



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ Μ/Υ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΝΑΥΤΙΛΙΑΣ ΚΑΙ ΒΙΟΜΗΧΑΝΙΑΣ
ΤΜΗΜΑΤΟΣ ΒΙΟΜΗΧΑΝΙΚΗΣ ΔΙΟΙΚΗΣΗΣ & ΤΕΧΝΟΛΟΓΙΑΣ
ΔΙΑΠΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟ-ΟΙΚΟΝΟΜΙΚΑ ΣΥΣΤΗΜΑΤΑ»



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΔΑΜΙΑΝΟΣ ΚΑΛΑΝΙΔΗΣ

"Μηχανική όραση για ανίχνευση ανωμαλιών σε οδικές υποδομές με χρήση νευρωνικών δικτύων"

Επιβλέπων καθηγητής: ΔΟΥΛΑΜΗΣ ΑΝΑΣΤΑΣΙΟΣ
ΑΝΑΠΛΗΡΩΤΗΣ ΚΑΘΗΓΗΤΗΣ Ε.Μ.Π

Μηχανική όραση

- Ως άνθρωποι, αντιλαμβανόμαστε την τρισδιάστατη δομή του κόσμου γύρω μας με εμφανή ευκολία.
- Η μηχανική όραση, υπολογιστική όραση ή τεχνητή όραση είναι ένα επιστημονικό πεδίο της τεχνητής νοημοσύνης το οποίο επιχειρεί να αναπαράγει αλγοριθμικά την αίσθηση της όρασης, συνήθως σε ηλεκτρονικό υπολογιστή ή ρομπότ.
- Οι επιστήμονες στη μηχανική όραση, αναπτύσσουν παράλληλα μαθηματικά μοντέλα για την ανάκτηση των τρισδιάστατων σχημάτων και εμφάνισης των αντικειμένων σε μια εικόνα.
- Ωστόσο, παρόλη την πρόοδο, η δυνατότητα του υπολογιστή να αναγνωρίσει μια εικόνα στο ίδιο επίπεδο όπως ένα παιδί λίγων ετών, μοιάζει για τώρα κάτι απατηλό.
- Η πιο δύσκολη εκδοχή αναγνώρισης είναι η αναγνώριση γενικής κατηγορίας (ή τάξης), η οποία μπορεί να περιλαμβάνει την αναγνώριση περιπτώσεων εξαιρετικά διαφορετικών κατηγοριών, όπως ζώα ή έπιπλα.

Αναγνώριση προτύπων (Pattern Recognition)

Η επιστημονική περιοχή της αναγνώρισης προτύπων ασχολείται με την αυτόματη ανακάλυψη των υποκείμενων κανονικοτήτων στα δεδομένα με τη χρήση αλγορίθμων έτσι ώστε να πάρει αποφάσεις, όπως η ταξινόμηση των δεδομένων σε διαφορετικές κατηγορίες.

Ιστορικά υπήρχαν οι δύο πρώτες θεμελιώδεις προσεγγίσεις για την εφαρμογή ενός συστήματος αναγνώρισης προτύπων με την τρίτη να έχει ραγδαία εξέλιξη τα τελευταία χρόνια:

1. Στατιστική Αναγνώριση Προτύπων (ή θεωρητική απόφαση)

Βασίζεται σε υποκείμενα στατιστικά μοντέλα για την περιγραφή των πρότυπων και των τάξεων τους όπως ιστογράμματα.

2. Συντακτική Αναγνώριση Προτύπων (ή δομική).

Οι τάξεις αναπαρίστανται μέσω τυπικών δομών όπως γραμματικές, automata, strings.

3. Νευρωνική Αναγνώριση Προτύπων

Αναγνώριση προτύπων (Pattern Recognition)

Η γενική αναγνώριση αντικειμένων εμπίπτει σε δύο μεγάλες κατηγορίες, συγκεκριμένα την αναγνώριση στιγμιότυπων (instance recognition) και αναγνώριση κλάσεων (class recognition).

- ▶ Η πρώτη περιλαμβάνει την εκ νέου αναγνώριση ενός γνωστού 2D ή 3D άκαμπτου αντικειμένου, το οποίο ενδεχομένως παρατηρείται από μια νέα οπτική γωνία, έναντι ενός σε ένα ακατάστατου φόντου και με επιμέρους σκιάσεις.
- ▶ Η δεύτερη περίπτωση, η οποία είναι επίσης γνωστή ως αναγνώριση επιπέδου κατηγορίας ή αναγνώριση γενικού αντικειμένου, είναι το πολύ πιο δύσκολο πρόβλημα της αναγνώρισης οποιαδήποτε στιγμιότυπου συγκεκριμένης γενικής τάξης, όπως «γάτα», «αυτοκίνητο» ή «δέντρο».

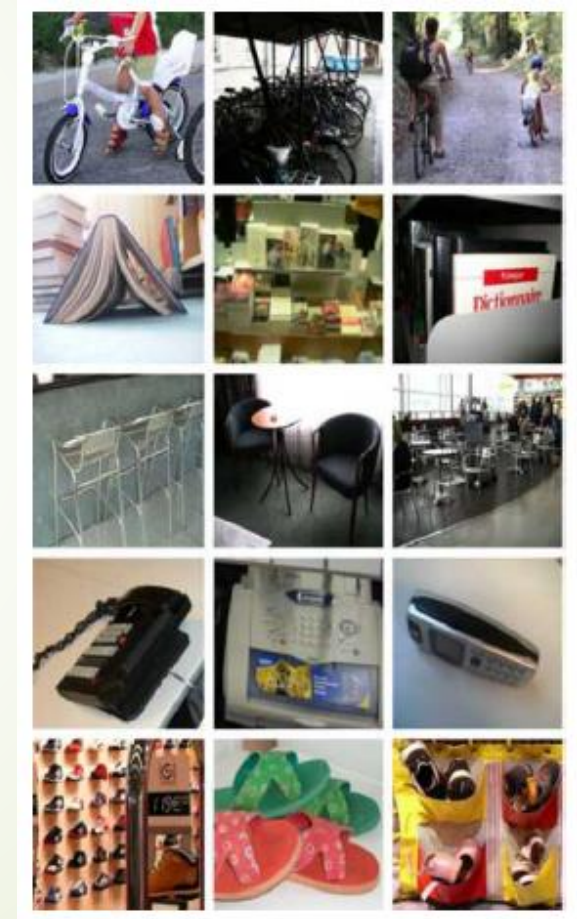
Αναγνώριση στιγμιότυπων (instance recognition)

- Στην αρχή οι προσεγγίσεις επικεντρώνονταν στην εξαγωγή γραμμών, περιγραμμάτων ή τρισδιάστατων επιφανειών από εικόνες και αντιστοίχιση τους με γνωστά μοντέλα τρισδιάστατων αντικειμένων.
- Καθώς ο αριθμός των αντικειμένων ή εικόνων σε μια συλλογή αρχίζει να γίνεται πολύ μεγάλος, ο χρόνος που χρειάζεται για να συγκρίνουμε μια νέα εικόνα με κάθε εικόνα της βάσης δεδομένων μπορεί να γίνει απαγορευτικός



Κατηγοριοποίηση εικόνων (image classification)

- Η οπτική αναγνώριση κατηγορίας είναι ένα εξαιρετικά απαιτητικό πρόβλημα
- Κανείς δεν έχει κατασκευάσει ακόμη ένα σύστημα που να προσεγγίζει το επίπεδο απόδοσης ενός παιδιού δύο ετών.
- Ωστόσο, η πρόοδος στον τομέα ήταν αρκετά δραματική, αν κριθεί το πόσο καλύτεροι είναι οι σημερινοί αλγόριθμοι σε σύγκριση με αυτούς της προηγούμενης δεκαετίας.



Κατηγοριοποίηση εικόνων (image classification)

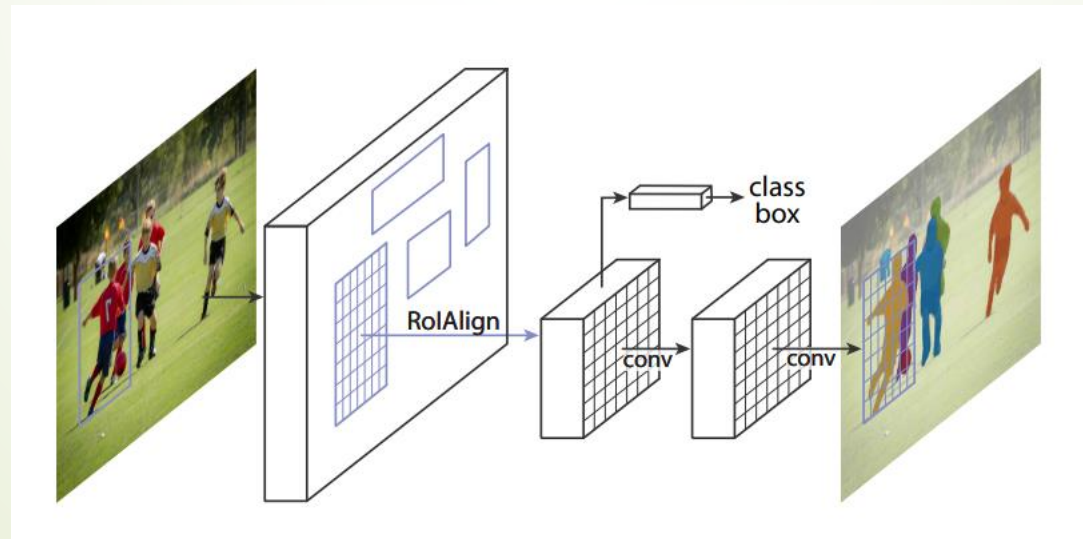
- ▶ Υπάρχουν διάφορες προσεγγίσεις για την κατηγοριοποίηση των εικόνων (Image classification) με κυριότερες να είναι:
 - Οι κλασικές προσεγγίσεις βασισμένες σε χαρακτηριστικά (part-based representations) που βασίζονται σε χειροποίητα χαρακτηριστικά και τα στατιστικά τους και προαιρετικά χρησιμοποιούν μηχανική εκμάθηση για την τελική κατηγοριοποίηση
 - Τα σύγχρονα συστήματα ταξινόμησης εικόνων, τα οποία βασίζονται στα βαθιά νευρωνικά δίκτυα (deep neural networks)

Τμηματικά μοντέλα (Part-based models)

- ▶ Αναγνώριση ενός αντικειμένου με την εύρεση των συστατικών του μερών και τη μέτρηση των γεωμετρικών τους
- ▶ Δεν αρκεί ο εντοπισμός των επιμέρους χαρακτηριστικών μέσα στην εικόνα, αλλά αναζητείται και η επιβεβαίωση συγκεκριμένων χωρικών σχέσεων μεταξύ αυτών των χαρακτηριστικών.

Αναγνώριση με τμηματοποίηση (Recognition with segmentation)

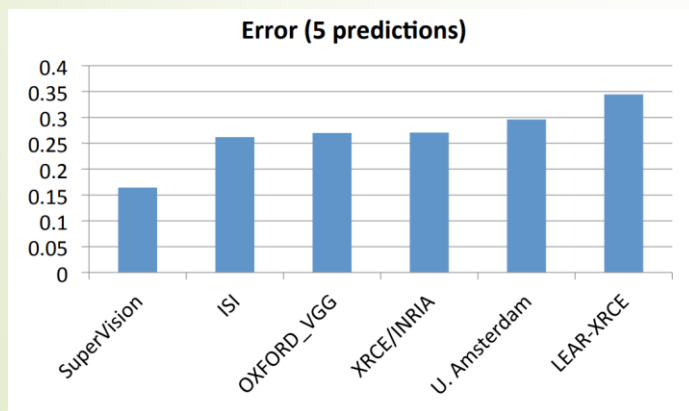
- ▶ Η κατάτμηση εικόνας είναι το έργο της εύρεσης ομάδων εικονοστοιχείων που «πάνε μαζί». Στη στατιστική, αυτό το πρόβλημα είναι γνωστό ως ανάλυση συστάδων και είναι μια ευρέως μελετημένη περιοχή με εκατοντάδες διαφορετικούς αλγορίθμους.



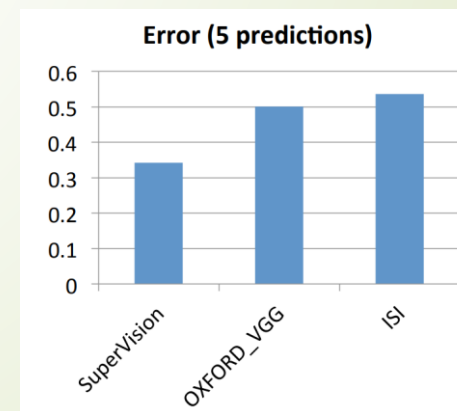
Βαθιά δίκτυα - Deep networks

- Με τον όρο βαθιά δίκτυα ονομάζουμε τα δίκτυα όπου περιέχουν παραπάνω από ένα εσωτερικά επίπεδα επεξεργασίας μεταξύ του επιπέδου εισόδου και μεταξύ του επιπέδου εξόδου, για την επίλυση ενός συγκεκριμένου προβλήματος.
- Από την προηγούμενη δεκαετία, τα βαθιά δίκτυα (deep networks) άρχισαν να ξεπερνούν τις προσεγγίσεις βασισμένες στη «ρηχή» μάθηση, με γνωστό παράδειγμα το διαγωνισμό ImageNet Large Scale Visual Recognition Challenge (ILSVRC) το 2012 όπου κέρδισε το βαθύ νευρωνικό δίκτυο (deep neural network) SuperVision/AlexNet των Krizhevsky, Sutskever, and Hinton.

Classification

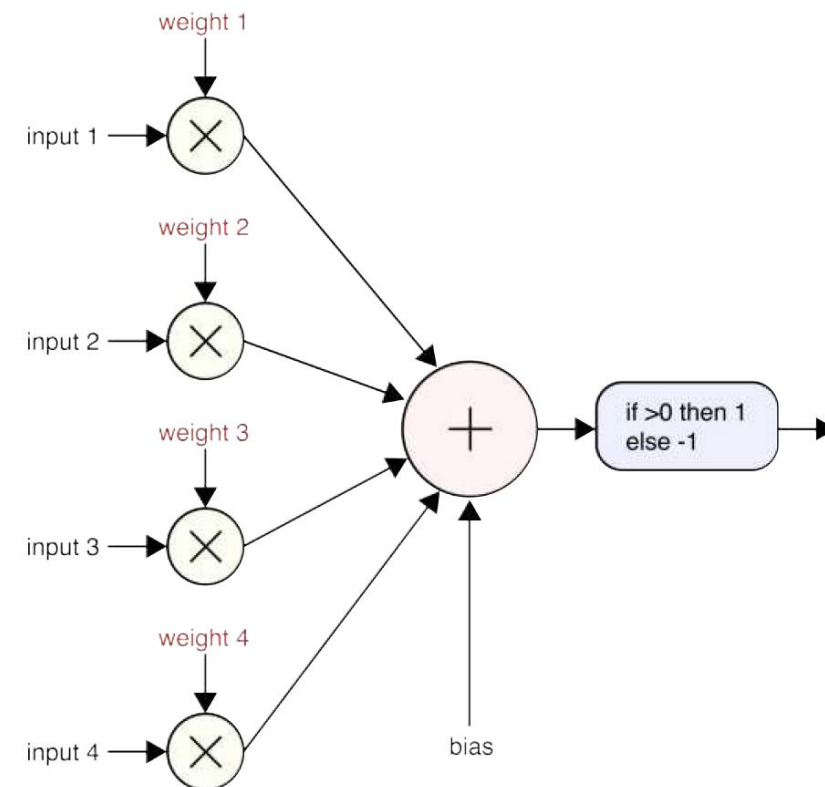


Detection



Νευρωνικά Δίκτυα - Neural networks

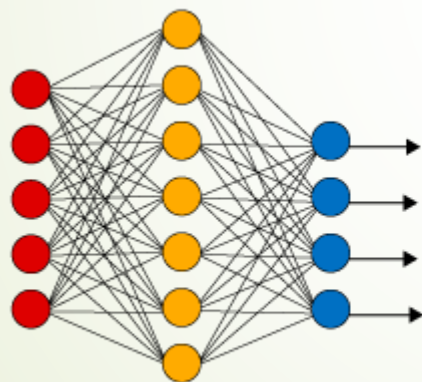
- Βασικό δομικό στοιχείο των Νευρωνικών Δικτύων είναι οι νευρώνες-μονάδες(units) όπου εκτελούν σταθμισμένα αθροίσματα των εισροών τους και στη συνέχεια αντιστοίχιση σε μια μη γραμμική συνάρτησης ενεργοποίησης.
- Κάθε μονάδα εκτελεί σταθμισμένα αθροίσματα των εισροών της $s_i = \mathbf{w}_i^T \mathbf{x}_i + b_i$ και στη συνέχεια αντιστοίχιση σε μια μη γραμμική συνάρτησης ενεργοποίησης $y_i = h(s_i)$
- Το perceptron είναι η απλούστερη μορφή νευρωνικού δικτύου και αποτελείται από ένα επίπεδο όπου η μονάδα νευρώνας παρέχει μόνο μια έξοδο αποτελέσματος.



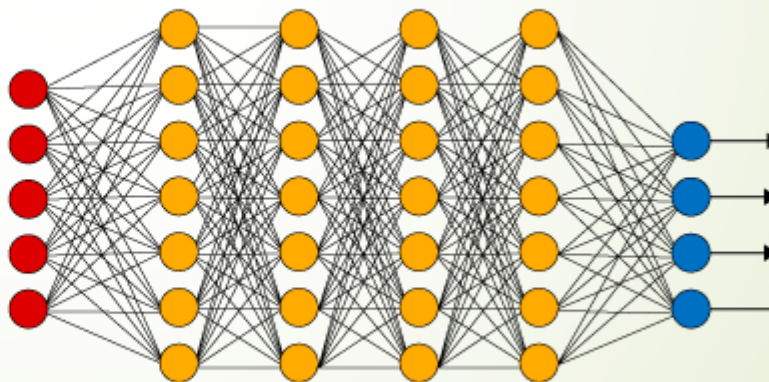
Deep neural networks (DNNs)

- Τα Βαθιά Νευρωνικά δίκτυα (DNNs) είναι νευρωνικά δίκτυα όπου περιέχουν παραπάνω από ένα εσωτερικά επίπεδα επεξεργασίας

Simple Neural Network



Deep Learning Neural Network



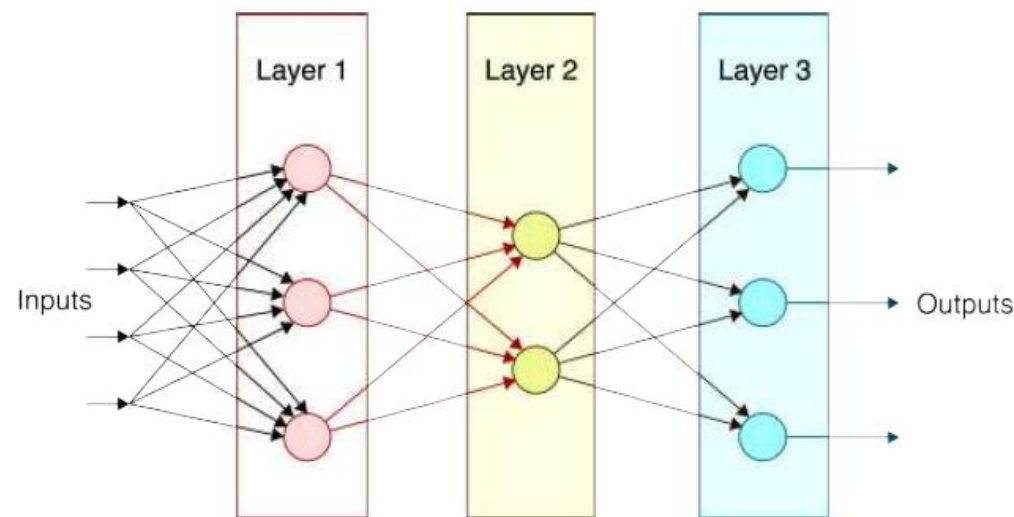
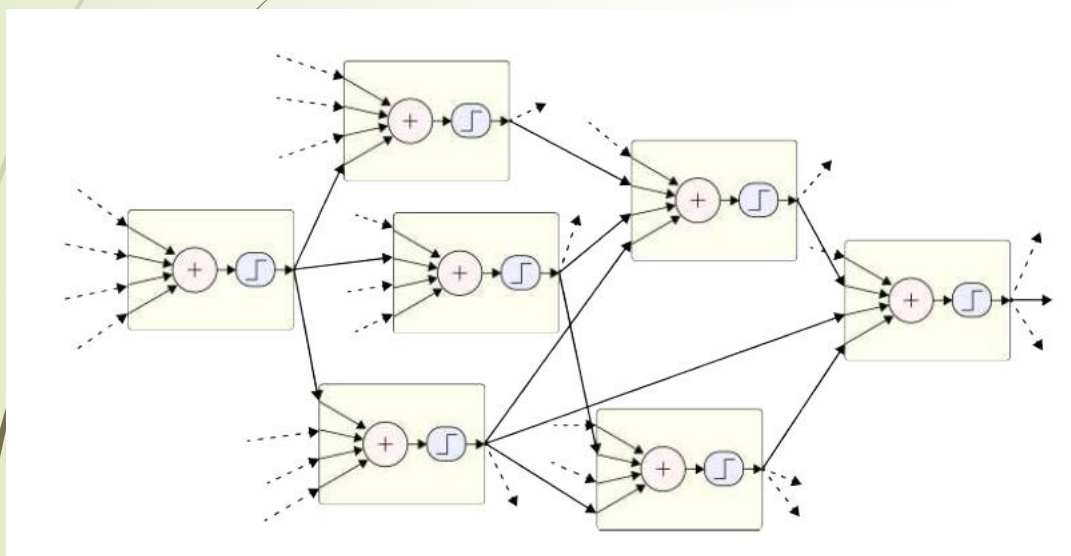
● Input Layer

● Hidden Layer

● Output Layer

Τρόποι διασύνδεσης μονάδων DNN

- Οι μονάδες δεν συνδέονται μεταξύ τους με τη μορφή ακανόνιστου γράφου όπως στην αριστερή εικόνα αλλά συνήθως οργανώνονται σε διαδοχικά στρώματα (layers) όπως στη δεξιά εικόνα.

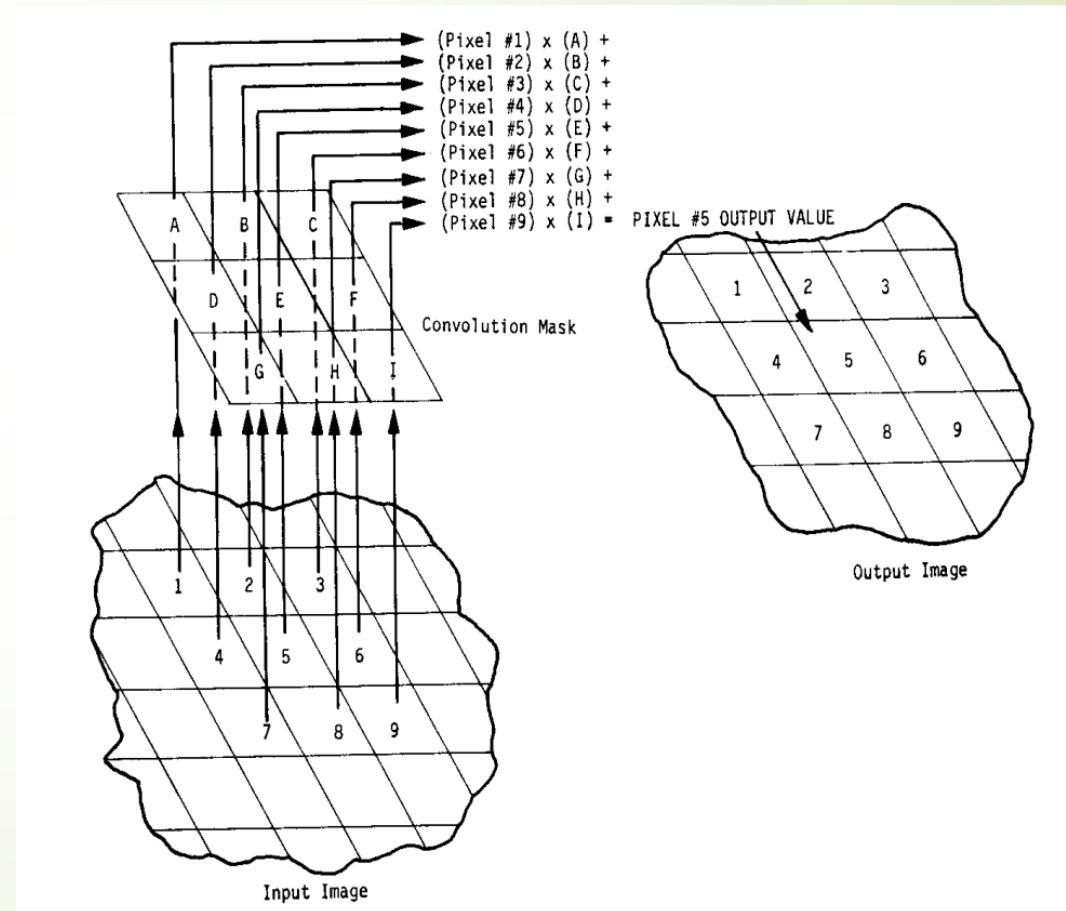


Συνελικτικά νευρωνικά δίκτυα (Convolutional neural networks)

- Η συνέλιξη (convolution) είναι μία μαθηματική μέθοδος που χρησιμοποιείται ευρέως στην ανάλυση σημάτων και αποτελεί μία θεμελιώδη διαδικασία μετατροπής δεδομένων εικόνας.

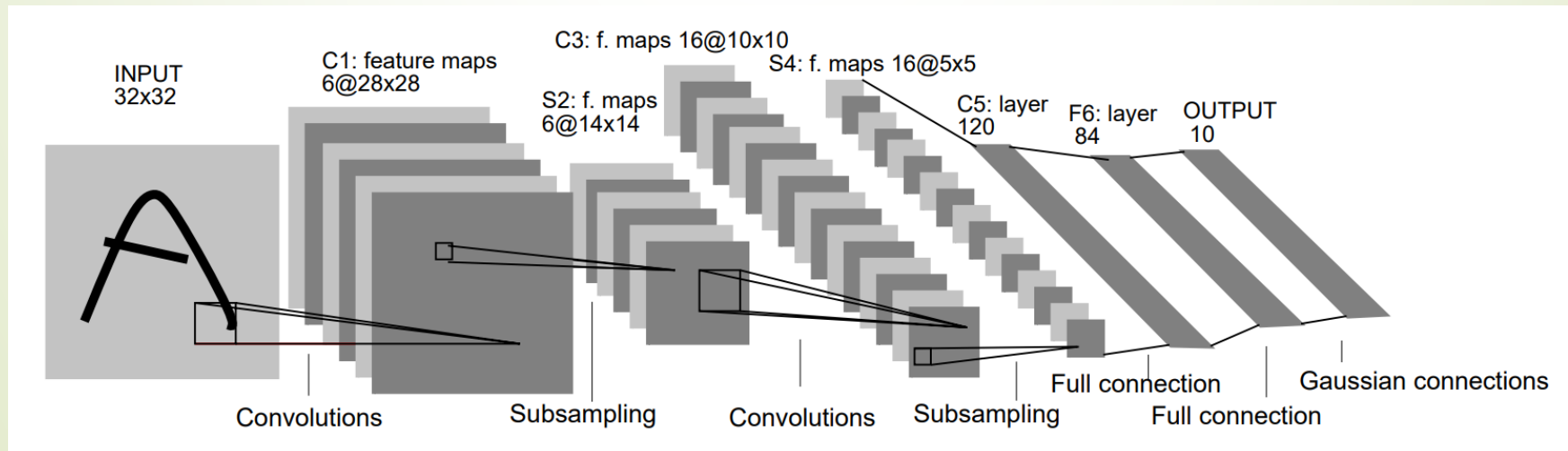
$$\begin{bmatrix} a & b & b \\ d & e & f \\ g & h & i \end{bmatrix} \times \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix} = [e']$$

Input *Kernel* *output*



Συνελικτικά νευρωνικά δίκτυα (Convolutional neural networks)

- ▶ Τα Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΔ) - Convolutional Neural Networks (CNN), χωρίζονται σε δυο μεγάλες κατηγορίες τα Αβαθή Νευρωνικά Δίκτυα (Shallow Neural Networks) και τα Βαθιά Νευρωνικά Δίκτυα (Deep Neural Networks).
- ▶ Ένα συνελικτικό επίπεδο (convolutional layer) είναι ουσιαστικά ένα σύνολο από νευρώνες που εκτελούν συνέλιξη των φίλτρων που έχουν προκαθοριστεί, με την εικόνα-διάνυσμα που δέχονται στην είσοδο.
- ▶ Αντί να συνδέουν όλες τις μονάδες ενός επιπέδου με όλες τις μονάδες ενός προηγούμενου επιπέδου, τα συνελικτικά δίκτυα οργανώνουν κάθε επίπεδο σε χάρτες χαρακτηριστικών.



Βάσεις δεδομένων αναγνώρισης εικόνων και σύνολα δοκιμών

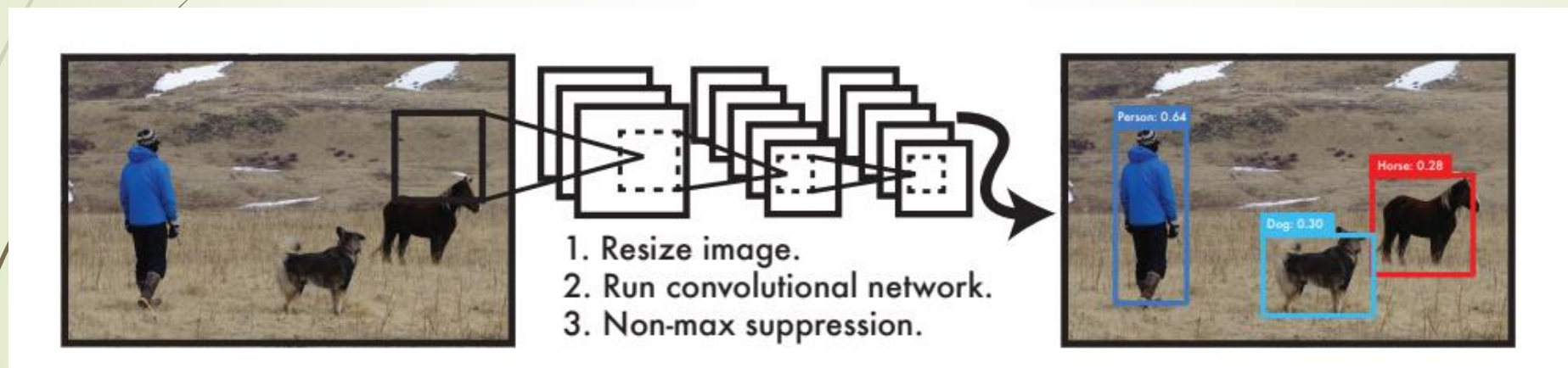
- Ένα από τα βασικά συστατικά για τη συνεχή βελτίωση των αλγορίθμων αναγνώρισης ήταν η αυξημένη διαθεσιμότητα και ποιότητα των βάσεων δεδομένων αναγνώρισης εικόνων



Δείγματα εικόνων από τον διαγωνισμό PASCAL Visual Object Classes Challenge 2008

YOLO v1

- Ο αλγόριθμος YOLO ανιχνεύει και αναγνωρίζει διάφορα αντικείμενα σε μια εικόνα σε πραγματικό χρόνο) και χρησιμοποιεί συνελκτικά νευρωνικά δίκτυα.
- Περιεγράφηκε το 2015 από τους Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi.



Η επεξεργασία εικόνων με το YOLO είναι απλή και απλή. Το σύστημά (1) αλλάζει το μέγεθος της εικόνας εισόδου σε 448×448 , (2) εκτελεί ένα ενιαίο συνελκτικό δίκτυο στην εικόνα και (3) ορίζει τα κατώφλια των ανιχνεύσεων που προκύπτουν με βάση την εμπιστοσύνη του μοντέλου

YOLO v1

- ▶ **Πλεονεκτήματα :** εξαιρετικά γρήγορο, λαμβάνει υπόψιν του συνολικά την εικόνα όταν κάνει προβλέψεις, μαθαίνει γενικεύσιμες αναπαραστάσεις αντικειμένων (πχ. Πρόσωπο σε εικόνα, έργα τέχνης)
- ▶ **Μειονεκτήματα:** Το YOLO εξακολουθεί να υστερεί σε σχέση με τα υπερσύγχρονα συστήματα ανίχνευσης στην ακρίβεια. Ενώ μπορεί να αναγνωρίσει γρήγορα αντικείμενα στις εικόνες, δυσκολεύεται να εντοπίσει με ακρίβεια ορισμένα αντικείμενα, ειδικά μικρά.
- ▶ **Περιορισμοί:** Το YOLO επιβάλλει ισχυρούς χωρικούς περιορισμούς στις προβλέψεις πλαισίων οριοθέτησης αφού κάθε κελί πλέγματος προβλέπει μόνο δύο πλαίσια και μπορεί να έχει μόνο μία κλάση. χρησιμοποιεί σχετικά χονδροειδή χαρακτηριστικά για την πρόβλεψη οριοθετημένων πλαισίων, λόγω πολλαπλών επίπεδων μείωσης δειγματοληψίας

YOLO v2 (YOLO9000)

Η επόμενη έκδοση του YOLO παρουσιάστηκε το 2016

Βελτιώσεις

	YOLO								YOLOv2
batch norm?		✓	✓	✓	✓	✓	✓	✓	✓
hi-res classifier?			✓	✓	✓	✓	✓	✓	✓
convolutional?				✓	✓	✓	✓	✓	✓
anchor boxes?				✓	✓				
new network?					✓	✓	✓	✓	✓
dimension priors?						✓	✓	✓	✓
location prediction?						✓	✓	✓	✓
passthrough?							✓	✓	✓
multi-scale?								✓	✓
hi-res detector?									✓
VOC2007 mAP	63.4	65.8	69.5	69.2	69.6	74.4	75.4	76.8	78.6

► Κανονικοποίηση παρτίδας (Batch Normalization)

Batch normalization στα convolutional layers. +2% mAP, αφαίρεση dropout από το μοντέλο χωρίς overfitting.

► Ταξινομητής υψηλής ανάλυσης (High Resolution Classifier).

Εκπαίδευση στα πρώτα 10 epochs σε 448x448, μετά 224x 224

► Πρόβλεψη μετατοπίσεων (Convolutional With Anchor Boxes)

Πιο ευέλικτη πρόβλεψη σε διάφορες περιοχές της εικόνας

► Dimension Clusters

Βελτιστοποίηση της περιοχής των bounding boxes

► Direct location prediction

► Fine-Grained Features

YOLO v3

► Η έκδοση v3 παρουσιάστηκε το 2018 έχει βελτιώσεις σε σχέση με την v2 στα ακόλουθα:

► **Bounding Box Prediction.**

Βελτιστοποιημένη πρόβλεψη σε σχέση με το YOLO v2

► **Class Prediction**

Ανίχνευση πολλαπλών κλάσεων εντός του bounding box

► **Predictions Across Scales**

► **Feature Extractor**

Χρησιμοποιείται ως βάση ένα νέο δίκτυο για την εκτέλεση εξαγωγής χαρακτηριστικών που έχει 53 συνελκτικά επίπεδα και ονομάζεται Darknet-53.

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1x	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2x	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8x	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8x	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4x	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
Connected		1000		
Softmax				

Δομή του Darknet-53

Οδοστρώματα - Τύποι Φθορών

Ρηγματώσεις (Cracking)



Ρηγμάτωση στα άκρα του οδοστρώματος (Edge cracking)



Οδοστρώματα - Τύποι Φθορών

Λακκούβες (Potholes)



Φθορές οριζόντιας σήμανσης (Markings wear)



ADAS (Advanced Driving Assistance Systems)

- Με τον όρο ADAS περιγράφουμε ένα σύνολο συστημάτων που είναι απαραίτητα για να παρακολουθούν διάφορες παραμέτρους που σχετίζονται με το όχημα και το περιβάλλον του οχήματος για τον εντοπισμό δυνητικά επικίνδυνων καταστάσεων σε ένα πρώιμο στάδιο.

SAE INTERNATIONAL

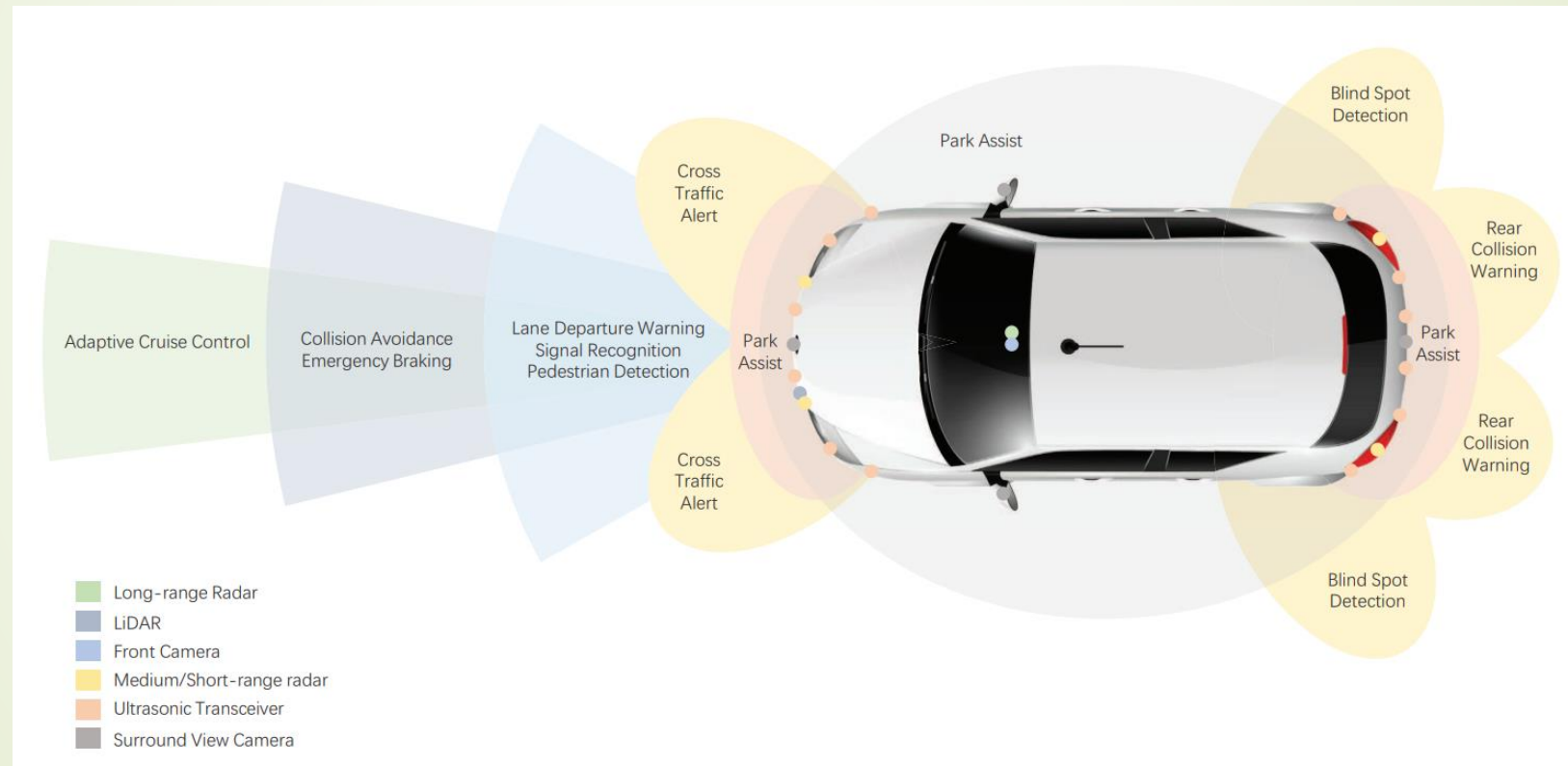
SAE J3016™ LEVELS OF DRIVING AUTOMATION™

Learn more here: [sae.org/standards/content/j3016_202104](https://www.sae.org/standards/content/j3016_202104)

Copyright © 2021 SAE International. The summary table may be freely copied and distributed AS-IS provided that SAE International is acknowledged as the source of the content.

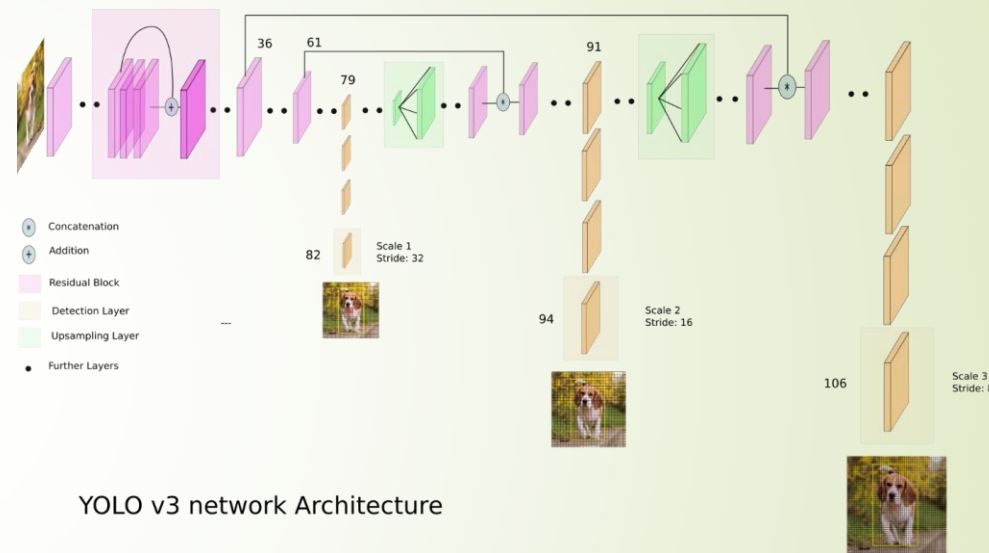
	SAE LEVEL 0™	SAE LEVEL 1™	SAE LEVEL 2™	SAE LEVEL 3™	SAE LEVEL 4™	SAE LEVEL 5™
What does the human in the driver's seat have to do?	You are driving whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering			You are not driving when these automated driving features are engaged – even if you are seated in “the driver’s seat”		
	You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety			When the feature requests, you must drive	These automated driving features will not require you to take over driving	
<small>Copyright © 2021 SAE International.</small>						
	These are driver support features			These are automated driving features		
What do these features do?	These features are limited to providing warnings and momentary assistance	These features provide steering OR brake/acceleration support to the driver	These features provide steering AND brake/acceleration support to the driver	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met		This feature can drive the vehicle under all conditions
Example Features	<ul style="list-style-type: none"> • automatic emergency braking • blind spot warning • lane departure warning 	<ul style="list-style-type: none"> • lane centering OR • adaptive cruise control 	<ul style="list-style-type: none"> • lane centering AND • adaptive cruise control at the same time 	<ul style="list-style-type: none"> • traffic jam chauffeur 	<ul style="list-style-type: none"> • local driverless taxi • pedals/steering wheel may or may not be installed 	<ul style="list-style-type: none"> • same as level 4, but feature can drive everywhere in all conditions

ADAS (Advanced Driving Assistance Systems)



Μεθοδολογία – Πειραματικό μέρος

- Για την εκτέλεση του πειραματικού μέρους της εργασίας χρησιμοποιήθηκε το **YOLO v3**
- 106 layers (εκ των οποίων τα 53 convolutional)
- Μέγεθος της παρτίδας (Batch size) : ο αριθμός των εικόνων που επεξεργαζόμαστε πριν από την ενημέρωση των βαρών δικτύου.

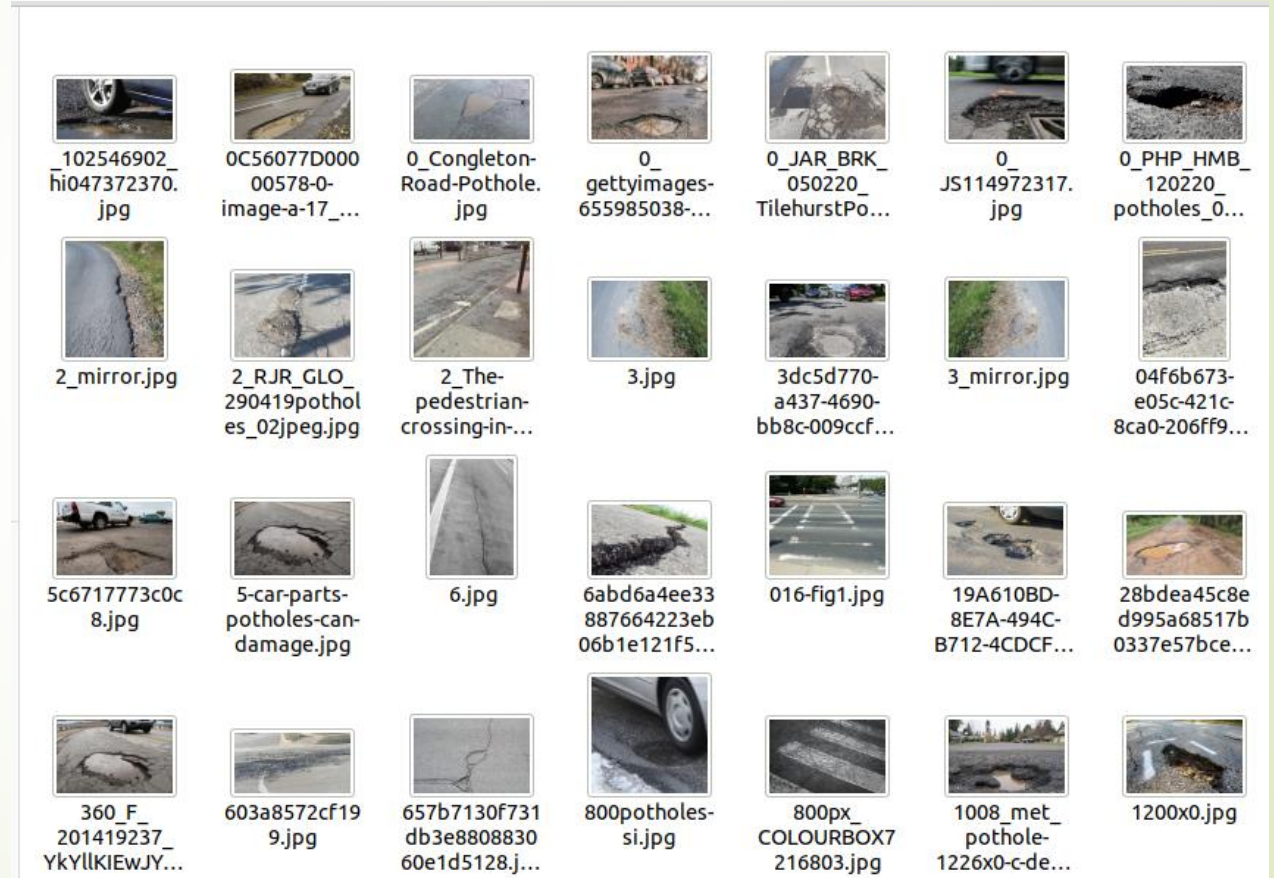


YOLO v3 network Architecture

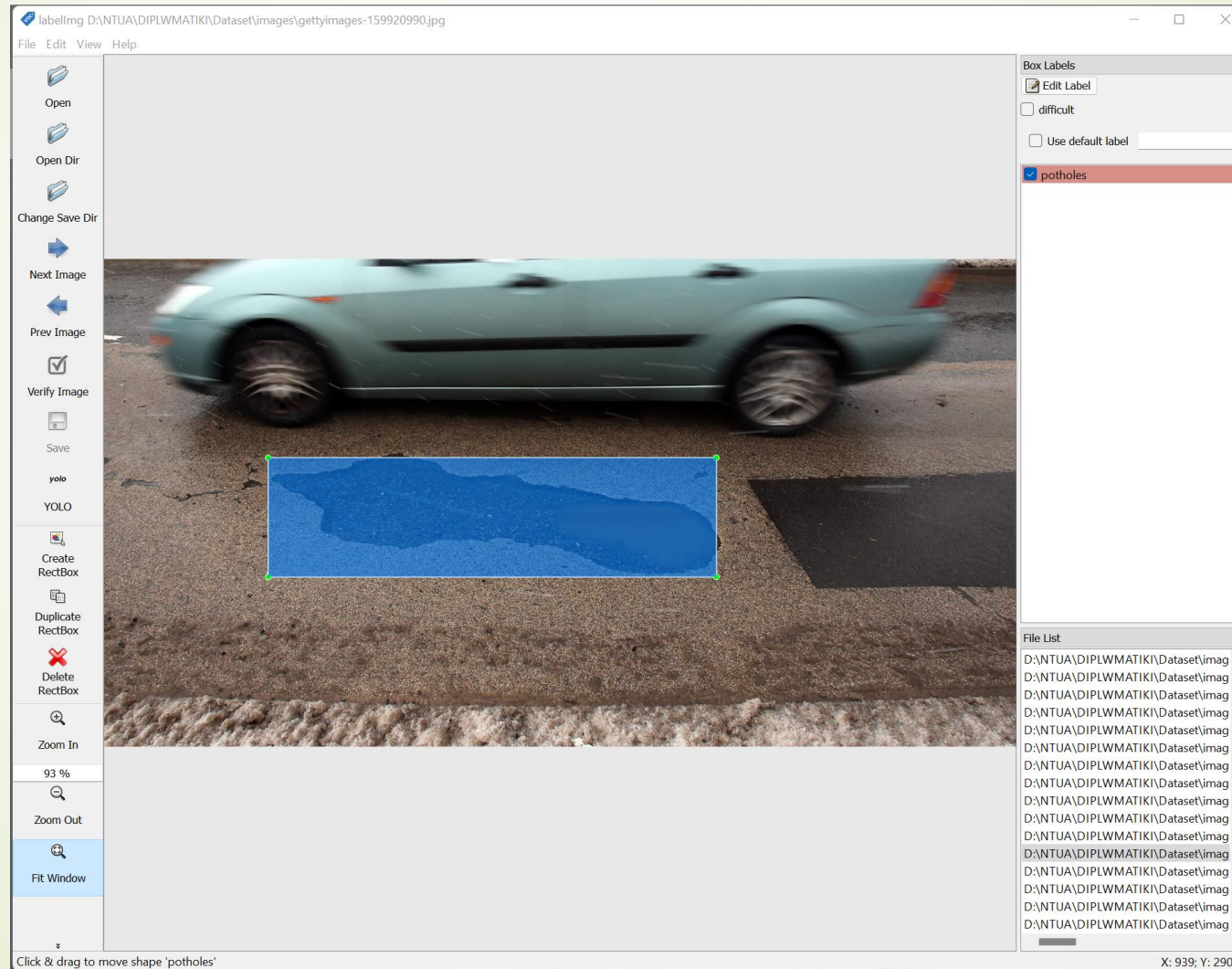
- Το μέγεθος της παρτίδας διαιρείται σε υποδιαιρέσεις (subdivisions) κατά τη φόρτωση του δικτύου, δηλαδή πόσες εικόνες μπορούν να χωρέσουν στη GPU που εκτελεί τον αλγόριθμο ταυτόχρονα. Με τις υποδιαιρέσεις προκύπτει το πλήθος των mini-batches (Batch/subdivisions = mini-batches).
- Μια επανάληψη (iteration) αποτελεί τον αριθμό των batches που χρειαζόμαστε για να ολοκληρωθεί ένα Epoch. Ένα Epoch έχουμε όταν κάνουμε μια πλήρης διαπέραση του training dataset.

Dataset Εκπαίδευσης - Δοκιμών

- Το dataset περιείχε ένα σύνολο από εικόνες που περιείχαν 4 κατηγορίες φθορών οδοστρώματος:
- 1. Ρωγμές (cracking)
- 2. αστοχίες στην άκρη(edge failure)
- 3. φθορά σημάτων (markings wear)
- 4. λακκούβες (potholes)

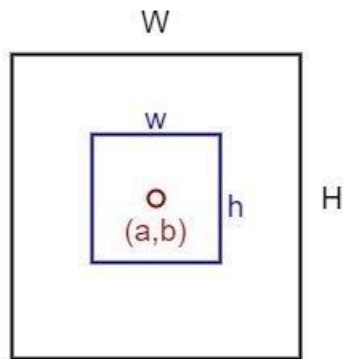


Επισήμανση περιοχών φθορών - Εργαλείο labeling



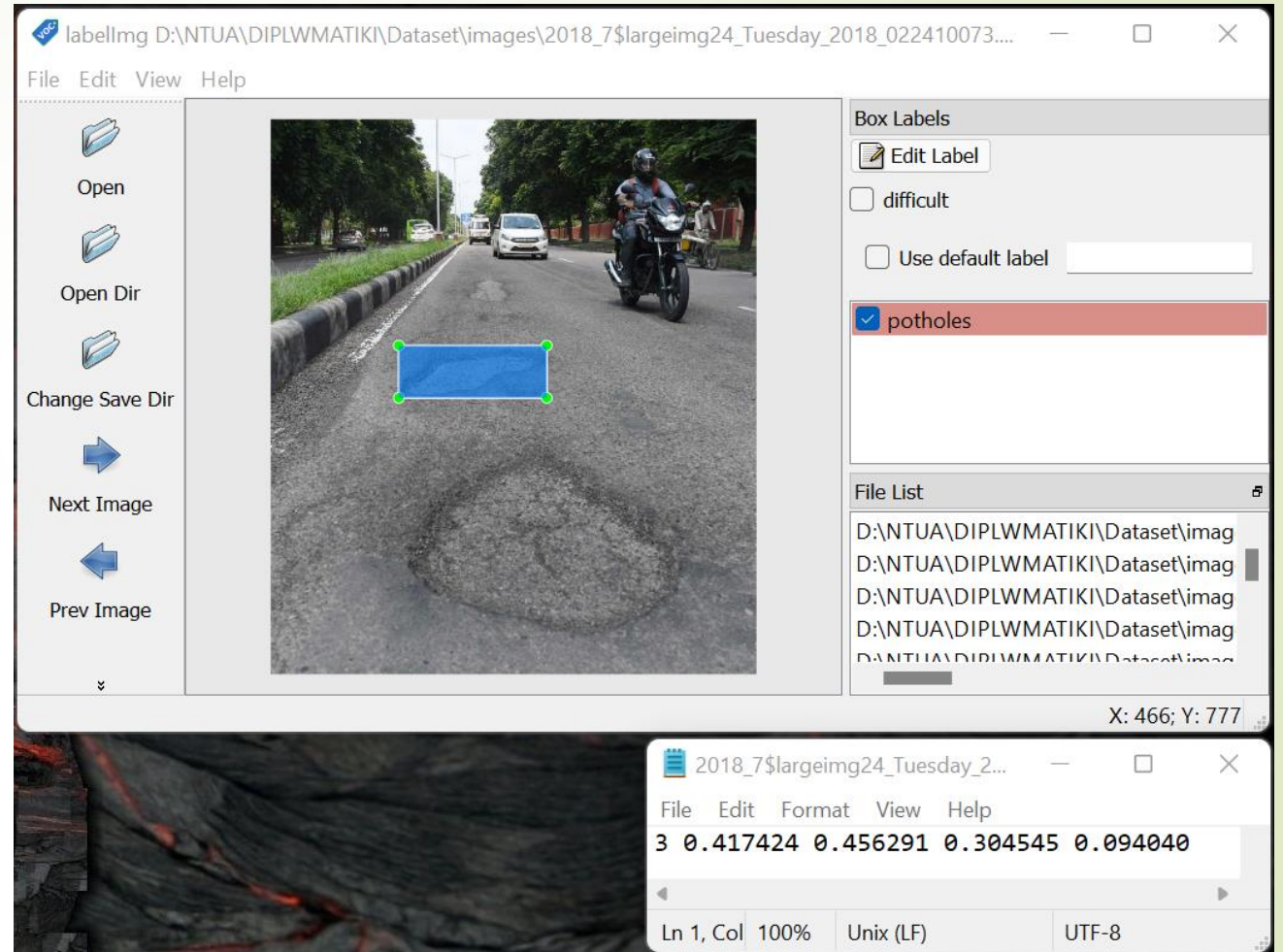
Δομή αρχείου Yolo

- Το Yolo Text file format είναι ένας τύπος αρχείου κειμένου, που χρησιμοποιείται από τον αλγόριθμο του YOLO. Αυτός ο τύπος αρχείου καθορίζει σε κάθε γραμμή του αρχείου την απόλυτη θέση του αντικειμένου στην εικόνα και τον αριθμό της κλάση στην οποία ανήκει, και αυτή την περιγραφή θα χρησιμοποιήσουμε για να κάνουμε την εκπαίδευση



W = Image Width
H = Image Height
w = Absolute Width
h = Absolute Height
(a,b) = Absolute x,y coordinates

YOLO FORMAT = <Class Number><a/W><b/H><w/W><h/H>



Dataset Εκπαίδευσης - Δοκιμών

Για το σύνολο των 382 εικόνων του dataset έγιναν τα ακόλουθα μαρκαρίσματα

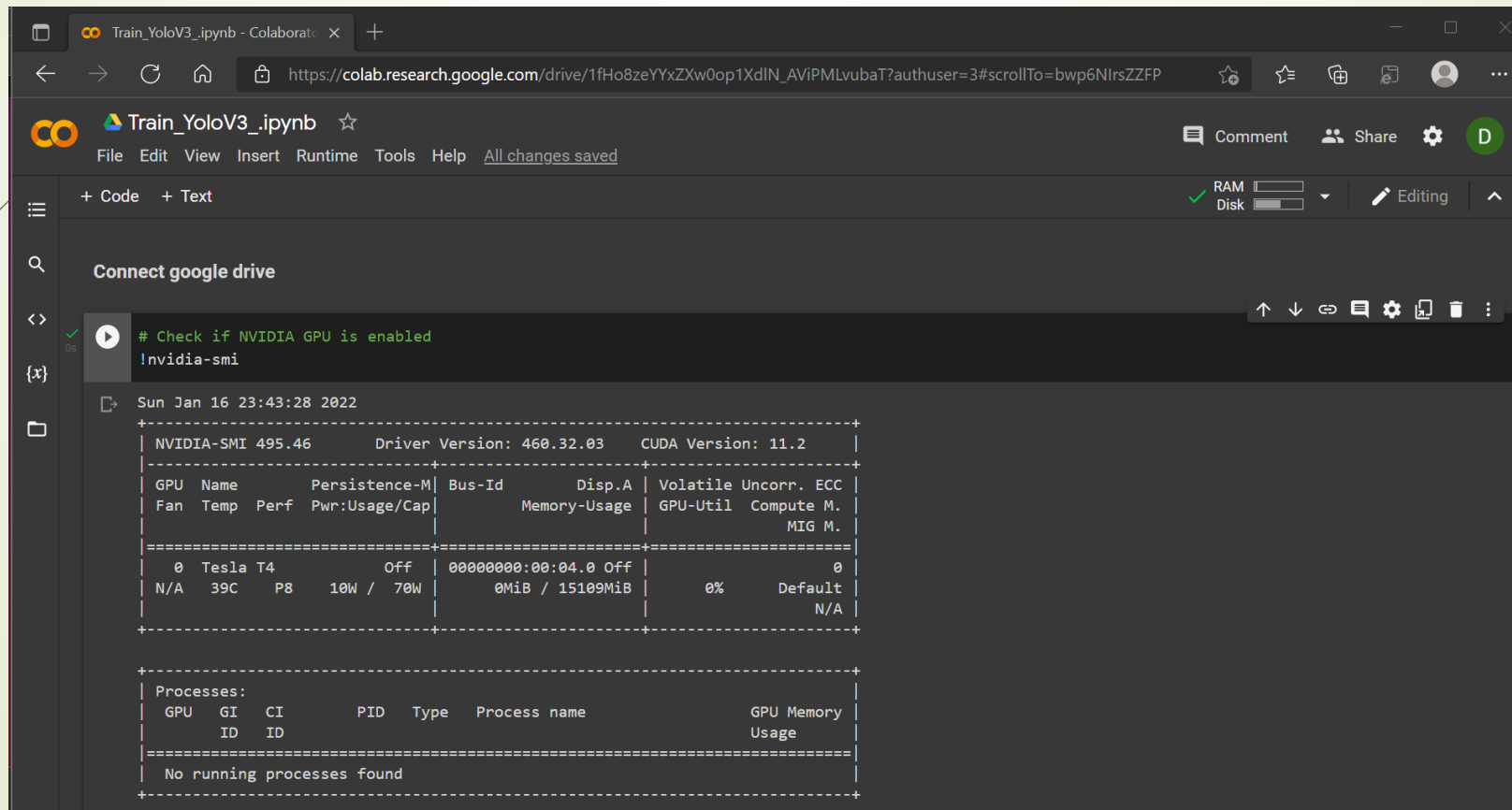
Κατηγορία φθορών οδοστρώματος	Σύνολο μαρκαρισμάτων
cracking	447
markings wear	406
edge failure	59
potholes	239

Επίσης για την χρησιμοποιήθηκε ένα υποσύνολο 354 εικόνων από το dataset με το ακόλουθο πλήθος μαρκαρισμένων περιοχών

Κατηγορία φθορών οδοστρώματος	Σύνολο μαρκαρισμάτων
cracking	211
markings wear	89
edge failure	17
potholes	79

Εκπαίδευση μοντέλου

- Για την εκπαίδευση του μοντέλου χρησιμοποίησα τη πλατφόρμα Google Colaboratory.



```
# Check if NVIDIA GPU is enabled
!nvidia-smi

Sun Jan 16 23:43:28 2022
+-----+
| NVIDIA-SMI 495.46      Driver Version: 460.32.03   CUDA Version: 11.2     |
+-----+-----+
| GPU Name      Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC | |
| Fan  Temp  Perf  Pwr:Usage/Cap|      Memory-Usage | GPU-Util  Compute M. |
|               |                  |              |   MIG M.               |
+-----+-----+
| 0   Tesla T4      Off          | 00000000:00:04:0  Off   |    0%      Default  |
| N/A   39C    P8      10W / 70W |  0MiB / 15109MiB |             |                      |
+-----+-----+

+-----+
| Processes:
| GPU   GI   CI        PID   Type   Process name                      GPU Memory
|   ID   ID                                     Usage
+-----+
| No running processes found
+-----+
```

Εκπαίδευση μοντέλου

Κατά τη διάρκεια της εκπαίδευσης του μοντέλου ρυθμίζονται κάποιες παράμετροι του Darknet ανάλογα πως θέλουμε να εκπαιδύσουμε το μοντέλο μας. Οι σημαντικότερες που αλλάζουμε κάθε φορά που εκτελούμε την εκπαίδευση του μοντέλου μας και που έχουμε περιγράψει νωρίτερα είναι :

`batch`: Το μέγεθος της παρτίδας (Batch size)

`subdivisions`: Οι υποδιαιρέσεις, δηλαδή το σύνολο των εικόνων που επεξεργάζεται μαζί η GPU

`max_batches`: Το μέγιστο πλήθος των iterations

Εκπαίδευση μοντέλου

- ▶ Παραμετροποίηση του αρχείου `yolov3_training.cfg` που περιέχει τις ρυθμίσεις για το training με τις ακόλουθες τιμές :

`batch=8, subdivisions=4, max_batches = 8000, classes=4, filters=27`

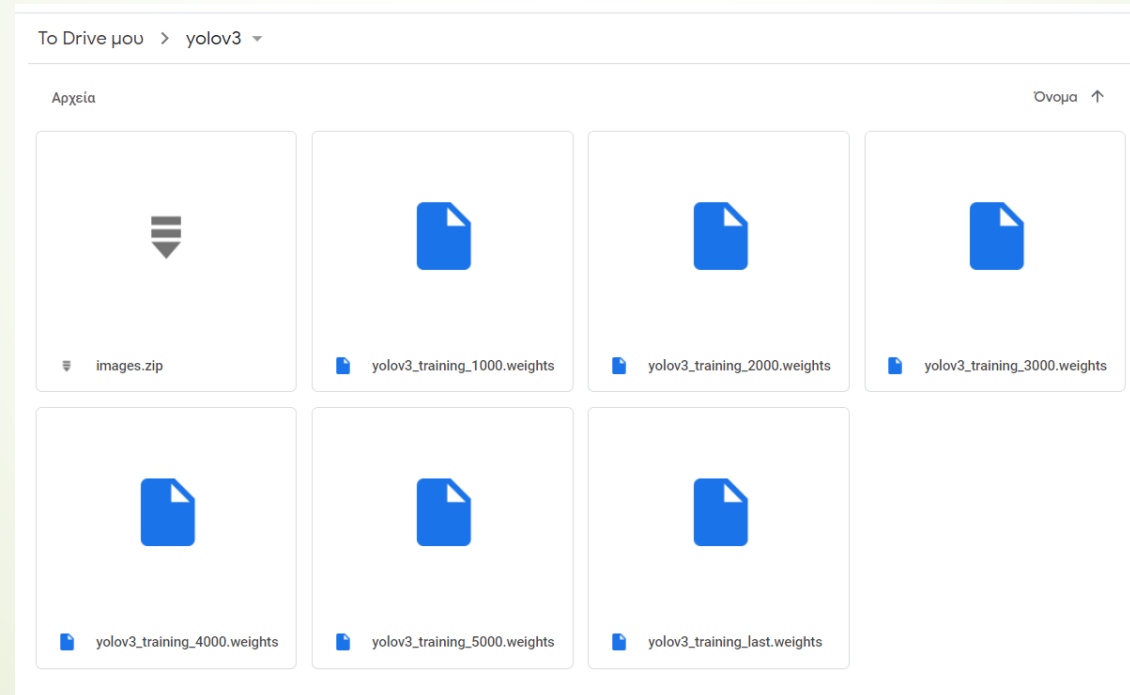
3) Configure Darknet network for training YOLO V3

```
✓ [5] !cp cfg/yolov3.cfg cfg/yolov3_training.cfg
```

```
✓ [6] !sed -i 's/batch=1/batch=8/' cfg/yolov3_training.cfg  
!sed -i 's/subdivisions=1/subdivisions=4/' cfg/yolov3_training.cfg  
!sed -i 's/max_batches = 500200/max_batches = 8000/' cfg/yolov3_training.cfg  
!sed -i '610 s@classes=80@classes=4@' cfg/yolov3_training.cfg  
!sed -i '696 s@classes=80@classes=4@' cfg/yolov3_training.cfg  
!sed -i '783 s@classes=80@classes=4@' cfg/yolov3_training.cfg  
!sed -i '603 s@filters=255@filters=27@' cfg/yolov3_training.cfg  
!sed -i '689 s@filters=255@filters=27@' cfg/yolov3_training.cfg  
!sed -i '776 s@filters=255@filters=27@' cfg/yolov3_training.cfg
```


Εκπαίδευση μοντέλου

- Με την ολοκλήρωση της διαδικασίας εκπαίδευσης, τα βάρη που προκύπτουν από την εκπαίδευση αποθηκεύονται στα κατάλληλα αρχεία

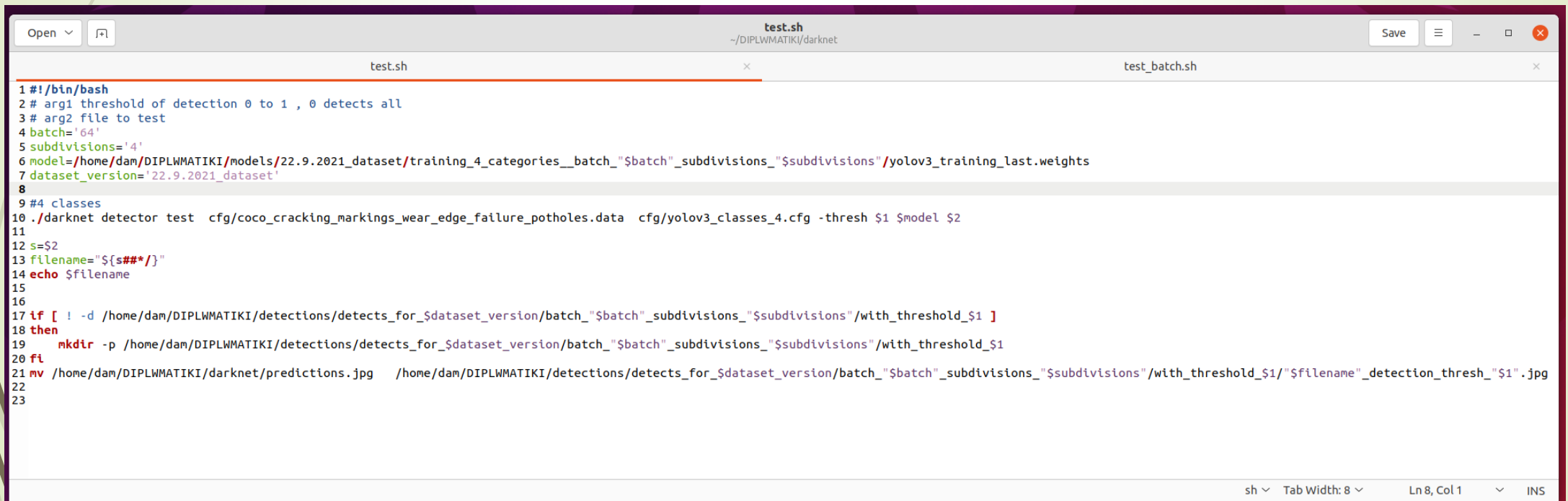


Δοκιμή μοντέλου και ανίχνευση

Η εκτέλεση των δοκιμών έγινε τοπικά σε υπολογιστή με λειτουργικό Linux όπου εγκαταστάθηκε το Darknet.

Η ανίχνευση κλάσεων σε μια εικόνα βάσει ενός μοντέλου εκπαίδευσης γίνεται με την ακόλουθη εντολή :

```
./darknet detector test <αρχείο πληροφοριών κλάσεων> <αρχείο ρυθμίσεων παραμέτρων εκτέλεσης> -thresh <επίπεδο threshold με τιμές από 0.0 έως 1.0> <μοντέλο εκπαίδευσης>  
<αρχείο που θα γίνει η δοκιμή ανίχνευσης των κλάσεων>
```



```
test.sh
~/DIPLWMATIKI/darknet
Save
test.sh
test_batch.sh

1 #!/bin/bash
2 # arg1 threshold of detection 0 to 1 , 0 detects all
3 # arg2 file to test
4 batch='64'
5 subdivisions='4'
6 model=/home/dam/DIPLWMATIKI/models/22.9.2021_dataset/training_4_categories__batch_"$batch"_subdivisions_"$subdivisions"/yolov3_training_last.weights
7 dataset_version='22.9.2021_dataset'
8
9 #4 classes
10 ./darknet detector test cfg/coco_cracking_markings_wear_edge_failure_potholes.data cfg/yolov3_classes_4.cfg -thresh $1 $model $2
11
12 s=$2
13 filename="{s##*/}"
14 echo $filename
15
16
17 if [ ! -d /home/dam/DIPLWMATIKI/detections/detects_for_${dataset_version}/batch_"$batch"_subdivisions_"$subdivisions"/with_threshold_$1 ]
18 then
19     mkdir -p /home/dam/DIPLWMATIKI/detections/detects_for_${dataset_version}/batch_"$batch"_subdivisions_"$subdivisions"/with_threshold_$1
20 fi
21 mv /home/dam/DIPLWMATIKI/darknet/predictions.jpg /home/dam/DIPLWMATIKI/detections/detects_for_${dataset_version}/batch_"$batch"_subdivisions_"$subdivisions"/with_threshold_$1/"$filename"_detection_thresh_"$1".jpg
22
23

sh Tab Width: 8 Ln 8, Col 1 INS
```

Δοκιμή μοντέλου και ανίχνευση

Στιγμιότυπο εκτέλεσης ανίχνευσης όπου φαίνεται η φόρτωση των βαρών του μοντέλου

```
0_gettyimages-655985038-170667a.jpg
layer  filters  size  input  output  BFLOPs
0 conv  32  3 x 3 / 1  608 x 608 x 3  ->  608 x 608 x 32  0.639 BFLOPs
1 conv  64  3 x 3 / 2  608 x 608 x 32  ->  304 x 304 x 64  3.407 BFLOPs
2 conv  32  1 x 1 / 1  304 x 304 x 64  ->  304 x 304 x 32  0.379 BFLOPs
3 conv  64  3 x 3 / 1  304 x 304 x 32  ->  304 x 304 x 64  3.407 BFLOPs
4 res   1  304 x 304 x 64  ->  304 x 304 x 64
5 conv  128 3 x 3 / 2  304 x 304 x 64  ->  152 x 152 x 128 3.407 BFLOPs
6 conv  64  1 x 1 / 1  152 x 152 x 128 ->  152 x 152 x 64  0.379 BFLOPs
7 conv  128 3 x 3 / 1  152 x 152 x 64  ->  152 x 152 x 128 3.407 BFLOPs
8 res   5  152 x 152 x 128 ->  152 x 152 x 128
9 conv  64  1 x 1 / 1  152 x 152 x 128 ->  152 x 152 x 64  0.379 BFLOPs
10 conv 128 3 x 3 / 1  152 x 152 x 64  ->  152 x 152 x 128 3.407 BFLOPs
11 res   8  152 x 152 x 128 ->  152 x 152 x 128
12 conv 256 3 x 3 / 2  152 x 152 x 128 ->  76 x 76 x 256  3.407 BFLOPs
13 conv 128 1 x 1 / 1  76 x 76 x 256  ->  76 x 76 x 128  0.379 BFLOPs
14 conv 256 3 x 3 / 1  76 x 76 x 128  ->  76 x 76 x 256  3.407 BFLOPs
15 res  12  76 x 76 x 256  ->  76 x 76 x 256
16 conv 128 1 x 1 / 1  76 x 76 x 256  ->  76 x 76 x 128  0.379 BFLOPs
17 conv 256 3 x 3 / 1  76 x 76 x 128  ->  76 x 76 x 256  3.407 BFLOPs
18 res  15  76 x 76 x 256  ->  76 x 76 x 256
```

Στιγμιότυπο εκτέλεσης ανίχνευσης όπου φαίνονται οι ανιχνεύσεις στην εξεταζόμενη εικόνα

```
Loading weights from /home/dam/DIPLWMATIKI/models/22.9.2021_dataset/training_4_categories__batch_64_subdivisions_8/yolo
t.weights...Done!
/home/dam/DIPLWMATIKI/dataset/images/0_gettyimages-655985038-170667a.jpg: Predicted in 10.163996 seconds.
potholes: 60%
potholes: 56%
markings_wear: 58%
cracking: 63%
cracking: 60%
cracking: 58%
cracking: 57%
```

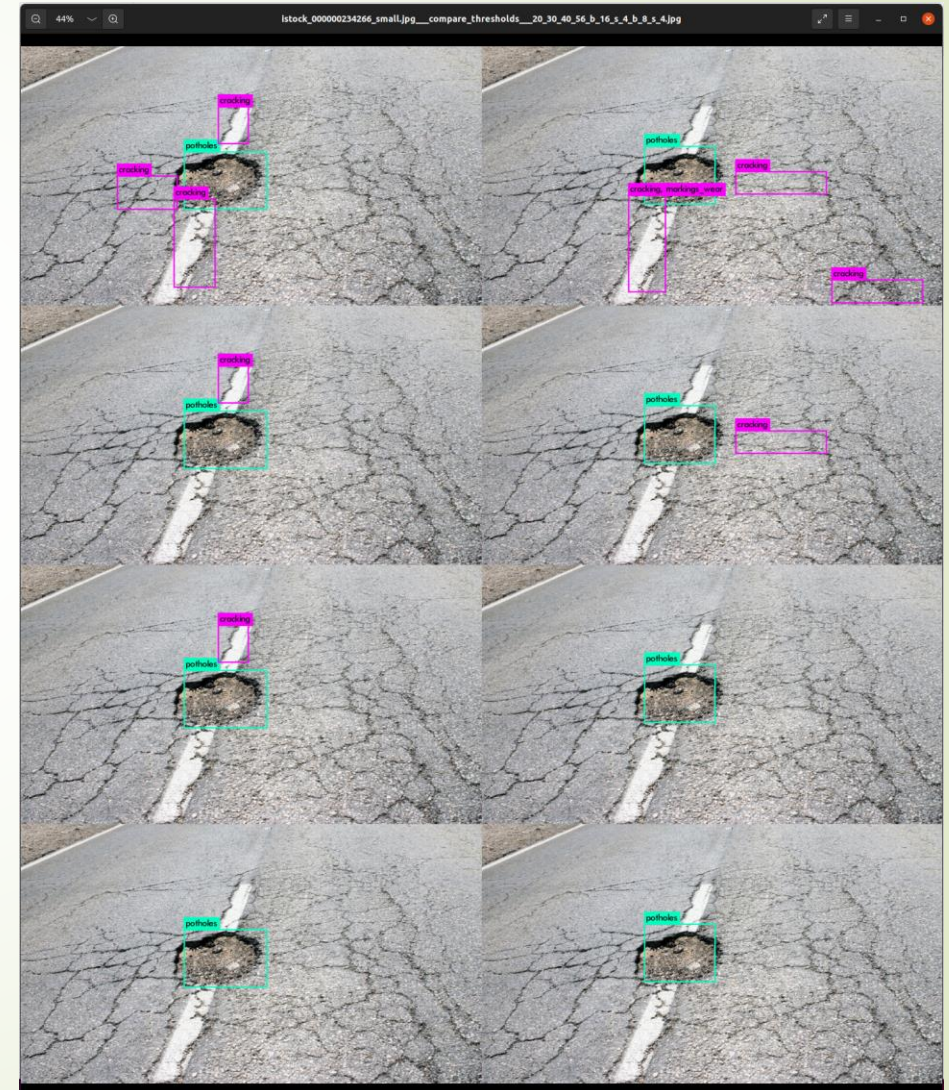
Δοκιμή μοντέλου και ανίχνευση

- Ανιχνεύσεις με διαφορετικά επίπεδα βεβαιότητας, ξεκινώντας από πάνω με threshold 20% , συνεχίζοντας με 30%, 40% και καταλήγοντας σε 56%. Παρατηρούμε ότι όσο χαμηλότερο το threshold η ανίχνευση είναι λιγότερο αυστηρή παρουσιάζοντας περιοχές που ανήκουν σε κάποια κλάση με μικρότερη βεβαιότητα.

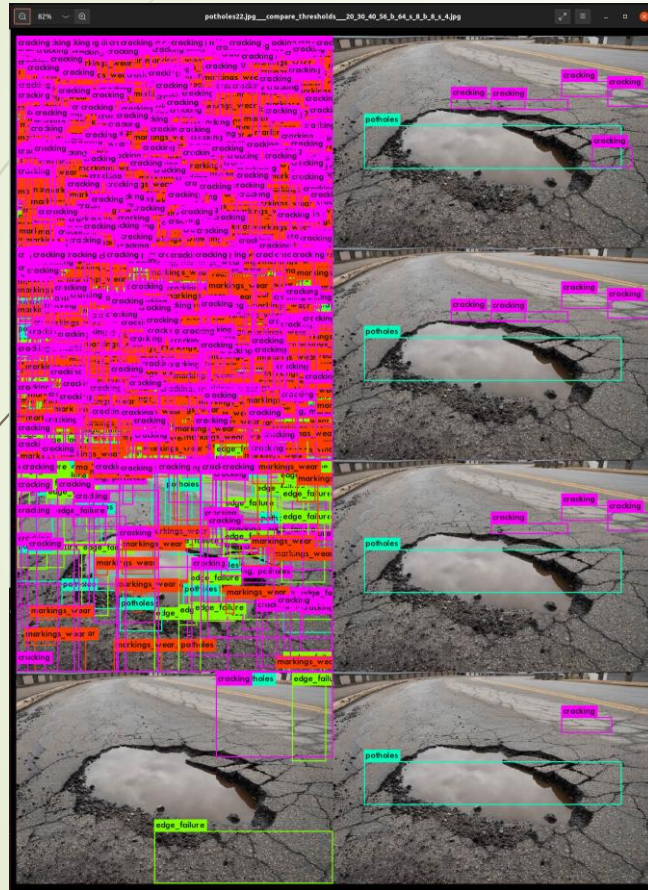


Δοκιμή μοντέλου και ανίχνευση

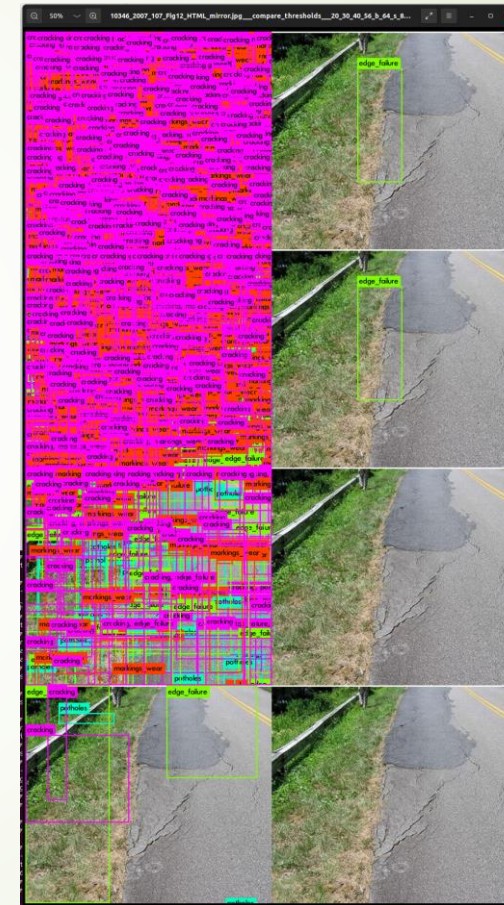
- Πλέγμα 2x4 εικόνων όπου στις στήλες έχουμε παραμέτρους batch=16, subdivision=4 για την 1η στήλη και batch=8, subdivision=4 για την 2η στήλη και τιμές threshold= 20,30,40,56 από πάνω προς τα κάτω
- Σε περιοχές μικρότερης βεβαιότητας η διαφορετική ομαδοποίηση (batch size) στο μοντέλο εκπαίδευσης, κατά την ανίχνευση επηρεάζει τα βάρη των μοντέλων με διαφορετικό τρόπο



Δοκιμή μοντέλου και ανίχνευση



Πλέγμα 2x4 εικόνων όπου στις στήλες έχουμε παραμέτρους batch=64, subdivision=8 για την 1η στήλη και batch=8, subdivision=4 για την 2η στήλη και τιμές threshold= 20,30,40,56 από πάνω προς τα κάτω



Πλέγμα 2x4 εικόνων όπου στις στήλες έχουμε παραμέτρους batch=64, subdivision=8 για την 1η στήλη και batch=8, subdivision=4 για την 2η στήλη και τιμές threshold= 20,30,40,56 από πάνω προς τα κάτω

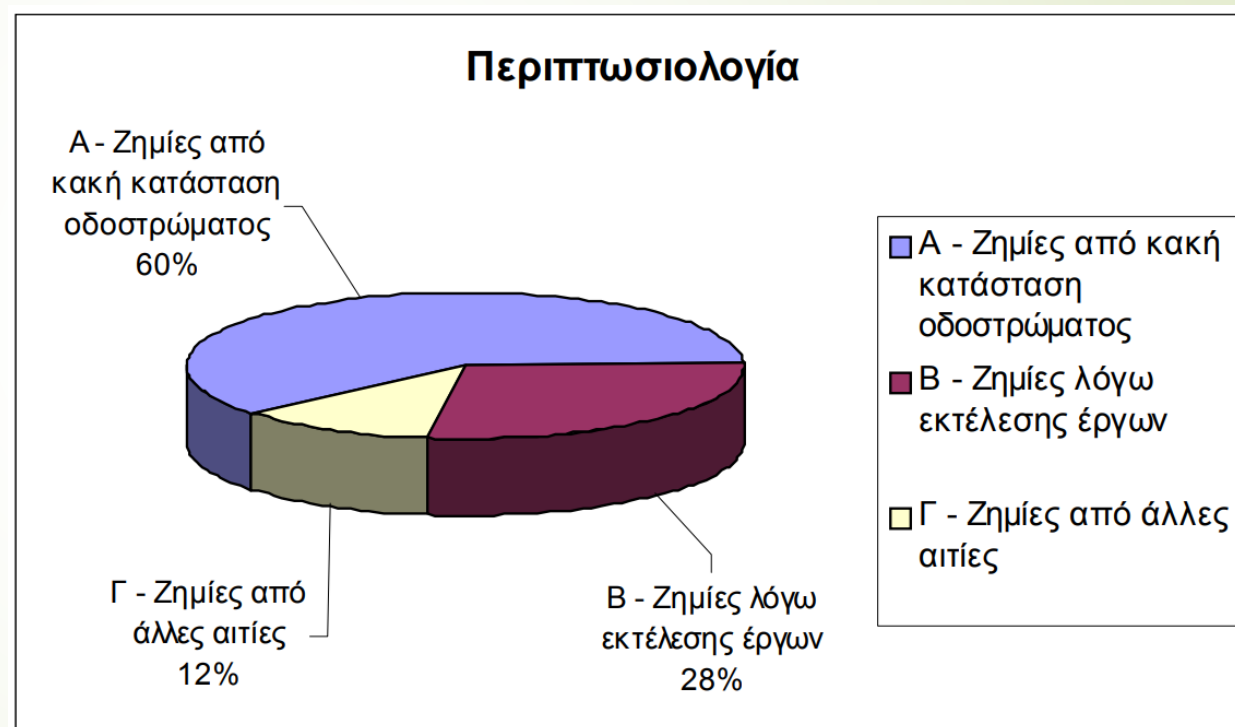
Δοκιμή μοντέλου και ανίχνευση

- Πλέγμα 2x4 εικόνων όπου στις στήλες έχουμε παραμέτρους $\text{batch}=64$, $\text{subdivision}=8$ για την 1η στήλη και $\text{batch}=64$, $\text{subdivision}=8$ για την 2η στήλη και τιμές $\text{threshold}=20,30,40,56$ από πάνω προς τα κάτω. Η αριστερή στήλη αναφέρεται σε ανιχνεύσεις με training από το υποσύνολο του dataset και η δεύτερη από το πλήρες



Τεχνο-οικονομική ανάλυση

Κατανομή από τις 207 αναφορές που το 2004 είχαν κατατεθεί στον Συνήγορο του Πολίτη



Τεχνο-οικονομικη ανάλυση

- Η συντήρηση των φθορών οδοστρώματος μπορεί να επεκτείνει τον εκτιμώμενο χρόνο ζωής των οδοστρωμάτων και να καθυστερήσει την ανάγκη για αποκατάσταση
- Η πραγματική διάρκεια ζωής μπορεί να διαφέρει εξαιτίας των τοπικών συνθηκών κυκλοφορίας, περιβάλλοντος και ποιότητας κατασκευής

Ρηγμάτωση λόγω κόπωσης				
Σοβαρότητα	Έκταση (% μήκους)	Πάχος ασφαλτικής στρώσης (mm)	Τυπικές επεμβάσεις	Εκτιμώμενη διάρκεια ζωής (χρόνια)
Χαμηλή	1-10	<200	“Μπάλωμα” ή τοπική πλήρωση ρωγμών	5
Μέση	11-25	>200	Επίστρωση 50mm	12
Μέση	11-25	<200	Επίστρωση 75mm	12
Μέση	25-50	>200	Επίστρωση 75mm	12
Μέση	25-50	<200	Επίστρωση 150mm	12
Υψηλή	>50	>200	Επίστρωση 100mm	12
Υψηλή	>50	<200	Ανακατασκευή	15

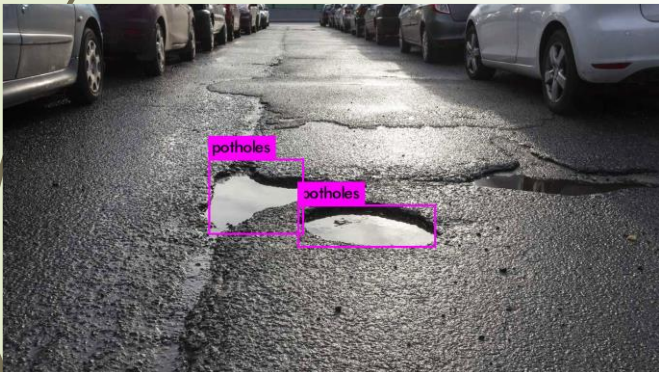
Εγκάρσια ρηγμάτωση				
Σοβαρότητα	Έκταση (απόσταση σε m)	Τυπικές επεμβάσεις		Εκτιμώμενη διάρκεια ζωής (χρόνια)
Χαμηλή	>15	Καθαρισμός και σφράγιση		8
Μέση	10-15	Καθαρισμός και σφράγιση		5
Μέση	10-15	Καθαρισμός και σφράγιση με ασφαλτόμιγμα τύπου Slurry		3
Υψηλή	<10	Καθαρισμός/σφράγιση και επίστρωση 50mm		10

Παραμορφώσεις - Αυλάκωση				
Σοβαρότητα (mm)	Έκταση (% μήκους)	Τυπικές επεμβάσεις		Εκτιμώμενη διάρκεια ζωής (χρόνια)
<6	>25	Καμία επέμβαση		-
6-12	>25	Ισοπέδωση		7
12-25	>25	Ισοπέδωση		4
>25	>50	Φρεζάρισμα και επίστρωση 50mm		10

Συμπεράσματα – Μελλοντική έρευνα

Μοντέλα τύπου YOLO framework

↑ Απαιτήσεις για μικρότερη επεξεργαστική ισχύ
↓ Λιγότερο λεπτομερής ανίχνευση (επισήμανση εντός λεπτομερούς πλαισίου)



Προηγμένα μοντέλα και μέθοδοι που χρησιμοποιούν Semantic Segmentation

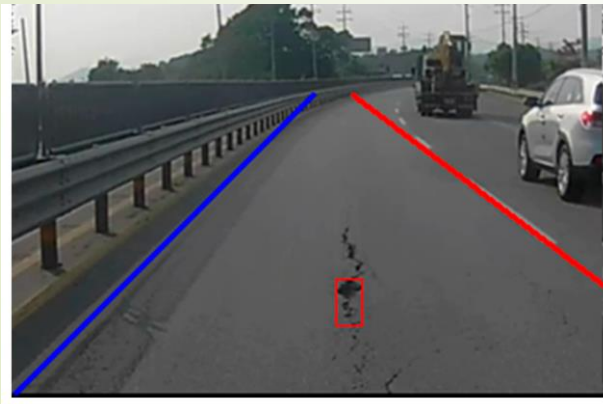
↑ η ανάλυση γίνεται εξαντλητικά σε επίπεδο pixel
↓ Απαιτήσεις για μεγαλύτερη επεξεργαστική ισχύ



Συμπεράσματα – Μελλοντική έρευνα

Μοντέλα τύπου YOLO framework

- Κατάλληλα για εφαρμογές πραγματικού χρόνου και ανάγκης άμεσης λήψης αποφάσεων
- Παραδείγματα εφαρμογών είναι η ανίχνευση ανωμαλιών σε πραγματικό χρόνο από ένα αυτοκίνητο που κινείται εντός αυτοκινητοδρόμου



Προηγμένα μοντέλα και μέθοδοι που χρησιμοποιούν Semantic Segmentation

- Κατάλληλα για εφαρμογές εκτέλεσης αναλυτικότερων εργασιών και μεγάλου όγκου δεδομένων και όχι απαραίτητα πραγματικού χρόνου
- Παραδείγματα εφαρμογών είναι η αναλυτική σκιαγράφιση των ανωμαλιών ενός αυτοκινητοδρόμου από την διαχειρίστρια εταιρία του αυτοκινητοδρόμου για την τεχνοοικονομική ανάλυση και την στοχευμένη δρομολόγηση των απαιτούμενων εργασιών επισκευής χρήση αυτόνομων μηχανημάτων επιδιόρθωσης





Ερωτήσεις;

