόλο στην έγκαιρη ανίχνευση πυρκαγιάς για ελαχιστοποίηση των ζημιών.
2. Γίνεται εκπαίδευση και fine-tuning της αρχιτεκτονικής AlexNet για την ανίχνευση πυρκαγιάς χρησιμοποιώντας μια στρατηγική μεταφοράς μάθησης. Το μοντέλο ξεπερνά τις συμβατικές μεθόδους ανίχνευσης πυρκαγιάς βασισμένες σε χειροκίνητα χαρακτηριστικά. Ωστόσο, παραμένει συγκριτικά μεγάλο σε μέγεθος (238 MB), καθιστώντας την υλοποίησή του δύσκολη σε εξοπλισμό με περιορισμένους πόρους.
3. Για να μειώσουν το μέγεθος του μοντέλου, προσαρμόζουν με προσοχή ένα μοντέλο με παρόμοια αρχιτεκτονική με το μοντέλο SqueezeNet [87] για ανίχνευση πυρκαγιάς στα αρχικά στάδια. Το μέγεθος του μοντέλου μειώθηκε από 238 MB σε 3 MB, εξοικονομώντας έτσι επιπλέον χώρο 235 MB, ελαχιστοποιώντας το κόστος και καθιστώντας την υλοποίησή του πιο εφικτή στα δίκτυα επιτήρησης. Επιπλέον, το προτεινόμενο μοντέλο απαιτεί 0.72

GFLOPS/εικόνα σε σύγκριση με το AlexNet, του οποίου η υπολογιστική πολυπλοκότητα είναι 2 GFLOPS/εικόνα. Αυτό καθιστά το προτεινόμενο μοντέλο πιο αποτελεσματικό όσον αφορά την εξαγωγή συμπερασμάτων, επιτρέποντάς του να επεξεργάζεται πολλαπλές ροές παρακολούθησης.

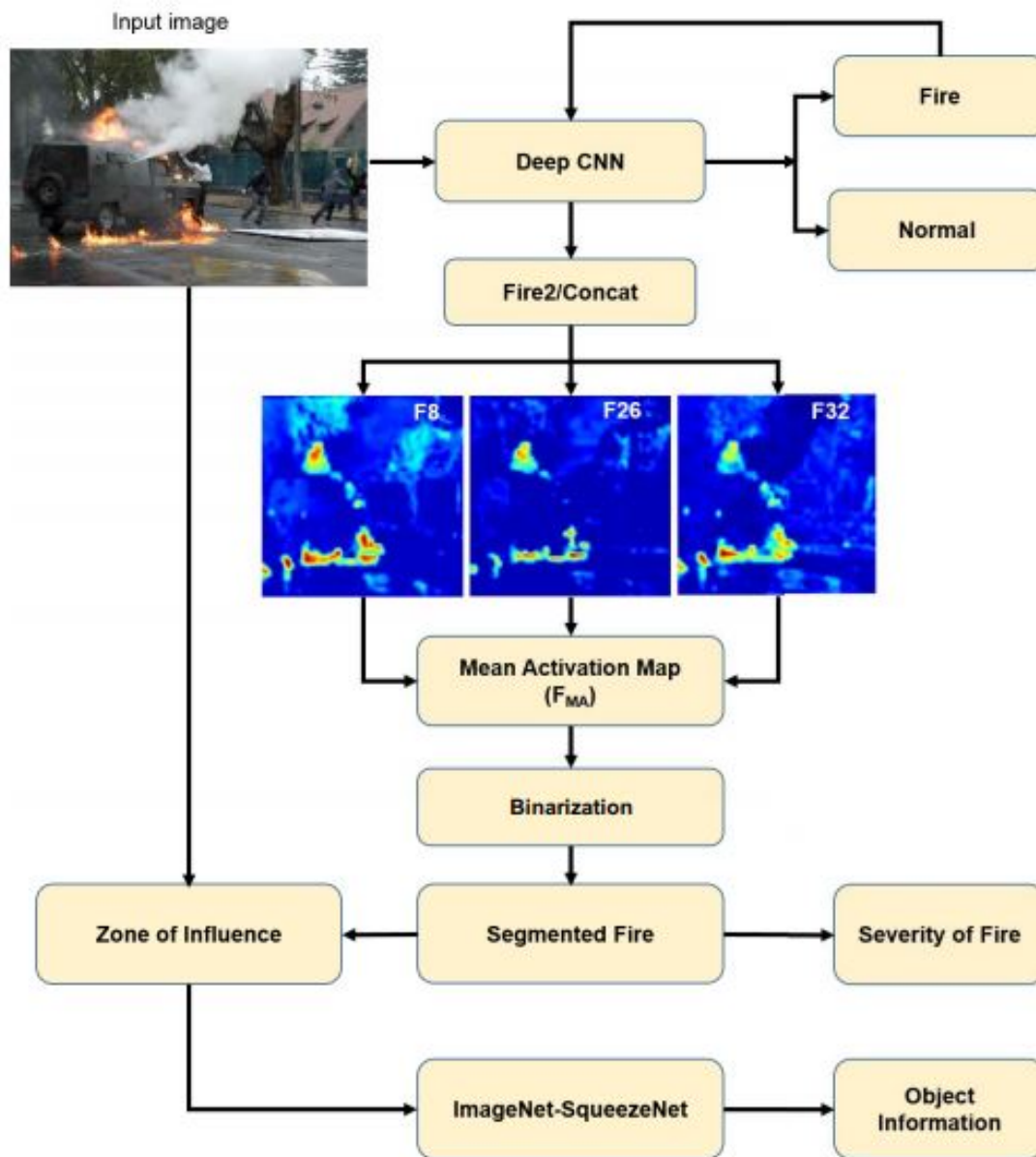
4. Αναπτύσσεται επίσης ένας αλγόριθμος επιλογής χαρτών χαρακτηριστικών, ο οποίος μπορεί έξυπνα να διαλέξει κατάλληλους χάρτες χαρακτηριστικών από τα συνελικτικά στρώματα του εκπαιδευμένου CNN, τα οποία είναι ευαίσθητα στις περιοχές πυρκαγιάς. Αυτοί οι χάρτες χαρακτηριστικών επιτρέπουν μια ακριβέστερη κατανομή της φωτιάς σε σύγκριση με άλλες μεθόδους. Οι πληροφορίες κατάτμησης μπορούν να αναλυθούν περαιτέρω για να εκτιμηθούν τα ουσιαστικά χαρακτηριστικά της φωτιάς, για παράδειγμα ο ρυθμός ανάπτυξής της. Χρησιμοποιώντας αυτή την προσέγγιση, μπορεί επίσης να προσδιοριστεί η σοβαρότητα της πυρκαγιάς ή/και ο βαθμός καύσης της. Ένα άλλο καινοτόμο χαρακτηριστικό του συστήματος αυτού είναι η ικανότητα εντοπισμού του αντικειμένου που καίγεται, χρησιμοποιώντας ένα pre-trained μοντέλο εκπαιδευμένο σε 1000 κατηγορίες αντικειμένων στο σύνολο δεδομένων ImageNet. Αυτό επιτρέπει την προσέγγιση αυτή να καθορίσει εάν η φωτιά βρίσκεται σε ένα αυτοκίνητο, ένα σπίτι, ένα δάσος ή οπουδήποτε αλλού. Χρησιμοποιώντας αυτές τις σημασιολογικές πληροφορίες, οι πυροσβέστες μπορούν να δώσουν προτεραιότητα στους στόχους τους επικεντρώνοντας κυρίως στις περιοχές με τη μεγαλύτερη επικινδυνότητα.

Μια δεδομένη εικόνα, τροφοδοτείται προς τα εμπρός μέσα από το βαθύ CNN, το οποίο αποδίδει μια ετικέτα "φωτιά" ή "κανονική" στην εικόνα εισόδου. Αυτή η ετικέτα αποδίδεται με βάση τις πιθανότητες που υπολογίζονται από το δίκτυο. Η υψηλότερη βαθμολογία πιθανότητας θεωρείται ότι είναι η τελική ετικέτα κατηγορίας της εικόνας εισόδου. Για να εντοπιστεί μια φωτιά σε μια εικόνα δείγματος, χρησιμοποιείται το πλαίσιο που δίνεται στο Σχήμα 4.16. Πρώτον, μια πρόβλεψη λαμβάνεται από το βαθύ CNN. Σε περιπτώσεις μη πυρκαγιάς, δεν γίνεται καμία άλλη ενέργεια. Αφού αναλυθούν όλοι οι χάρτες χαρακτηριστικών των διαφόρων επιπέδων του προτεινόμενου CNN χρησιμοποιώντας τον αλγόριθμο του σχήματος 4.16, οι χάρτες χαρακτηριστικών 8, 26 και 32 της στρώσης "Fire2/Concat" βρίσκονται ευαίσθητοι ως προς την ύπαρξη πυρκαγιάς και είναι κατάλληλοι για τον εντοπισμό της. Συνεπώς, αυτοί οι τρεις χάρτες χαρακτηριστικών συγχωνεύτηκαν και εφαρμόστηκαν στην αποικοδόμηση για τον χωρισμό της φωτιάς. Η κατακερματισμένη πυρκαγιά χρησιμοποιείται για δύο περαιτέρω σκοπούς:

- καθορισμός της σοβαρότητας βαθμού/καύσης της σκηνής υπό παρατήρηση και
- εύρεση της ζώνης επιρροής (ZOI) από την εικόνα φωτιάς εισόδου.

Ο βαθμός καύσης μπορεί να προσδιοριστεί από τον αριθμό των εικονοστοιχείων στην κατακερματισμένη φωτιά. Το ZOI μπορεί να υπολογιστεί αφαιρώντας τις κατακερματισμένες περιοχές πυρκαγιάς από την αρχική εικόνα εισόδου. Η προκύπτουσα εικόνα ZOI μεταφέρεται στη συνέχεια από το αρχικό μοντέλο SqueezeNet, το οποίο προβλέπει την ετικέτα της από 1000 αντικείμενα. Οι πληροφορίες αντικειμένων μπορούν να χρησιμοποιηθούν για τον προσδιορισμό της κατάστασης στη σκηνή, όπως πυρκαγιά σε σπίτι, δάσος ή όχημα. Αυτές οι πληροφορίες, μαζί με τη σοβαρότητα της

πυρκαγιάς, μπορούν να αναφερθούν στην πυροσβεστική υπηρεσία για να λάβουν τα κατάλληλα μέτρα.



Σχήμα 4.16: Προτεινόμενος αλγόριθμος ανίχνευσης φωτιάς των Muhammad et al.

Συνολικά για το πείραμα χρησιμοποιήθηκαν 68457 εικόνες, οι οποίες προέρχονται από δύο datasets: 1) Foggia et al. [88] (Dataset1) και Chino et al. [49] (Dataset2). Το σύνολο δεδομένων Dataset1 περιέχει συνολικά 31 βίντεο που έχουν καταγραφεί σε διαφορετικά περιβάλλοντα. Από αυτά τα βίντεο, 14 βίντεο περιλαμβάνουν πυρκαγιά, ενώ 17 είναι κανονικά βίντεο. Μια ποικιλία προκλήσεων, συμπεριλαμβανομένου του μεγαλύτερου μεγέθους σε σύγκριση με άλλα διαθέσιμα σύνολα δεδομένων, καθιστούν αυτό το σύνολο δεδομένων ιδιαίτερα κατάλληλο για αυτά τα πειράματα. Για παράδειγμα, μερικά από τα κανονικά βίντεο περιλαμβάνουν αντικείμενα που θυμίζουν πυρκαγιά. Αυτό καθιστά την ανίχνευση πυρκαγιάς πιο δύσκολη και, συνεπώς, οι μέθοδοι ανίχνευσης πυρκαγιάς που χρησιμοποιούν τις χρωματικές λειτουργίες ενδέχε-

ται να ταξινομήσουν εσφαλμένα αυτά τα πλαίσια. Επιπλέον, ένα σύνολο βίντεο συλλαμβάνεται σε ορεινές περιοχές και περιέχει σύννεφα και ομίχλη, για τα οποία τα συστήματα ανίχνευσης πυρκαγιάς που βασίζονται σε κίνηση ενδέχεται να μην λειτουργούν σωστά. Αυτές οι καταστάσεις μπορούν να εμφανιστούν στον πραγματικό κόσμο και, ως εκ τούτου, εισάγονται σε αυτό το σύνολο δεδομένων για να το κάνουν όσο πιο δύσκολο γίνεται. Τέλος, τα ποσοστά επιτυχίας της εργασίας αυτής ήταν **94.5%** και **91%** στο 1ο και 2ο dataset αντίστοιχα.

Κεφάλαιο 5

Βάσεις Δεδομένων

5.1 Corsican Database

Λαμβάνοντας υπόψη τι ήταν ενδιαφέρον στα αναφερόμενα σύνολα δεδομένων και τι έπρεπε να βελτιωθεί, αναπτύχθηκε η Corsican Fire Database (CFDB) [89]. Περιέχει εικόνες που έχουν ληφθεί στο ορατό και κοντινό υπέρυθρο φάσμα, ακολουθίες βίντεο, σχόλια και λεπτομέρειες σχετικά με τα χαρακτηριστικά της εικόνας, το περιβάλλον κλπ. Επιπλέον, κατασκευάστηκε για να είναι μια εξελισσόμενη βάση δεδομένων έτοιμη για εξωτερικές συνεισφορές.

5.1.1 Προέλευση των Εικόνων

Προκειμένου να δημιουργηθεί η βάση δεδομένων, υπήρξε μια κλήση για εικόνες φωτιάς και περισσότερες από 2000 εικόνες που καταγράφηκαν στα ορατά φάσματα συλλέχθηκαν από εταίρους και ερευνητές. Αυτές οι εικόνες προήλθαν από διαφορετικά μέρη του κόσμου, έχουν διαφορετικές μορφές και τραβήχτηκαν από κάμερες με διαφορετικές παραμέτρους. Στο σετ αυτό επιλέχθηκαν 500 εικόνες ώστε να έχουν ετερογενή χρώματα πυρκαγιάς και υφές, περιβάλλοντα, συνθήκες φωτισμού και βλάστηση. Για κάθε εικόνα, κατασκευάστηκε από ειδικό μια αναφορά φωτιάς με χειροκίνητη κατάτμησή της στην εικόνα. Αυτό το τμήμα της βάσης δεδομένων χρησιμοποιήθηκε στο [54].

Η βάση δεδομένων περιέχει επίσης 100 πολυτροπικές εικόνες φωτιάς και 5 σειρές πολλαπλών μεταδόσεων πυρκαγιάς σε διάδοση. Οι πολυτροπικές εικόνες λήφθηκαν χρησιμοποιώντας την κάμερα JAI AD-080GE. Αυτή η πολυφασματική κάμερα 2-CCD που βασίζεται σε πρίσμα αποκτά ταυτόχρονα μια εικόνα στα ορατά φάσματα και μια εικόνα στα φάσματα του εγγύς υπέρυθρου NIR (700nm - 900nm) μέσω του ίδιου οπτικού. Ένα παράδειγμα ορατών και κοντινών υπέρυθρων εικόνων που λαμβάνονται ταυτόχρονα φαίνεται στο Σχήμα 5.1. Οι πολυτροπικές εικόνες που λαμβάνονται απευθείας από τη φωτογραφική μηχανή JAI AD-080GE δεν είναι ευθυγραμμισμένες λόγω του γεγονότος ότι οι αισθητήρες ορατών και NIR δεν είναι ακριβώς συναρμολογημένοι. Μια καταγραφή εικόνας με βάση τον υπολογισμό μετασχηματισμού μήτρας ομογραφίας έγινε για να ευθυγραμμιστούν οι πολυτροπικές εικόνες που είναι διαθέσιμες στη βάση δεδομένων. Ο χρόνος λήψης της κάμερας επιλέχθηκε ανάλογα με τη φωτεινότητα περιβάλλοντος και το εστιακό μήκος ρυθμίστηκε στα 6mm. Οι αλληλουχίες διατηρήθηκαν με ρυθμό πλαισίου 1 fps. Οι εικόνες και των δύο φάσεων έχουν μέγεθος 1024 × 768 pixel. Όλες οι εικόνες της

βάσης δεδομένων βρίσκονται σε μορφή png χωρίς απώλειες.



Σχήμα 5.1: Παραδείγματα εικόνων τραβηγμένες με multi-spectral camera

5.1.2 Περιγραφητές των Εικόνων

Κάθε εικόνα της βάσης δεδομένων σχολιάζεται χρησιμοποιώντας αρκετούς περιγραφητές. Ορισμένοι από αυτούς σχολιάστηκαν χειροκίνητα και άλλοι αυτομάτως χρησιμοποιούσαν μια διαδικασία επεξεργασίας εικόνας. Οι περιγραφείς χωρίζονται σε δύο κύριες κατηγορίες, σε γενικούς περιγραφείς και σε περιγραφές πυρκαγιάς και περιβάλλοντος. Οι σχολιασμοί έχουν δύο σκοπούς:

- μπορούν να βοηθήσουν τους χρήστες στην ανάκτηση συγκεκριμένων ιδιοτήτων, για παράδειγμα, πυρκαγιές με συγκεκριμένο χρώμα ή υφή και
- επιτρέπει στον χρήστη να προσαρμόζει γρήγορα σεντ δοκιμών για συγκεκριμένους ερευνητικούς σκοπούς.

Για παράδειγμα, για τη δοκιμή ή την ανάπτυξη αλγορίθμων ανίχνευσης για συγκεκριμένες συνθήκες φωτιάς.

Καθολικοί Περιγραφητές

Οι *General Information* είναι διοικητικές περιγραφές, όπως ένα μοναδικό αναγνωριστικό, ο αριθμός ακολουθίας και ο αριθμός της εικόνας στην ακολουθία (τα οποία είναι μη μηδενικά αν η εικόνα ανήκει σε μια ακολουθία). Ένας άλλος σχολιασμός περιέχει πληροφορίες σχετικά με την τοποθεσία και την ημερομηνία. Οι τελευταίοι περιγραφείς αυτής της υποομάδας δίνουν τα ονόματα της ορατής εικόνας, της σχετικής υπέρυθρης εικόνας (αν υπάρχει) και της γroud τρυτη εικόνας.

Το πεδίο *Acquisition Settings* παρέχει πληροφορίες σχετικά με τον εξοπλισμό και οι ρυθμίσεις που χρησιμοποιούνται για τη λήψη των εικόνων: ο περιγραφέας <Material> δίνει το μοντέλο της φωτογραφικής μηχανής που χρησιμοποιείται, ο <Focal> υποδηλώνει το εστιακό μήκος της, το <Sensibility> παρουσιάζει την ευαισθησία ISO και το <Exposure> περιέχει τον χρόνο έκθεσης.

Εάν μια υπέρυθρη εικόνα συσχετίζεται με μια ορατή εικόνα, το πεδίο <Spectra> υποδεικνύει το φασματικό πεδίο: κοντά στην υπέρυθρη ακτινοβολία (NIR), υπέρυθρη μικρού μήκους κύματος (SWIR), υπέρυθρη μέσου μήκους κύματος (MWIR) ή μεγάλο μήκος κύματος υπέρυθρων (LWIR).

Περιγραφητές Φωτιάς και Περιβάλλοντος

Οι περιγραφητές αυτοί παρέχουν πληροφορίες για τις περιοχές/pixel της εικόνας όπου βρίσκεται η φωτιά και το υπόλοιπο φόντο αντίστοχα. Επίσης, δίνουν και άλλες πληροφορίες, όπως το εάν και πού υπάρχει καπνός στην εικόνα, την απόσταση της εστίας από το φακό, και τη φωτεινότητα του περιβάλλοντος.

Επιπλέον, αυτή η βάση δεδομένων περιέχει ακολουθίες βίντεο που συλλαμβάνονται ταυτόχρονα σε χρώματα και φάσματα NIR. Αυτές οι ακολουθίες εικόνων μπορούν να χρησιμεύσουν στη μελέτη αλγορίθμων πολυφασματικής σύντηξης, στην ανάλυση της απόδοσης τμηματοποίησης πυρκαγιάς σε αυτά τα φάσματα, στη χρήση κίνησης για κατακερματισμό πυρκαγιάς κλπ. Στο σημείο αυτό θα ήθελα να ευχαριστήσω την καθηγήτρια Lucile Rossi του Πανεπιστημίου της Κορσικής της Γαλλίας για την παροχή του dataset.

5.2 MIVIA Dataset

5.2.1 Φωτιά

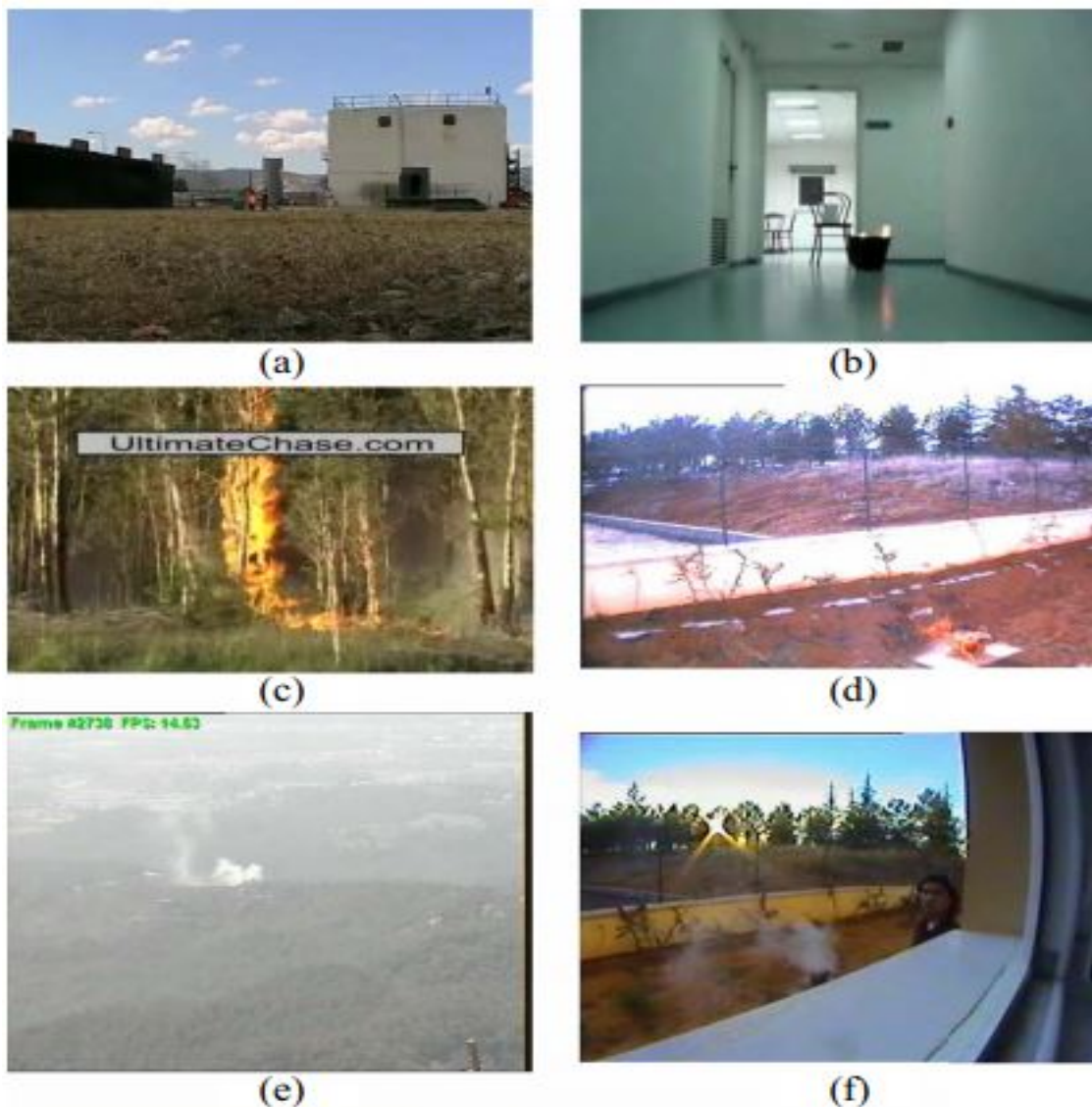
Αυτή η συλλογή βίντεο έχει αποκτηθεί τον Ιούλιο του 2012 στο Πανεπιστήμιο του Σαλέρνο της Ιταλίας και χρησιμοποιήθηκε για να ελεγχθούν οι μέθοδοι ανίχνευσης πυρκαγιάς και καπνού [88]. Αποτελείται από 31 βίντεο που αποκτήθηκαν σε πραγματικό περιβάλλον και κατεβάζτηκαν από το Internet. Το σύνολο δεδομένων μπορεί να θεωρηθεί ότι αποτελείται από δύο κύρια μέρη: τα πρώτα 14 βίντεο που χαρακτηρίζονται από την παρουσία της φωτιάς και τα τελευταία 17 βίντεο που δεν περιέχουν κανένα γεγονός ενδιαφέροντος. Συγκεκριμένα, αυτό το δεύτερο μέρος περιέχει κρίσιμες καταστάσεις που είναι δύσκολο να θεωρηθούν ως πυρκαγιά, όπως τα κόκκινα αντικείμενα που μετακινούνται στη σκηνή, καπνοί ή σύννεφα. Ένα δείγμα των δεδομένων αυτών φαίνεται στο παρακάτω Σχήμα 5.2:

5.2.2 Καπνός

Αυτό το σύνολο δεδομένων αποτελείται από 149 βίντεο, έκαστο διαρκεί περίπου 15 λεπτά, με αποτέλεσμα περισσότερες από 35 ώρες εγγραφής. Πρόκειται για ένα πολύ δύσκολο σύνολο δεδομένων, τόσο για τον έλεγχο των ανιχνευτών καπνού όσο και των ανιχνευτών πυρκαγιάς, καθώς περιέχει κόκκινα σπίτια σε μια μεγάλη κοιλάδα, βουνά στο ηλιοβασίλεμα, αντανάκλασεις του ήλιου στην κάμερα, και αρκετά σύννεφα, όπως φαίνεται και στο Σχήμα 5.3:

5.3 Λοιπά Datasets

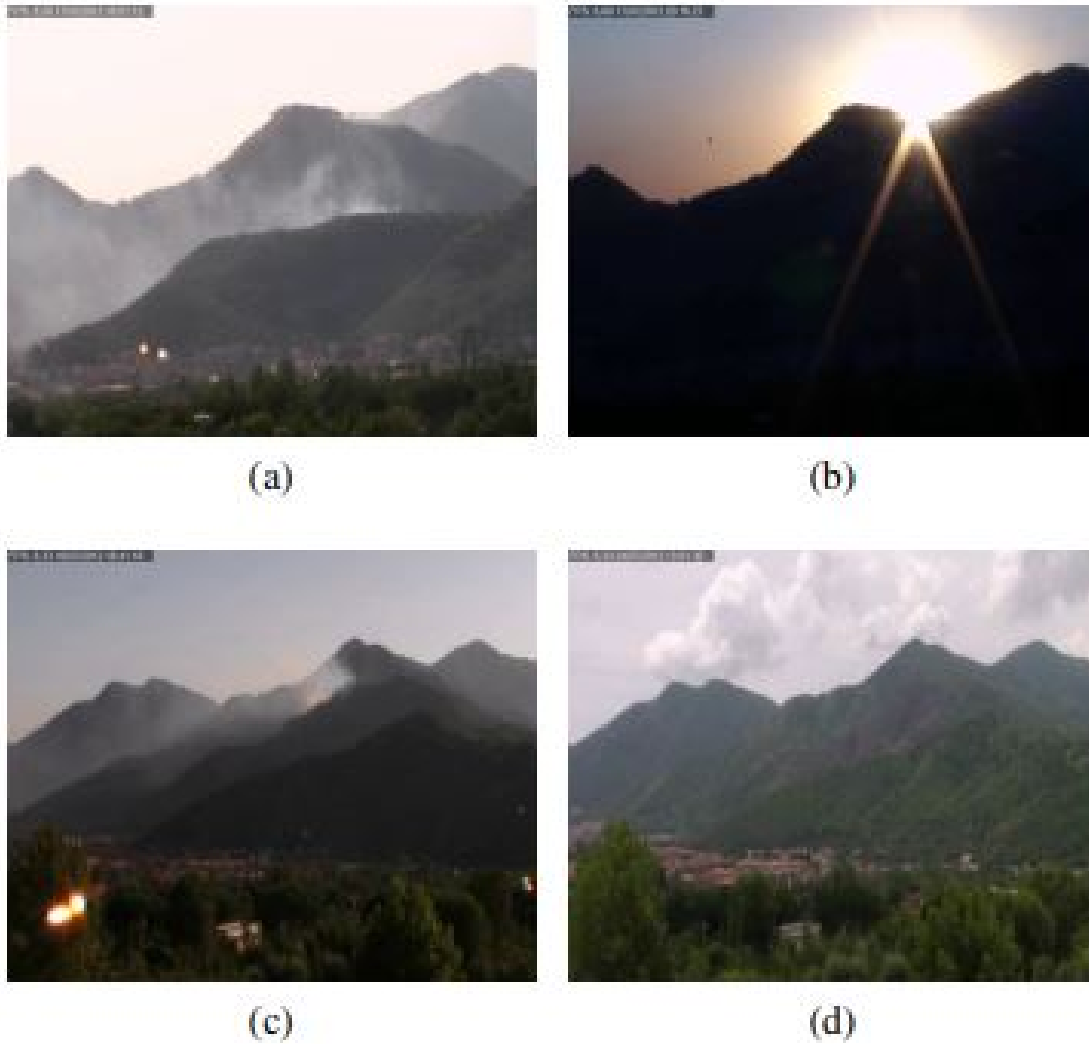
Εκτός από τα 2 παραπάνω datasets, τα οποία είναι και τα περισσότερο χρησιμοποιημένα στη βιβλιογραφία για αναγνώριση φωτιάς, υπάρχουν μερικά ακόμα που όμως είναι μικρά αλλά και αρκετά παλιά. Παρόλα αυτά, αποτέλεσαν τη βάση για τη δημιουργία των σύγχρονων βάσεων δεδομένων για αναγνώριση φωτιάς, και κυρίως του MIVIA dataset.



Σχήμα 5.2: Παραδείγματα βίντεο φωτιάς του MIVIA Dataset

Firesense

Το Firesense dataset είναι το πρώτο dataset που χρησιμοποιήθηκε στη βιβλιογραφία για αναγνώριση φωτιάς. Αποτελείται συνολικά από 264 αναλογικά βίντεο. Μερικά δείγματα για φωτιά και καπνό απεικονίζονται στα Σχήματα 5.4 και 5.5 αντίστοιχα.



Σχήμα 5.3: Παραδείγματα βίντεο καπνού του MIVIA Dataset



Σχήμα 5.4: Παραδείγματα βίντεο φωτιάς του Firesense dataset

Το dataset των Sharma et al. ([59])

Το συγκεκριμένο dataset, όπως αναφέρθηκε και στο Κεφάλαιο 4, δημιουργήθηκε από εικόνες που συλλέχθηκαν από το διαδίκτυο. Παρόλο που δεν περιέχει τόσο αντιπροσωπευτικές φωτογραφίες διασικής πυρκαγιάς, περιλαμβάνει όμως ικανοποιητικό αριθμό δειγμάτων τα οποία είναι δύσκολο να κατηγοριοποιηθούν, όπως ηλιοβασίλεμα, φώτα με διαφορετικές αποχρώσεις του κόκκινου και του κίτρινου.



Σχήμα 5.5: Παραδείγματα βίντεο καπνού του Firesense dataset

Κεφάλαιο 6

Προτεινόμενο Σύστημα

6.1 Χαρακτηριστικά Εικόνων

Μια εικόνα μπορεί να περιγραφεί από ένα σύνολο χαρακτηριστικών, δηλαδή τοποθεσίες ενδιαφέροντος (σημεία "κλειδιά") με τις περιγραφές τους. Τα χαρακτηριστικά μπορεί να αντιστοιχούν σε ακμές, γωνίες, υφές, χρώματα κ.α. Μια τυπική εφαρμογή είναι η αντιστοίχιση εικόνων συγκρίνοντας τα χαρακτηριστικά τους. Γι' αυτό το λόγο, ανάλογα με την εκάστοτε εργασία, τα χαρακτηριστικά πρέπει να έχουν ιδιότητες όπως γεωμετρική αμεταβλητότητα (μετακίνηση, περιστροφή, κλίμακα, μετασχηματισμός) και φωτομετρική αμεταβλητότητα (φωτεινότητα, χρώμα, έκθεση, σκιάσεις). Άλλες επιθυμητές ιδιότητες είναι η ευρωστία (σε θόρυβο, θόλωμα, συμπίεση) η ικανότητα διάκρισης (διαφορετικές περιγραφές για διαφορετικά σημεία ενδιαφέροντος) και η αποδοτικότητα (υπολογισμοί σε πραγματικό χρόνο). Στην περίπτωση που κάθε pixel της εικόνας θεωρείται σημείο κλειδί, το αποτέλεσμα είναι μια εικόνα χαρακτηριστικών. Ενώ στις συνήθεις εικόνες το βάθος κωδικοποιεί το χρώμα του pixel, στις εικόνες χαρακτηριστικών κωδικοποιεί μια πολυδιάστατη περιγραφή χαρακτηριστικού.

6.1.1 Είδη Χαρακτηριστικών

Οι αρχιτεκτονικές εξαγωγής χαρακτηριστικών συνθέτονται από δύο συστήματα: την ανίχνευση σημείων κλειδιών και η περιγραφή σημείων κλειδιών (χρησιμοποιώντας πληροφορία της εικόνας γύρω από αυτό το σημείο). Δύο ευρέως χρησιμοποιούμενα παραδείγματα χαρακτηριστικών είναι οι SIFT (Scale Invariant Feature Transform - Αμετάβλητος ως προς την Κλίμακα Μετασχηματισμός Χαρακτηριστικών) [90] και ORB (Oriented Fast and Rotated BRIEF) [91]. Είναι ιδιαίτερα σχεδιασμένες ώστε να ικανοποιούν αμεταβλητότητες και να εξασφαλίζουν εκτέλεση σε πραγματικό χρόνο.

SIFT

Οι SIFT περιγραφητές χαρακτηριστικών έχουν σχεδιαστεί ώστε να είναι αμετάβλητοι ως προς την κλίμακα, την περιστροφή, τους ομοπαράλληλους μετασχηματισμούς και τις φωτομετρικές αλλαγές. Οι SIFT περιγράφουν τα σημεία-κλειδιά με χρήση ενός 16×16 παραθύρου pixel γύρω τους, και γκαουσιανών πυραμίδων ώστε να επιλεγεί η κατάλληλη κλίμακα. Για κάθε pixel, υπολογίζονται οι κλίσεις (χωρικές παράγωγοι). Το παράθυρο αυτό χωρίζεται σε τέσσερα τεταρτημόρια και σε καθένα από αυτά υπολογίζεται το ιστόγραμμα προσανατολισμού κλίσης, με οχτώ διαστήματα. Το αποτέλεσμα αυτής της διαδικασίας είναι 128 μη αρνητικές τιμές οι

οποίες κανονικοποιούνται και ψαλιδίζονται ώστε να σχηματίσουν ένα 128-διάστατο περιγραφητή.

SURF

Οι SURF ταξινομητές είναι μια πιο γρήγορη εκδοχή των SIFT. Το SURF πηγαίνει λίγο περισσότερο και προσεγγίζει το LoG με το Box Filter. Επίσης, το SURF βασίζεται στον προσδιοριστή του Hessian matrix τόσο για την κλίμακα όσο και για την τοποθεσία. Για αντιστοίχιση προσανατολισμού, το SURF χρησιμοποιεί αποκρίσεις κύματος σε οριζόντια και κάθετη κατεύθυνση για μια γειτονιά μεγέθους $6s$. Επίσης, εφαρμόζονται επαρκή βάρη Gaussian. Ο κυρίαρχος προσανατολισμός υπολογίζεται με τον υπολογισμό του αθροίσματος όλων των αποκρίσεων μέσα σε ένα παράθυρο ολίσθησης με γωνία 60 μοίρες. Ενδιαφέρον είναι ότι, η απάντηση wavelet μπορεί να βρεθεί χρησιμοποιώντας ολοκληρωμένες εικόνες πολύ εύκολα σε οποιαδήποτε κλίμακα. Για την περιγραφή χαρακτηριστικών, το SURF χρησιμοποιεί αποκρίσεις Wavelet σε οριζόντια και κάθετη κατεύθυνση (και πάλι, η χρήση ολοκληρωμένων εικόνων διευκολύνει τα πράγματα). Μια γειτονιά μεγέθους $20s \times 20s$ λαμβάνεται γύρω από το σημείο κλειδί όπου s είναι το μέγεθος. Διαιρείται σε υποπεριφέρειες 4×4 . Για κάθε υποπεριοχή, λαμβάνονται οριζόντιες και κάθετες αποκρίσεις κύματος και ένα διάνυσμα. Με λίγα λόγια, το SURF προσθέτει πολλά χαρακτηριστικά για τη βελτίωση της ταχύτητας σε κάθε βήμα. Η ανάλυση δείχνει ότι είναι 3 φορές ταχύτερη από το SIFT, ενώ η απόδοση είναι συγκρίσιμη με το SIFT. Το SURF είναι καλό στο χειρισμό εικόνων με θόλωση και περιστροφή, αλλά δεν είναι καλό στο χειρισμό της αλλαγής απόψεων και της αλλαγής του φωτισμού.

ORB

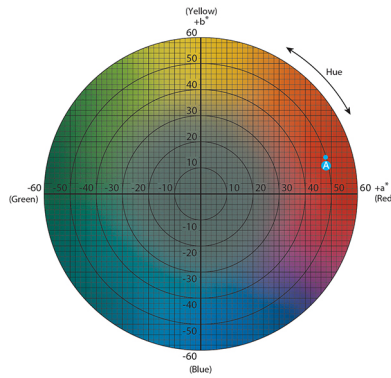
Ο ORB είναι ένα υπολογιστικά αποδοτικό, μη κατοχυρωμένο με δίπλωμα ευρεσιτεχνίας σύστημα εξαγωγής χαρακτηριστικών. Χρησιμοποιεί μια τροποποιημένη εκδοχή του FAST εξαγωγέα σημείων κλειδιών [92] ώστε να συνυπολογίζει περιστροφές, και μια ευαίσθητη στις περιστροφές υλοποίηση του BRIEF περιγραφητή [93].

LAB Histogram

Ο χρωματικός χώρος CIELAB (επίσης γνωστός ως CIE $L^*a^*b^*$ ή μερικές φορές συντετημημένος ως απλά "LAB") είναι ένας χρωματικός χώρος που εκφράζει χρώμα ως τρεις τιμές: L^* για τη φωτεινότητα από το μαύρο (0) στο λευκό (100), a^* από πράσινο (-) έως κόκκινο (+), και b^* από μπλε (-) έως κίτρινο (+), όπως φαίνεται και στο Σχήμα 6.1. Το CIELAB σχεδιάστηκε έτσι ώστε η ίδια ποσότητα αριθμητικής αλλαγής σε αυτές τις τιμές να αντιστοιχεί σε περίπου την ίδια ποσότητα οπτικά αντιληπτών αλλαγών.

6.2 Δημιουργία Βάσης

Για την κατασκευή του συστήματος μας, υλοποιήσαμε αρχικά τη βάση (baseline), η οποία είναι ουσιαστικά μία κλασική προσέγγιση της μηχανικής μάθησης με όραση υπολογιστών, όπου γίνεται εξαγωγή χαρακτηριστικών (feature extraction) με τη βοήθεια



Σχήμα 6.1: Χρωματικός χώρος LAB

των περιγραφητών που ορίστηκαν στην Ενότητα 6.1. Στη συνέχεια, τα χαρακτηριστικά εισάγονται σε μια μηχανή διανυσματικής υποστήριξης (SVM) όπου λαμβάνεται η τελική απόφαση για το εάν μία εικόνα περιέχει φωτιά ή όχι. Παρακάτω, περιγράφονται πιο αναλυτικά αυτά τα βήματα.

6.2.1 Εξαγωγή Χαρακτηριστικών

Για την εξαγωγή των χαρακτηριστικών βασιστήκαμε στο μοντέλο των Jiang et al. [39], το οποίο περιγράφηκε στην Ενότητα 4.1 παίρνοντας τον συνδυασμό που έδωσε το καλύτερο αποτέλεσμα, δηλαδή **SURF + LAB + SVM** με RBF kernel.

6.2.2 Περιγραφή του Dataset

Συλλέξαμε **600** εικόνες πυρκαγιάς που ποικίλλουν στο χρώμα, το σχήμα και την έντασή τους και **600** εικόνες χωρίς πυρκαγιά που περιλαμβάνουν μια ποικιλία απεικονίσεων, όπως δάση, αγροτικές περιοχές, σύννεφα και ουρανό, τα οποία έχουν απλό χρώμα και υφή. Τα δεδομένα αυτά προέρχονται από το CFDB αλλά και από το Google. Επιλέγουμε τυχαία το 75% από τα μπαλώματα για εκπαίδευση (train set) και 25% για αξιολόγηση (test set). Το πείραμα πραγματοποιήθηκε σε υπολογιστή με Intel Core i5 2.7GHz CPU και 8GB RAM. Παραδείγματα των εικόνων του dataset φαίνονται στο Σχήμα 6.2 παρακάτω:

Το καθολικό χρωματικό ιστόγραμμα (global color histogram) διαφορετικού χώρου χρώματος αποτελείται από 96 bins, που είναι η συνένωση τριών ιστογραμμάτων 32-bin για κάθε κανάλι. Οι περιγραφείς SURF υπολογίζονται πάνω από τα βασικά σημεία που ανιχνεύονται από τον Fast Hessian Detector, αποδίδοντας 64-διάστατα διανύσματα (vectors). Επειδή η ποιότητα και η ευκρίνεια των εικόνων είναι σχετικά χαμηλή, το κατώφλι του ανιχνευτή έχει οριστεί σε 100 για να ανιχνεύσει περισσότερα σημεία-κλειδιά. Τα τοπικά χρωματικά ιστογράμματα (local color histograms) υπολογίζονται εντός του πλαισίου κάθε πυρήνα SURF. Το μέγεθος του είναι 24 bin, με 8 bins για κάθε κανάλι. Συνδυάζονται για να σχηματίσουν ένα 88-διάστατο διάνυσμα χαρακτηριστικών, το οποίο στη συνέχεια εισέρχεται ως είσοδος στον SVM. Το μοντέλο υλοποιήθηκε με χρήση του εργαλείου scikit-learn.



Σχήμα 6.2: Παραδείγματα εικόνων με και χωρίς πυρκαγιά του dataset για τη λύση του συστήματος βάσης

6.3 Προσέγγιση με ΣΝΔ

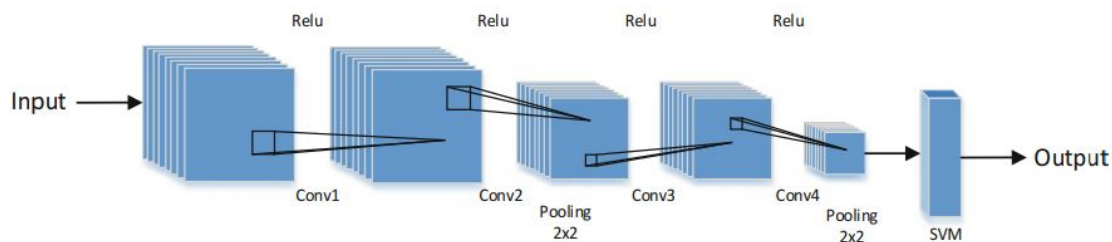
Το επόμενο βήμα στην υλοποίηση του συστήματος ανίχνευσης πυρκαγιάς είναι η χρήση βαθιού συνελικτικού νευρωνικού δικτύου (CNN), το οποίο κάνει feature extraction απευθείας από τα δεδομένα/εικόνες (raw data). Η διαδικασία αυτή συνοψίζεται ως εξής: Πρώτον, γίνεται εξαγωγή χαρακτηριστικών από το CNN. Δεύτερον, τα αποτελέσματα της εξαγωγής ταξινομούνται ως θετικά δείγματα και αρνητικά δείγματα του CNN. Στη συνέχεια, πραγματοποιείται εκπαίδευση (training) του CNN και αποθήκευση του μοντέλου. Τρίτον, αντικαθιστούμε το πλήρως συνδεδεμένο στρώμα CNN από ένα SVM και συνεχίζουμε την εκπαίδευση του σετ δειγμάτων. Τέλος, δοκιμάζουμε το σετ αξιολόγησης (test set) με το εκπαιδευμένο μοντέλο για να λάβουμε στην έξοδο το αποτέλεσμα της τελικής ταξινόμησης της εικόνας, εάν δηλαδή περιέχει ή όχι φωτιά.

6.3.1 Αρχιτεκτονική του Συστήματος

Πρώτα απ'όλα, δίνεται ένας μεγάλος αριθμός εικόνων το μέγεθος των οποίων έχει μειωθεί στο 32×32 και το μέγεθος πυρήνα συνέλιξης των τεσσάρων συνελικτικών στρωμάτων (convolutional layers) είναι 3×3 . Μετά την εκπαίδευση του μοντέλου CNN, χρησιμοποιούμε το SVM, όπου τα δεδομένα εκπαίδευσης είναι ο φορέας της παραγωγής του τέταρτου συνελικτικού στρώματος από το μοντέλο, για ταξινόμηση αντί για το πλήρως συνδεδεμένο στρώμα του CNN.

Η παραπάνω μέθοδος διατηρεί το εκπαιδευμένο μοντέλο CNN, έτσι ώστε η έξοδος του τελευταίου στρώματος συνέλιξης να μπορεί να ληφθεί απευθείας από το μοντέλο CNN. Η διαδικασία εκπαίδευσης του CNN είναι ουσιαστικά ίδια με τη διαδικασία ρύθμισης των παραμέτρων, έτσι ώστε να έχουμε ένα καλύτερο αποτέλεσμα καθώς οι παράμετροι του στρώματος συνέλιξης βελτιώνονται περισσότερο. Τα πειράματα δείχνουν ότι αν τα αποτελέσματα του τελευταίου στρώματος συνέλιξης στη διαδικασία εκμάθησης χρησιμοποιούνται ως είσοδοι στο SVM, τα αποτελέσματα ταξινόμησης θα είναι πολύ μικρότερα από το παραπάνω μοντέλο. Έτσι, προτείνουμε μια αρχιτεκτονική CNN-SVM, όπως φαίνεται στο Σχήμα 6.3. Η διαδικασία είναι η εξής:

Πρώτον, έχουμε την τελευταία έξοδο στρώματος συνέλιξης από τη συντήρηση του μοντέλου CNN ως είσοδο ταξινομητή SVM. Ο πυρήνας λειτουργίας της ακτινικής βάσης (RBF) χαρτογραφεί ένα δείγμα σε υψηλότερη διάσταση. Σε σύγκριση με τον πολυωνυμικό πυρήνα, ο RBF πρέπει να καθορίσει τις παραμέτρους να είναι μικρότερες, αφού ο αριθμός των παραμέτρων λειτουργίας του πυρήνα επηρεάζουν άμεσα την πολυπλοκότητα του συστήματος. Έτσι, επιλέγεται ο ταξινομητής SVM με τη λειτουργία του πυρήνα ως RBF. Στη συνέχεια, πραγματοποιείται η εκπαίδευση του SVM και στη συνέχεια η αποθήκευση του μοντέλου, για την ανίχνευση της πυρκαγιάς.



Σχήμα 6.3: Η αρχιτεκτονική του δικτύου CNN-SVM

6.3.2 Περιγραφή του Πειράματος

Όπως περιγράφηκε προηγουμένως, το σύστημα αποτελείται από 4 στρώματα συνέλιξης μεγέθους 64, 128, 256 και 512 αντίστοιχα, με μέγεθος πυρήνα 3×3 και Max Pooling 2×2 . Η συνάρτηση ενεργοποίησης είναι η ReLU. Επίσης, χρησιμοποιήσαμε ρυθμό εκπαίδευσης (learning rate) ίσο με 0.001, momentum = 0.9, batch size = 64 και cross-entropy loss. Ο optimizer μας είναι ο Adam.

Όπως θα φανεί και στην Ενότητα 6.6, το συγκεκριμένο CNN χρησιμοποιήθηκε τόσο για το κομμάτι της αναγνώρισης της φωτιάς, όσο και για αυτό του καπνού. Συνολικά δηλαδή έγιναν δύο πειράματα.

Για το δίκτυο αναγνώρισης πυρκαγιάς, κάναμε χρήση όλων των datasets που παρουσιάστηκαν στο Κεφάλαιο 5. Ειδικότερα, συγκεντρώσαμε συνολικά 10319 εικόνες, από τις οποίες οι 5061 περιέχουν φωτιά, ενώ οι υπόλοιπες 5258 όχι. Αντίστοιχα, για το σύστημα αναγνώρισης καπνού χρησιμοποιήθηκε κυρίως το MIVIA dataset, αφού είναι το πιο αντιπροσωπευτικό για την περίπτωση του καπνού. Τα δεδομένα μας είναι 8915 εικόνες με καπνό και 9752 χωρίς, συνολικά 18667 φωτογραφίες. Και τα δύο σύνολα δεδομένων διασπάτηκαν ως εξής: 60% για train, 20% για validation και 20% για test. Η εκπαίδευση έγινε σε 10 εποχές, ενώ και τα δύο μοντέλα υλοποιήθηκαν με χρήση του Keras (σε Tensorflow backend).

6.4 Real-Time

Όσον αφορά την κατασκευή του συστήματος το οποίο θα λειτουργεί σε πραγματικό χρόνο, χρησιμοποιούμε Μεταφορά Μάθησης (Transfer Learning). Πιο συγκεκριμένα, εκπαιδεύτηκε το pre-trained δίκτυο Inception-V1 σε 20593 εικόνες συνολικά (προστέθηκαν και εικόνες από το [94]), πυρκαγιάς και μη. Η αρχιτεκτονική του δικτύου φαίνεται στο Σχήμα 6.4.

Scikit-Learn

Η Scikit-learn (ή sklearn) είναι επίσης μια βιβλιοθήκη μηχανικής μάθησης ανοιχτού κώδικα της γλώσσας Python. Διαθέτει αλγόριθμους για ταξινόμηση (classification), παλινδρόμηση (regression), μηχανές διανυσμάτων υποστήριξης (SVM), αλλά και αλγόριθμους μη επιβλεπόμενης μάθησης όπως k -means και DBSCAN. Έχει σχεδιαστεί για να συνεργάζεται με τις αριθμητικές και επιστημονικές βιβλιοθήκες της Python, NumPy και SciPy.

NumPy

Η NumPy, μια ακόμα βιβλιοθήκη ανοιχτού κώδικα, είναι το θεμελιώδες πακέτο για επιστημονικούς υπολογισμούς με Python. Κάνει εφικτούς υπολογισμούς με πολυδιάστατους πίνακες χρησιμοποιώντας στο υπόβαθρο αποδοτικές ρουτίνες γραμμένες σε C. Η Numpy χρησιμοποιήθηκε σε όλη την έκταση του προγράμματος επιταχύνοντας ιδιαίτερα κατά τη δημιουργία της βάσης μέσω του sklearn.

SciPy

Η ανοιχτού κώδικα βιβλιοθήκη SciPy [100] περιέχει μονάδες για βελτιστοποίηση, γραμμική άλγεβρα και άλλα εργαλεία επιστημονικών υπολογισμών. Έχει γραφτεί με χρήση NumPy. Η συνάρτηση επίλυσης του προβλήματος γραμμικής αντιστοίχισης του SciPy προσέφερε μια αποδοτική υλοποίηση του Ουγγρικού Αλγορίθμου [101].

6.6 Αποτελέσματα

Όπως προκύπτουν από τα αποτελέσματα του Πίνακα 6.1, το αρχικό μας σύστημα (baseline) με τους εξαγωγείς χαρακτηριστικών και τη μηχανή SVM, δίνει σκορ κοντά στο state-of-the-art [39].

Model	Accuracy
SURF+LAB+SVM	85.87%
Jiang et al.	92.26%

Πίνακας 6.1: Σύγκριση των αποτελεσμάτων του προτεινόμενου συστήματος βάσης με αυτό των Jiang et al.

Στη συνέχεια, παραθέτουμε στον επόμενο Πίνακα 6.2 τα ποσοστά ακρίβειας των δύο αρχιτεκτονικών συνελικτικών νευρωνικών δικτύων για φωτιά και καπνό. Παρατηρούμε ότι το σκορ είναι σαφώς υψηλότερο σε σχέση με την κλασική μέθοδο του baseline, και επίσης υπερνικά αρκετές εργασίες της βιβλιογραφίας. Επιπλέον, πετυχαίνεται πολύ χαμηλός δείκτης εσφαλμένου συναγερμού (false alarm rate - FAR).

Τέλος, το real-time μοντέλο δίνει accuracy **93%**, ενώ ο δείκτης FAR είναι ίσος με **0.17**. Στα Σχήματα 6.5, 6.6 και 6.7 παρακάτω, παρουσιάζουμε μερικά παραδείγματα από την εφαρμογή του μοντέλου αυτού σε βίντεο πυρκαγιών που συλλέχθηκαν από το

YouTube, όπου η κόκκινη ένδειξη του πλαισίου σημαίνει ότι υπάρχει φωτιά (FIRE), ενώ η πράσινη ότι η περιοχή είναι καθαρή (CLEAR).

Model	Accuracy	FAR
CNN Fire	97.51%	2.03%
CNN Smoke	98.13%	3.26%

Πίνακας 6.2: Result metrics για τα δύο CNN



Σχήμα 6.5: Παράδειγμα εξόδου του συστήματος ανίχνευσης πυρκαγιάς σε πραγματικό χρόνο



Σχήμα 6.6: Παράδειγμα εξόδου του συστήματος ανίχνευσης πυρκαγιάς σε πραγματικό χρόνο



Σχήμα 6.7: Παράδειγμα εξόδου του συστήματος ανίχνευσης πυρκαγιάς σε πραγματικό χρόνο

Κεφάλαιο 7

Εφαρμογή

Στο παρόν Κεφάλαιο παρουσιάζονται μέθοδοι και τεχνικές με τις οποίες μπορεί το σύστημα που περιγράφηκε στις προηγούμενες ενότητες να αποτελέσει μία πραγματική εφαρμογή μέσω ενός αυτόματου συστήματος παρακολούθησης για την αποφυγή φαινομένων δασικής (και όχι μόνο) πυρκαγιάς.

7.1 Κάμερες

Φυσικά, για να μπορέσουμε να παρατηρούμε συνεχώς μια μεγάλη περιοχή, θα χρειαστούν πολλές κάμερες. Επιπλέον, αυτές οι κάμερες πρέπει να αποθηκεύσουν τα δεδομένα που λαμβάνουν κάπου, είτε τοπικά είτε σε απομακρυσμένη τοποθεσία. Ένα βίντεο υψηλότερης ποιότητας θα πάρει πολύ περισσότερη μνήμη από μια κατώτερη ποιότητα. Επιπλέον, μια ροή εισόδου RGB είναι 3 φορές μεγαλύτερη από μια ασπρόμαυρη ροή εισόδου. Δεδομένου ότι μπορούμε να αποθηκεύσουμε μόνο ένα πεπερασμένο ποσό της ροής εισόδου, η ποιότητα συχνά μειώνεται για να μεγιστοποιηθεί η αποθήκευση. Επομένως, ένα κλιμακωτό σύστημα παρακολούθησης θα πρέπει να μπορεί να ερμηνεύει εικόνες χαμηλής ποιότητας. Ως εκ τούτου, ο αλγόριθμος Deep Learning πρέπει να εκπαιδεύεται και σε εικόνες χαμηλής ποιότητας, όπως δηλαδή και έγινε (βλ. Κεφάλαιο 6).

7.2 Επεξεργασία Δεδομένων

Τώρα που επιλύσαμε τον περιορισμό της συλλογής των δεδομένων, μπορούμε να απαντήσουμε σε μια μεγαλύτερη ερώτηση. Από πού επεξεργαζόμαστε τα δεδομένα που λαμβάνονται από τις κάμερες. Υπάρχουν δύο μέθοδοι για να γίνει αυτό:

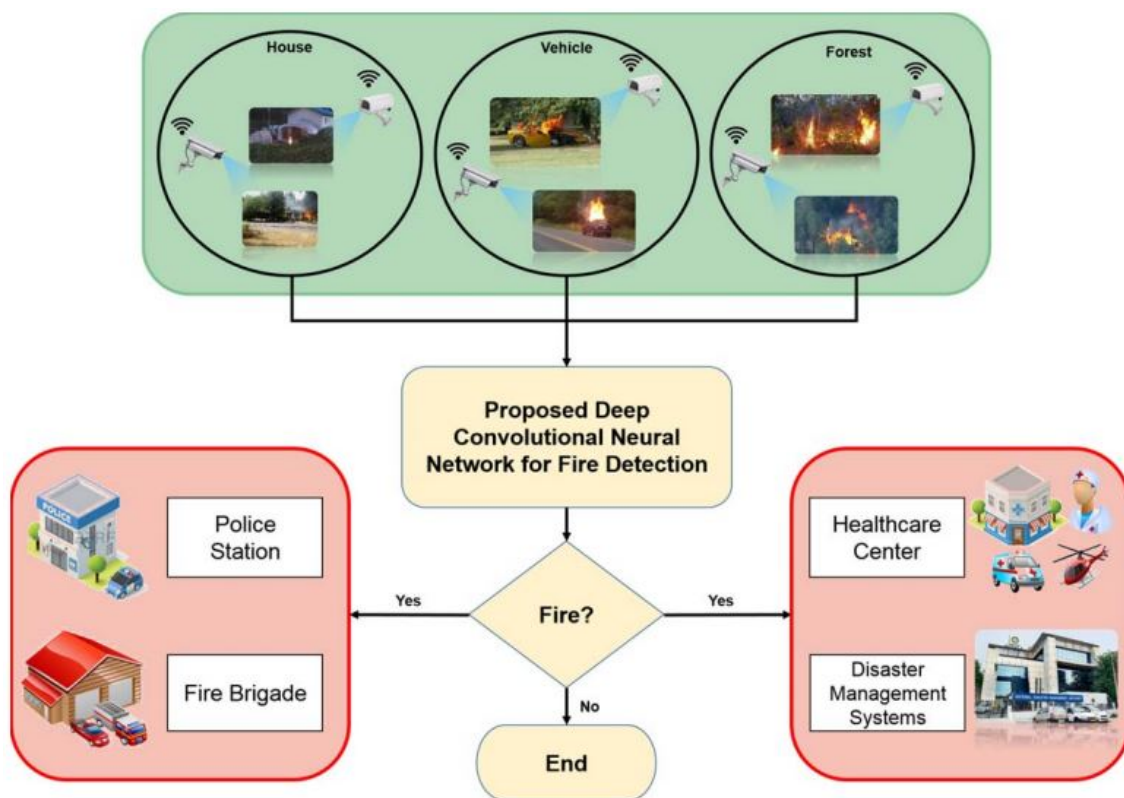
7.2.1 Επεξεργασία σε Κεντρικό Server

Οι ροές βίντεο από τις κάμερες επεξεργάζονται ανά καρέ σε έναν απομακρυσμένο διακομιστή ή έναν server. Αυτή η μέθοδος είναι ισχυρή και μάς επιτρέπει να αποκομίζουμε τα οφέλη των πολύπλοκων μοντέλων με υψηλή ακρίβεια. Το προφανές πρόβλημα είναι η λανθάνουσα κατάσταση. Απαιτείται μια γρήγορη σύνδεση στο Internet για περιορισμένη καθυστέρηση του βίντεο (fps). Επιπλέον, αν δεν χρησιμοποιείται ένα εμπορικό API, το κόστος εγκατάστασης και συντήρησης του διακομιστή μπορεί να είναι υψηλό.

7.2.2 Επεξεργασίας στο Νέφος

Η επεξεργασία στο νέφος (processing on the edge) μπορεί να γίνει τοποθετώντας ένα μικρό μικροελεγκτή, όπου μπορούμε να λάβουμε συμπεράσματα σε πραγματικό χρόνο για την ίδια την κάμερα. Δεν υπάρχει καθυστέρηση μετάδοσης και οι ανωμαλίες μπορούν να περιοριστούν ταχύτερα από την προηγούμενη μέθοδο. Επιπλέον, αυτό είναι ένα σημαντικό πλεονέκτημα για τις συσκευές που είναι κινητά, έτσι ώστε να μη χρειάζεται να περιοριστούν από το φάσμα των διαθέσιμων WiFi/Bluetooth. Για το λόγο αυτό είναι προτιμότερη η χρήση αυτής της μεθόδου.

Αυτός είναι και ο σκοπός της συγκεκριμένης διατριβής, η ενσωμάτωση δηλαδή του προτεινόμενου συστήματος με χρήση βαθιάς μάθησης σε εφαρμογή. Αυτό μπορεί να γίνει με χρήση πλατφόρμων όπως το Raspberry Pi [102]. Πιο συγκεκριμένα, το μοντέλο μπορεί να "τρέχει" στο raspberry το οποίο θα είναι συνδεδεμένο με κάμερα (π.χ. GoPro) και σε περίπτωση που ανιχνεύσει μια πιθανή εστία πυρκαγιάς, θα στέλνεται σήμα (alarm) στους αρμόδιους φορείς (π.χ. πυροσβεστική). Ένας ακόμα τρόπος είναι με χρήση drones [103]. Συμπερασματικά, ένα συνολικό σύστημα αναγνώρισης φωτιάς μπορεί να συμπεριληφθεί στο ακόλουθο Σχήμα 7.1:



Σχήμα 7.1: Προτεινόμενο σύστημα παρακολούθησης και ανίχνευσης φωτιάς

Κεφάλαιο 8

Συμπεράσματα

8.1 Αξιολόγηση

Η παρούσα Διπλωματική Εργασία περιγράφει την κατασκευή ενός συστήματος αναγνώρισης πιθανής εστίας πυρκαγιάς σε δασικό κυρίως περιβάλλον. Για το σκοπό αυτό, ακολουθήσαμε ποικίλες μεθόδους και τεχνικές. Ξεκινώντας από ένα απλό σύστημα εξαγωγής χαρακτηριστικών με γραμμικό ταξινομητή SVM, δείξαμε ότι ακόμα και μία τόσο "φτωχή" αρχιτεκτονική μπορεί να επιτύχει αξιοσημείωτα αποτελέσματα, της τάξης του 86%, νούμερο το οποίο είναι αρκετά κοντά σε αυτά άλλων πιο περίπλοκων state-of-the-art τεχνικών. Στη συνέχεια, το επόμενο βήμα ήταν η υλοποίηση ενός βαθιού συνελικτικού νευρωνικού δικτύου, όπου οι χάρτες χαρακτηριστικών από τα βαθιά επίπεδα του δικτύου ανίχνευσης κωδικοποιούν την εμφάνιση της φωτιάς ή του καπνού. Τα σκορ του βαθιού ΣΝΔ τόσο για φωτιά όσο και για καπνό είναι άκρως ενθαρρυντικά, φτάνοντας στο 98% και υπερνικώντας πολλές παρόμοιες τεχνικές στη βιβλιογραφία της ανίχνευσης πυρκαγιάς. Τέλος, δοκιμάσαμε το σύγχρονο προεκπαιδευμένο νευρωνικό δίκτυο GoogleNet, γνωστό και ως Inception για τη δημιουργία ενός συστήματος το οποίο θα μπορεί να λειτουργεί σε πραγματικό χρόνο.

8.2 Μελλοντικές Επεκτάσεις

Υπάρχουν αρκετές μελλοντικές κατευθύνσεις για βελτίωση του συστήματος που προτείνεται. Μία μέθοδος που θα μπορούσε να εφαρμοστεί στο μέλλον είναι η δοκιμή και άλλων σύγχρονων state-of-the-art αρχιτεκτονικών νευρωνικών δικτύων, όπως τα AlexNet, VGG και ResNet. Επιπλέον, η εφαρμογή του προτεινόμενου συστήματος της παρούσας διατριβής σε πραγματικές εγκαταστάσεις, όπως η τοποθέτηση Raspberry Pi συνδεδεμένα με κάμερα σε πυλώνες του ΟΤΕ σε επιλεγμένα σημεία της Ελλάδας όπου υπάρχει υψηλή επικινδυνότητα για εκδήλωση πυρκαγιάς, είναι ένας πολύ καλός τρόπος αξιολόγησης. Επίσης, το σύστημα θα μπορούσε να εκπαιδευτεί κυρίως με εικόνες καπνού, αφού είναι αναγκαία ο έγκαιρος εντοπισμός της φωτιάς. Άλλωστε, όπως αναφέραμε κι από την αρχή της εργασίας, ο σκοπός είναι η μείωση των φαινομένων δασικής πυρκαγιάς και να μην υπάρχει απειλή της ανθρώπινης ζωής.

Βιβλιογραφία

- [1] B Ugur Toreyin and A Enis Cetin. «Online detection of fire in video». In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 2007, pp. 1–5.
- [2] Byoung Chul Ko, Kwang-Ho Cheong, and Jae-Yeal Nam. «Fire detection based on vision sensor and support vector machines». In: *Fire Safety Journal* 44.3 (2009), pp. 322–329.
- [3] Paulo Vinicius Koerich Borges and Ebroul Izquierdo. «A probabilistic approach for vision-based fire detection in videos». In: *IEEE transactions on circuits and systems for video technology* 20.5 (2010), pp. 721–731.
- [4] Ren C Luo and Kuo L Su. «Autonomous fire-detection system using adaptive sensory fusion for intelligent security robot». In: *Ieee/Asme Transactions on Mechatronics* 12.3 (2007), pp. 274–281.
- [5] Primož Podržaj and Hideki Hashimoto. «Intelligent space as a framework for fire detection and evacuation». In: *Fire technology* 44.1 (2008), pp. 65–76.
- [6] Glenn Healey et al. «A system for real-time fire detection». In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 1993, pp. 605–606.
- [7] Che-Bin Liu and Narendra Ahuja. «Vision based fire detection». In: *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004*. Vol. 4. IEEE. 2004, pp. 134–137.
- [8] Klaus Schmidt-Rohr. «Why combustions are always exothermic, yielding about 418 kJ per mole of O₂». In: *Journal of Chemical Education* 92.12 (2015), pp. 2094–2099.
- [9] Wikipedia contributors. *Fire — Wikipedia, The Free Encyclopedia*. [Online; accessed 29-May-2019]. 2019. URL: <https://en.wikipedia.org/w/index.php?title=Fire&oldid=897870787>.
- [10] William J Bond and Jon E Keeley. «Fire as a global ‘herbivore’: the ecology and evolution of flammable ecosystems». In: *Trends in ecology & evolution* 20.7 (2005), pp. 387–394.
- [11] Philippe Julien et al. «Freely-propagating flames in aluminum dust clouds». In: *Combustion and Flame* 162.11 (2015), pp. 4241–4253.
- [12] Jesús San-Miguel-Ayanz et al. «Comprehensive monitoring of wildfires in Europe: the European forest fire information system (EFFIS)». In: *Approaches to Managing Disaster-Assessing Hazards, Emergencies and Disaster Impacts*. IntechOpen, 2012.

- [13] Sameer Karki. *Community involvement in and management of forest fires in South East Asia*. Citeseer, 2002.
- [14] Javier Solobera. «Detecting forest fires using wireless sensor networks». In: *Libelium World* (2010).
- [15] E Thomson. «Preventing forest fires with tree power». In: *MIT Tech Talk* (2008).
- [16] Wikipedia contributors. *Wildfire — Wikipedia, The Free Encyclopedia*. [Online; accessed 29-May-2019]. 2019. URL: <https://en.wikipedia.org/w/index.php?title=Wildfire&oldid=897216351>.
- [17] Olga Raggio. «The myth of Prometheus: Its survival and metamorphoses up to the eighteenth century». In: *Journal of the Warburg and Courtauld Institutes* 21.1/2 (1958), pp. 44–62.
- [18] Wikipedia contributors. *List of natural disasters by death toll — Wikipedia, The Free Encyclopedia*. [Online; accessed 2-June-2019]. 2019. URL: https://en.wikipedia.org/w/index.php?title=List_of_natural_disasters_by_death_toll&oldid=897376796.
- [19] Christopher M Bishop. *Bishop pattern recognition and machine learning*. 2001.
- [20] Vadim Kantorov et al. «Contextlocnet: Context-aware deep network models for weakly supervised localization». In: *European Conference on Computer Vision*. Springer. 2016, pp. 350–365.
- [21] Ross Girshick et al. «Region-based convolutional networks for accurate object detection and segmentation». In: *IEEE transactions on pattern analysis and machine intelligence* 38.1 (2015), pp. 142–158.
- [22] Wenlu Zhang et al. «Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation». In: *NeuroImage* 108 (2015), pp. 214–224.
- [23] Jamil Ahmad et al. «Object-oriented convolutional features for fine-grained image retrieval in large surveillance datasets». In: *Future Generation Computer Systems* 81 (2018), pp. 314–330.
- [24] Yann LeCun et al. «Gradient-based learning applied to document recognition». In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324.
- [25] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. «Imagenet classification with deep convolutional neural networks». In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.
- [26] John Nickolls. «Scalable parallel programming with CUDA introduction». In: *2008 IEEE Hot Chips 20 Symposium (HCS)*. IEEE. 2008, pp. 1–9.
- [27] Matthew D. Zeiler and Rob Fergus. «Visualizing and Understanding Convolutional Networks». In: *CoRR* abs/1311.2901 (2013). arXiv: 1311.2901. URL: <http://arxiv.org/abs/1311.2901>.
- [28] Christian Szegedy et al. «Going Deeper with Convolutions». In: (2014). arXiv: 1409.4842. URL: <http://arxiv.org/abs/1409.4842>.

- [29] Christian Szegedy, Sergey Ioffe, and Vincent Vanhoucke. «Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning». In: *CoRR* abs/1602.07261 (2016). arXiv: 1602.07261. URL: <http://arxiv.org/abs/1602.07261>.
- [30] K. Simonyan and A. Zisserman. «Very Deep Convolutional Networks for Large-Scale Image Recognition». In: *CoRR* abs/1409.1556 (2014).
- [31] Kaiming He et al. «Deep residual learning for image recognition». In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.
- [32] Kaiming He et al. «Identity Mappings in Deep Residual Networks». In: *CoRR* abs/1603.05027 (2016). arXiv: 1603.05027. URL: <http://arxiv.org/abs/1603.05027>.
- [33] Lisa Torrey and Jude Shavlik. «Transfer learning». In: *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. IGI Global, 2010, pp. 242–264.
- [34] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.
- [35] Qingjie Zhang et al. «Deep convolutional neural networks for forest fire detection». In: *2016 International Forum on Management, Education and Information Technology Application*. Atlantis Press. 2016.
- [36] N Dedeoglu et al. «Real-time fire and flame detection in video». In: *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005*. Vol. 2. IEEE. 2005, pp. ii–669.
- [37] Yangqing Jia et al. «Caffe: Convolutional architecture for fast feature embedding». In: *Proceedings of the 22nd ACM international conference on Multimedia*. ACM. 2014, pp. 675–678.
- [38] Jonathan Long, Evan Shelhamer, and Trevor Darrell. «Fully convolutional networks for semantic segmentation». In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 3431–3440.
- [39] Bo Jiang et al. «Towards a solid solution of real-time fire and flame detection». In: *Multimedia Tools and Applications* 74.3 (2015), pp. 689–705.
- [40] David G Lowe et al. «Object recognition from local scale-invariant features.» In: *iccv*. Vol. 99. 2. 1999, pp. 1150–1157.
- [41] Navneet Dalal and Bill Triggs. «Histograms of oriented gradients for human detection». In: *international Conference on computer vision & Pattern Recognition (CVPR'05)*. Vol. 1. IEEE Computer Society. 2005, pp. 886–893.
- [42] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. «Surf: Speeded up robust features». In: *European conference on computer vision*. Springer. 2006, pp. 404–417.
- [43] Sebastien Frizzi et al. «Convolutional neural network for video fire and smoke detection». In: *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*. IEEE. 2016, pp. 877–882.
- [44] Bing Xu et al. «Empirical evaluation of rectified activations in convolutional network». In: *arXiv preprint arXiv:1505.00853* (2015).

- [45] Andrew L Maas, Awni Y Hannun, and Andrew Y Ng. «Rectifier nonlinearities improve neural network acoustic models». In: *Proc. icml*. Vol. 30. 1. 2013, p. 3.
- [46] Frédéric Bastien et al. «Theano: new features and speed improvements». In: *arXiv preprint arXiv:1211.5590* (2012).
- [47] James Bergstra et al. «Theano: a CPU and GPU math expression compiler». In: *Proceedings of the Python for scientific computing conference (SciPy)*. Vol. 4. 3. Austin, TX. 2010.
- [48] Steven Verstockt et al. «State of the art in vision-based fire and smoke detection». In: *14th International Conference on Automatic Fire Detection*. Vol. 2. University of Duisburg-Essen. Department of Communication Systems. 2009, pp. 285–292.
- [49] Daniel YT Chino et al. «Bowfire: detection of fire in still images by integrating pixel color and texture analysis». In: *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images*. IEEE. 2015, pp. 95–102.
- [50] Kumarguru Poobalan and Siau-Chuin Liew. «Fire detection algorithm using image processing techniques». In: *Proceedings of the 3rd International Conference on Artificial Intelligence and Computer Science (AICS2015)*. 2015, pp. 160–168.
- [51] Jing Shao, Guanxiang Wang, and Wei Guo. «An image-based fire detection method using color analysis». In: *2012 International Conference on Computer Science and Information Processing (CSIP)*. IEEE. 2012, pp. 1008–1011.
- [52] B Uğur Töreyn et al. «Computer vision based method for real-time fire and flame detection». In: *Pattern recognition letters* 27.1 (2006), pp. 49–58.
- [53] Jerome Vicente and Philippe Guillemant. «An image processing technique for automatically detecting forest fire». In: *International Journal of Thermal Sciences* 41.12 (2002), pp. 1113–1120.
- [54] Tom Toulouse et al. «Automatic fire pixel detection using image processing: a comparative analysis of rule-based and machine learning-based methods». In: *Signal, Image and Video Processing* 10.4 (2016), pp. 647–654.
- [55] Olga Russakovsky et al. «Imagenet large scale visual recognition challenge». In: *International journal of computer vision* 115.3 (2015), pp. 211–252.
- [56] Chongyuan Tao, Jian Zhang, and Pan Wang. «Smoke detection based on deep convolutional neural networks». In: *2016 International conference on industrial informatics-computing technology, intelligent technology, industrial information integration (ICIICII)*. IEEE. 2016, pp. 150–153.
- [57] Dengyi Zhang et al. «Image based forest fire detection using dynamic characteristics with artificial neural networks». In: *2009 International Joint Conference on Artificial Intelligence*. IEEE. 2009, pp. 290–293.
- [58] Wen-Bing Horng and Jian-Wen Peng. «Image-based fire detection using neural networks». In: *9th Joint International Conference on Information Sciences (JCIS-06)*. Atlantis Press. 2006.

- [59] Jivitesh Sharma et al. «Deep convolutional neural networks for fire detection in images». In: *International Conference on Engineering Applications of Neural Networks*. Springer. 2017, pp. 183–193.
- [60] Diederik P Kingma and Jimmy Ba. «Adam: A method for stochastic optimization». In: *arXiv preprint arXiv:1412.6980* (2014).
- [61] Chenggang Yan et al. «Efficient parallel framework for HEVC motion estimation on many-core processors». In: *IEEE Transactions on Circuits and Systems for Video Technology* 24.12 (2014), pp. 2077–2089.
- [62] B Uğur Töreyn, Yiğithan Dedeoğlu, and A Enis Cetin. «Wavelet based real-time smoke detection in video». In: *2005 13th European Signal Processing Conference*. IEEE. 2005, pp. 1–4.
- [63] Thou-Ho Chen, Ping-Hsueh Wu, and Yung-Chuen Chiou. «An early fire-detection method based on image processing». In: *2004 International Conference on Image Processing, 2004. ICIP'04*. Vol. 3. IEEE. 2004, pp. 1707–1710.
- [64] Nobuyuki Fujiwara and Kenji Terada. «Extraction of a smoke region using fractal coding». In: *IEEE International Symposium on Communications and Information Technology, 2004. ISCIT 2004*. Vol. 2. IEEE. 2004, pp. 659–662.
- [65] ByoungChul Ko, JunOh Park, and Jae-Yeal Nam. «Spatiotemporal bag-of-features for early wildfire smoke detection». In: *Image and Vision Computing* 31.10 (2013), pp. 786–795.
- [66] Audrey Chenebert, Toby P Breckon, and Anna Gaszczak. «A non-temporal texture driven approach to real-time fire detection». In: *2011 18th IEEE International Conference on Image Processing*. IEEE. 2011, pp. 1741–1744.
- [67] Feiniu Yuan. «A fast accumulative motion orientation model based on integral image for video smoke detection». In: *Pattern Recognition Letters* 29.7 (2008), pp. 925–932.
- [68] Hongda Tian et al. «Smoke detection in videos using non-redundant local binary pattern-based features». In: *2011 IEEE 13th International workshop on multimedia signal processing*. IEEE. 2011, pp. 1–4.
- [69] Ruggero Donida Labati et al. «Wildfire smoke detection using computational intelligence techniques enhanced with synthetic smoke plume generation». In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 43.4 (2013), pp. 1003–1012.
- [70] Feiniu Yuan. «Video-based smoke detection with histogram sequence of LBP and LBPV pyramids». In: *Fire safety journal* 46.3 (2011), pp. 132–139.
- [71] Yu Chunyu et al. «Video fire smoke detection using motion and color features». In: *Fire technology* 46.3 (2010), pp. 651–663.
- [72] Tao Wang, Yuan Liu, and Zhen-Ping Xie. «Flutter analysis based video smoke detection». In: *Dianzi Yu Xinxi Xuebao (Journal of Electronics and Information Technology)* 33.5 (2011), pp. 1024–1029.
- [73] Sheng Luo et al. «Smoke detection based on condensed image». In: *Fire Safety Journal* 75 (2015), pp. 23–35.

- [74] Yanmin Luo et al. «Fire smoke detection algorithm based on motion characteristic and convolutional neural networks». In: *Multimedia Tools and Applications* (2018), pp. 1–18.
- [75] Jasper RR Uijlings et al. «Selective search for object recognition». In: *International journal of computer vision* 104.2 (2013), pp. 154–171.
- [76] Joao Carreira and Cristian Sminchisescu. «CPMC: Automatic object segmentation using constrained parametric min-cuts». In: *IEEE Transactions on pattern analysis and machine intelligence* 34.7 (2011), pp. 1312–1328.
- [77] Sheng Luo and YZ Jiang. «State-of-art of video based smoke detection algorithms». In: *Journal of Image and Graphics* 18.10 (2013), pp. 1225–1236.
- [78] Hongda Tian et al. «A novel video-based smoke detection method using image separation». In: *2012 IEEE International Conference on Multimedia and Expo*. IEEE. 2012, pp. 532–537.
- [79] Ivan Kolesov et al. «Fire and smoke detection in video with optimal mass transport based optical flow and neural networks». In: *2010 IEEE International Conference on Image Processing*. IEEE. 2010, pp. 761–764.
- [80] Salman Khan et al. «Energy-Efficient Deep CNN for Smoke Detection in Foggy IoT Environment». In: *IEEE Internet of Things Journal* (2019).
- [81] Gaohua Lin et al. «Smoke Detection on Video Sequences Using 3D Convolutional Neural Networks». In: *Fire Technology* (2019), pp. 1–21.
- [82] Du Tran et al. «Learning spatiotemporal features with 3d convolutional networks». In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 4489–4497.
- [83] Jeffrey Donahue et al. «Long-term recurrent convolutional networks for visual recognition and description». In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 2625–2634.
- [84] Andrej Karpathy et al. «Large-scale video classification with convolutional neural networks». In: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2014, pp. 1725–1732.
- [85] Geoffrey E. Hinton et al. «Improving neural networks by preventing co-adaptation of feature detectors». In: *CoRR* abs/1207.0580 (2012). arXiv: 1207.0580. URL: <http://arxiv.org/abs/1207.0580>.
- [86] Khan Muhammad et al. «Efficient deep CNN-based fire detection and localization in video surveillance applications». In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 99 (2018), pp. 1–16.
- [87] Forrest N Iandola et al. «SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size». In: *arXiv preprint arXiv:1602.07360* (2016).
- [88] Pasquale Foggia, Alessia Saggese, and Mario Vento. «Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion». In: *IEEE TRANSACTIONS on circuits and systems for video technology* 25.9 (2015), pp. 1545–1556.

- [89] Tom Toulouse et al. «Computer vision for wildfire research: An evolving image dataset for processing and analysis». In: *Fire Safety Journal* 92 (2017), pp. 188–194.
- [90] David G Lowe. «Distinctive image features from scale-invariant keypoints». In: *International journal of computer vision* 60.2 (2004), pp. 91–110.
- [91] Ethan Rublee et al. «ORB: An efficient alternative to SIFT or SURF.» In: *ICCV*. Vol. 11. 1. Citeseer. 2011, p. 2.
- [92] Edward Rosten and Tom Drummond. «Machine learning for high-speed corner detection». In: *European conference on computer vision*. Springer. 2006, pp. 430–443.
- [93] Michael Calonder et al. «Brief: Binary robust independent elementary features». In: *European conference on computer vision*. Springer. 2010, pp. 778–792.
- [94] A. Dunning and T.P. Breckon. «Experimentally defined Convolutional Nerual Network Architecture Variants for Non-temporal Real-time Fire Detection». In: *Proc. International Conference on Image Processing*. IEEE, Sept. 2018, pp. 1558–1562. DOI: 10.1109/ICIP.2018.8451657.
- [95] Martin Abadi et al. «TensorFlow: Large-scale machine learning on heterogeneous systems, 2015». In: *Software available from tensorflow. org* 1.2 (2015).
- [96] Sharan Chetlur et al. «cudnn: Efficient primitives for deep learning». In: *arXiv preprint arXiv:1410.0759* (2014).
- [97] François Chollet et al. «Keras: Deep learning library for theano and tensorflow». In: *URL: https://keras. io/k* 7.8 (2015), T1.
- [98] The Theano Development Team et al. «Theano: A Python framework for fast computation of mathematical expressions». In: *arXiv preprint arXiv:1605.02688* (2016).
- [99] Dong Yu et al. «An introduction to computational networks and the computational network toolkit». In: *Microsoft Technical Report MSR-TR-2014-112* (2014).
- [100] Eric Jones, Travis Oliphant, and Pearu Peterson. «{SciPy}: Open source scientific tools for {Python}». In: (2014).
- [101] Harold W Kuhn. «The Hungarian method for the assignment problem». In: *Naval research logistics quarterly* 2.1-2 (1955), pp. 83–97.
- [102] Eben Upton and Gareth Halfacree. *Raspberry Pi user guide*. John Wiley & Sons, 2014.
- [103] Anastasios Kokkalis and Theodore I Lekas. «Drones Surveillance—Challenges and Techniques». In: *Surveillance in Action*. Springer, 2018, pp. 185–194.