Κατάτμηση κυτταρολογικών εικόνων από τεστ ΠΑΠ με το μοντέλο splinedist

Η Μεταπτυχιακή Διπλωματική Εργασία

υποβάλλεται στην ορισθείσα

από τη Συνέλευση

του Τμήματος Μηχανικών Η/Υ και Πληροφορικής

Εξεταστική Επιτροπή

από την

Φωτεινή Πλακούτση

ως μέρος των υποχρεώσεων για την απόκτηση του

ΔΙΠΛΩΜΑΤΟΣ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΣΤΗ ΜΗΧΑΝΙΚΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΜΕ ΕΙΔΙΚΕΥΣΗ

ΣΤΗΝ ΕΠΙΣΤΗΜΗ ΚΑΙ ΜΗΧΑΝΙΚΗ ΔΕΔΟΜΕΝΩΝ

Πανεπιστήμιο Ιωαννίνων

Πολυτεχνική Σχολή

Ιωάννινα 2025

Εξεταστική Επιτροπή:

- **Χριστόφορος Νίκου**, Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων (Επιβλέπων)
- **Λυσίμαχος-Παύλος Κόντης**, Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
- **Γεώργιος Μανής**, Αναπλ. Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων

Αφιερωση

Θα ήθελα να αφιερώσω τη διπλωματική μου εργασία στην οικογένειά μου, η οποία με στήριξε σε όλη τη διάρκεια αυτού του δύσκολου αλλά ευχάριστου ταξιδιού μου.

.

Ευχαριστιες

Αρχικά, θα ήθελα να ευχαριστήσω θερμά, τους καθηγητές μου, κ.Χριστόφορο Νίκου και κ.Πλησίτη Μαρίνα, για την σωστή καθοδήγηση και επίβλεψη που μου παρείχαν σε όλη τη διάρκεια της διπλωματικής μου εργασίας. Στη συνέχεια, θα ήθελα να ευχαριστήσω την οικογένειά μου, και συγκεκριμένα, τους γονείς μου,την αδερφή μου, και άλλα συγγενικά και αγαπημένα μου πρόσωπα, οι οποίοι με υποστήριξαν σε όλους τους τομείς και κυρίως ψυχολογικά, σε αυτή τη δύσκολη διαδρομή που είχα να διανύσω.

Περιεχομενα

Κατάλογος Σχημάτων

Κατάλογος Πινάκων

Περιληψη

Η ακριβής αναγνώριση και κατάτμηση κυττάρων σε εικόνες μικροσκοπίου είναι απαραίτητη για την αξιόπιστη ποσοτική ανάλυση σε επίπεδο μεμονωμένων κυττάρων. Αυτή η διαδικασία επιτρέπει τη μελέτη των μορφολογικών χαρακτηριστικών των κυττάρων και οδηγεί στον εντοπισμό μη φυσιολογικών και παθολογικών κυττάρων.

Οι τεχνικές βαθιάς μάθησης έχουν συνεισφέρει σημαντικά στην επεξεργασία κυτταρολογικών εικόνων, οι οποίες παρουσιάζουν πολλές προκλήσεις , όπως χαμηλή αντίθεση, μεταβαλλόμενες κυτταρικές μορφές, επικάλυψηκυττάρων. Μια από τις πρόσφατες τεχνικές είναι το μοντέλο StarDist, που βασίζεται σε κυρτά αστεροειδή πολύγωνα (star-convex polygons), και έχει σημειώσει αξιόλογα αποτελέσματα. Ωστόσο, παρουσιάζει περιορισμούς στην κατάτμηση μη κυρτών αντικειμένων.

Στην παρούσα εργασία, μελετάται το μοντέλο SplineDist, που είναι μια επέκταση του StarDist για την αντιμετώπιση αυτών των προκλήσεων. Το SplineDist είναι ένα νέο πλαίσιο κατάτμησης που αντικαθιστά τα άκαμπτα πολύγωνα με εύκαμπτες παραμετρικές καμπύλες τύπου spline. Βασίζεται σε αρχιτεκτονική U-Net και προβλέπει ένα σταθερό σύνολο spline σημείων ελέγχου, το καθένα ορισμένο μέσω γωνίας και ακτίνας ως προστο κέντρο του κυττάρου. Αυτό επιτρέπει μεγαλύτερη ευελιξία στην περιγραφή των ορίων, ακόμη και σε πολύπλοκες μορφές.

Για την αξιολόγηση του μοντέλου πραγματοποιήθηκαν πειράματα στο σύνολο δεδομένων SIPaKMeD, το οποίο περιλαμβάνει 4049 εικόνες κυττάρων. Τα αποτελέσματα δείχνουν ότι το μοντέλο SplineDist επιτυγχάνει υψηλή ακρίβεια στην κατάτμηση, με ταυτόχρονα χαμηλό αριθμό παραμέτρων. Τα παραγόμενα περιγράμματα είναι πιο ομαλά και ακριβή, ακόμη και με μικρό αριθμό σημείων ελέγχου.

Συνολικά, το μοντέλο SplineDist συνδυάζει κλασικές γεωμετρικές μεθόδους με σύγχρονες τεχνικές βαθιάς μάθησης, προσφέροντας ένα αποδοτικό και ευέλικτο εργαλείο για την αυτόματη κατάτμηση βιολογικών αντικειμένων σε εικόνες μικροσκο-

πίου.

Extended Abstract

Accurate cell detection and segmentation in microscopy images is essential for reliable quantitative analysis at the single-cell level. This process enables the study of cellular morphological characteristics and facilitates the identification of abnormal and pathological cells.

Deep learning techniques have significantly contributed to cytological image processing, which presents numerous challenges such as low contrast, variable cellular shapes, and cell overlap. One of the recent techniques is the StarDist model, which is based on star-convex polygons and has achieved notable results. However, it has limitations when it comes to segmenting non-convex objects.

In this study, we examine the SplineDist model, which is an extension of StarDist designed to address these challenges. SplineDist is a novel segmentation framework that replaces rigid polygons with flexible spline-type parametric curves. It is based on a U-Net architecture and predicts a fixed set of spline control points, each defined by an angle and a radius relative to the cell center. This allows for greater flexibility in boundary representation, even in complex shapes.

To evaluate the model, experiments were conducted on the SIPaKMeD dataset, which includes 4,049 cell images. The results show that SplineDist achieves high segmentation accuracy while maintaining a low parameter count. The generated contours are smoother and more precise, even with a small number of control points.

Overall, the SplineDist model combines classical geometric methods with modern deep learning techniques, offering an efficient and flexible tool for automatic segmentation of biological objects in microscopy images.

Κεφαλαιο 1

Εισαγωγη

Η ανάλυση δεδομένων σε μία βιολογική έρευνα που εξετάζονται με μικροσκόπιο συχνά εξαρτάται από τον ακριβή εντοπισμό και την οριοθέτηση μεμονωμένων αντικειμένων μέσα σε μία εικόνα, μια διαδικασία γνωστή ως κατάτμηση στιγμιότυπου. Η κατάτμηση επιτρέπει την εξαγωγή λεπτομερών πληροφοριών σε επίπεδο μονοκυττάρου, όπως η έκφραση πρωτεΐνης και τα μορφολογικά χαρακτηριστικά. Επειδή υπάρχουν τεράστιες ποσότητες δεδομένων απεικόνισης που παράγονται στα σύγχρονα βιολογικά πειράματα, η ανάπτυξη τεχνικών αυτοματοποιημένης κατάτμησης αποτελεί επίκεντρο της ανάλυσης βιοϊατρικών εικόνων για δεκαετίες. Τα τελευταία δέκα χρόνια, η βαθιά μάθηση έχει αλλάξει δραματικά αυτόν τον τομέα. Τα συνελικτικά νευρωνικά δίκτυα (Convolutional Neural Networks, CNN), ιδιαίτερα το ευρέως χρησιμοποιούμενο μοντέλο U-Net, έχουν επιδείξει αξιοσημείωτη ακρίβεια στις εργασίες κατάτμησης . Μεταξύ των πιο επιτυχημένων εργαλείων που βασίζονται σε βαθιά μάθηση είναι ο αλγόριθμος StarDist, ο οποίος ξεχωρίζει για την αποτελεσματικότητά του και την ευκολία χρήσης του. Ο StarDist χρησιμοποιεί μια αρχιτεκτονική U-Net για να προβλέψει ένα σταθερό σύνολο σημείων κατά μήκος του περιγράμματος ενός αντικειμένου, κατασκευάζοντας το τελικό του σχήμα ως ένα κυρτό πολύγωνο με αστέρι. Αυτή η αναπαράσταση έχει κοινές ομοιότητες με τις παραμετρικές μεθόδους περιγράμματος B-spline, οι οποίες περιγράφουν τα όρια αντικειμένων ως ομαλές, συνεχείς καμπύλες βελτιστοποιημένες χρησιμοποιώντας τεχνικές ελαχιστοποίησης ενέργειας. Ενώ αυτές οι προσεγγίσεις που βασίζονται σε spline παρέχουν ακριβείς αναπαραστάσεις περιγράμματος, η εξάρτησή τους από χειροποίητα κριτήρια βελτιστοποίησης τις καθιστά λιγότερο προσιτές σε μη ειδικούς. Για να αντιμε-

τωπιστεί αυτός ο περιορισμός, προηγούμενες εργασίες [10] εισήγαγαν μια μέθοδο βασισμένη στο CNN για την άμεση πρόβλεψη περιγραμμάτων με παρεμβολή spline από δυαδικές εικόνες που περιέχουν μεμονωμένα αντικείμενα. Ωστόσο, αυτή η αρχική προσέγγιση δεν είχε την ικανότητα να χειρίζεται περίπλοκες εικόνες πολλαπλών αντικειμένων. Βασιζόμενοι σε αυτό το θεμέλιο, παρουσιάζουμε τον SplineDist, έναν αλγόριθμο κατάτμησης που ενσωματώνει τα πλεονεκτήματα του StarDist με αυτά της μοντελοποίησης που βασίζεται σε spline. Ο SplineDist διατηρεί την ισχυρή απόδοση ανίχνευσης του StarDist ενώ ενισχύει την ποιότητα τμηματοποίησης μέσω της προσέγγισης spline. Σε αντίθεση με τον προκάτοχό του, ο SplineDist μπορεί να επεξεργαστεί ακατέργαστες εικόνες που περιέχουν πολλά αντικείμενα, παρέχοντας μια πλήρως αυτοματοποιημένη λύση τμηματοποίησης από άκρο σε άκρο. Επιπλέον, εξαλείφει την εξάρτηση του StarDist στους περιορισμούς κυρτότητας αστεριών, διευρύνοντας τη δυνατότητα εφαρμογής του. Γεφυρώνοντας τη βαθιά μάθηση και τις παραδοσιακές μεθόδους που βασίζονται σε spline, ο SplineDist προσφέρει μια ισχυρή και ευέλικτη εναλλακτική λύση για την κατάτμηση βιοϊατρικών εικόνων. Στην παρούσα εργασία, εφαρμόζεται ο αλγόριθμος SplineDist στο σύνολο δεδομένων SIPakMeD, το οποίο αποτελείται από εικόνες μεμονωμένων κυττάρων. Σκοπός είναι να πραγματοποιηθεί κατάτμηση των κυττάρων σε μικροσκοπικές εικόνες, μοντελοποιώντας το περίγραμμα των αντικειμένων ως παραμετρικές καμπύλες, ακόμη και σε αντικείμενα που δεν αποτελούν αστεροειδή σχήματα.

Η εργασία είναι δομημένη σε εφτά κεφάλαια, όπου το πρώτο κεφάλαιο αποτελεί μία εισαγωγή, στο δεύτερο κεφάλαιο γίνεται αναφορά στη μηχανική μάθηση και στα βαθιά νευρωνικά δίκτυα και στο τρίτο κεφάλαιο περιγράφεται η δομή των μοντέλων που χρησιμοποιήθηκαν. Επιπλέον, στο τέταρτο κεφάλαιο περιγράφεται το σύνολο δεδομένων το οποίο χρησιμοποιήθηκε, ενώ στο πέμπτο κεφάλαιο αναφέρεται η διαδικασία που ακολουθήθηκε στην εκτέλεση των πειραμάτων, των οποίων τα αποτελέσματα αναφέρονται στο έκτο κεφάλαιο. Στο έβδομο κεφάλαιο αναφέρονται τα συμπεράσματα και οι προτάσεις για μελλοντική εργασία. Τέλος, υπάρχουν τρία παραρτήματα, όπου εξηγούνται αναλυτικά ορισμένοι δείκτες απόδοσης του μοντέλου.

Κεφαλαιο 2

Μηχανικη Μαθηση και Βαθια Νευρωνικα Δικτυα

Σε αυτό το κεφάλαιο γίνεται μία εισαγωγή στη μηχανική μάθηση και στη βαθιά μάθηση. Ειδικότερα, γίνεται μία αναφορά στα νευρωνικά δίκτυα και στα συνελικτικά νευρωνικά δίκτυα.

2.1 Μηχανική Μάθηση

Η μηχανική μάθηση είναι ένας κλάδος της τεχνητής νοημοσύνης ο οποίος δίνει τη δυνατότητα στους υπολογιστές και στις μηχανές να μιμούνται τον τρόπο με τον οποίο μαθαίνουν οι άνθρωποι να εκτελούν εργασίες αυτόνομα και να βελτιώνουν την απόδοση και την ακρίβεια, μαθαίνοντας όλο και περισσότερα δεδομένα. Οι πιο κοινοί αλγόριθμοι μηχανικής μάθησης είναι οι εξής : νευρωνικά δίκτυα, γραμμική παλινδρόμηση, λογιστική παλινδρόμηση, ομαδοποίηση(clustering), δέντρα απόφασης και random forest [11].

2.1.1 Επιβλεπόμενη Μάθηση (Supervised Learning)

Η επιβλεπόμενη μάθηση αφορά τη διαδικασία εκπαίδευσης αλγορίθμων για την ταξινόμηση δεδομένων και την ακριβή πρόβλεψη των αποτελεσμάτων, μέσω ενός συνόλου δεδομένων. Καθώς τα δεδομένα εισόδου περνάνε μέσα στο μοντέλο, αυτό προσαρμόζει τα βάρη του ανάλογα. Ορισμένες δημοφιλείς μέθοδοι επιβλεπόμενης μάθησης είναι: νευρωνικά δίκτυα, Naive Bayes, υποστηρικτικές μηχανές διανυσμάτων (SVM), γραμμική παλινδρόμηση, λογιστική παλινδρόμηση και τυχαία δάση (random forest). Από αυτές, μόνο τα νευρωνικά δίκτυα βασίζονται σε πολυεπίπεδη δομή εμπνευσμένη από τον ανθρώπινο εγκέφαλο [11].

2.1.2 Μη επιβλεπόμενη μάθηση (Unsupervised Learning)

Η μάθηση χωρίς επίβλεψη, χρησιμοποιεί αλγορίθμους μηχανικής μάθησης προκειμένου να αναλύσει και να ομαδοποιήσει σύνολα δεδομένων χωρίς ετικέτες. Αυτοί οι αλγόριθμοι έχουν ως στόχο την αναγνώριση ορισμένων χαρακτηριστικών ή μοτίβων, χωρίς την παρέμβαση ανθρώπινης φύσης . Η ικανότητα της μάθησης χωρίς επίβλεψη είναι να αναγνωρίζει ομοιότητες και διαφορές μέσω των πληροφοριών που λαμβάνει και επίσης χρησιμοποιείται για να μειώσει τον αριθμό των χαρακτηριστικών σε ένα μοντέλο κατά τη διάρκεια της μείωσης των διαστάσεων, όπως για παράδειγμα συμβαίνει στην μέθοδο Principal Component Analysis (PCA). Άλλοι αλγόριθμοι μη επιβλεπόμενης μάθησης, που περιέχουν νευρωνικά δίκτυα, είναι οι εξής : k-means και πιθανοτικές μέθοδοι ομαδοποίησης [11].

2.2 Νευρωνικά Δίκτυα

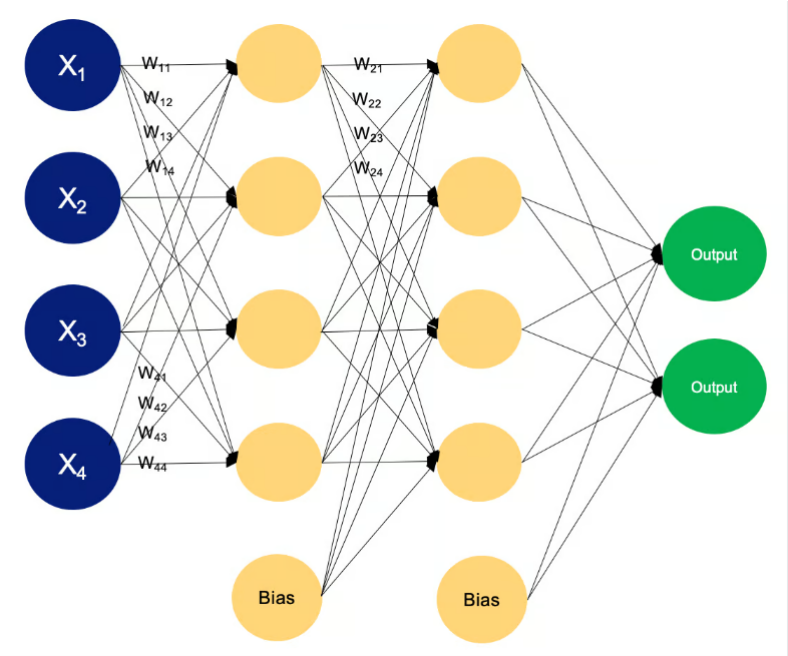
Τα νευρωνικά δίκτυα είναι θεμελιώδη εργαλεία στη μηχανική μάθηση, που τροφοδοτούν πολλούς αλγορίθμους και εφαρμογές τελευταίας τεχνολογίας σε διάφορους τομείς, συμπεριλαμβανομένης της όρασης υπολογιστών, της επεξεργασίας φυσικής γλώσσας, της ρομποτικής και άλλων. Ένα νευρωνικό δίκτυο αποτελείται από συνδεδεμένους κόμβους, που ονομάζονται νευρώνες, οργανωμένοι σε στρώματα. Κάθε νευρώνας λαμβάνει σήματα εισόδου, εκτελεί έναν υπολογισμό σε αυτά χρησιμοποιώντας μια συνάρτηση ενεργοποίησης και παράγει ένα σήμα εξόδου που μπορεί να περάσει σε άλλους νευρώνες του δικτύου. Μια συνάρτηση ενεργοποίησης καθορίζει την έξοδο ενός νευρώνα δεδομένης της εισόδου του. Αυτές οι συναρτήσεις εισάγουν μη γραμμικότητα στο δίκτυο, επιτρέποντάς του να μάθει πολύπλοκα μοτίβα στα δεδομένα. Το δίκτυο συνήθως οργανώνεται σε επίπεδα, ξεκινώντας από το επίπεδο εισόδου, όπου εισάγονται δεδομένα. Ακολουθούν κρυφά επίπεδα όπου εκτελούνται οι υπολογισμοί και τέλος, το επίπεδο εξόδου όπου λαμβάνονται προβλέψεις ή αποφάσεις. Το πολυεπίπεδο perceptron (Multilayer Perceptron, MLP) είναι ένας τύπος νευρωνικού δικτύου τροφοδοσίας που αποτελείται από πλήρως συνδεδεμένους νευρώνες με μη γραμμικό είδος συνάρτησης ενεργοποίησης. Χρησιμοποιείται ευρέως για τη διάκριση δεδομένων που δεν διαχωρίζονται γραμμικά.Το επίπεδο εισόδου αποτελείται από κόμβους ή νευρώνες που λαμβάνουν τα αρχικά δεδομένα εισόδου. Κάθε νευρώνας αντιπροσωπεύει ένα χαρακτηριστικό ή μια διάσταση των δεδομένων εισόδου. Ο αριθμός των νευρώνων στο επίπεδο εισόδου καθορίζεται από τη διάσταση των δεδομένων εισόδου. Μεταξύ των επιπέδων εισόδου και εξόδου, μπορεί να υπάρχουν ένα ή περισσότερα στρώματα νευρώνων. Κάθε νευρώνας σε ένα κρυφό στρώμα λαμβάνει εισόδους από όλους τους νευρώνες του προηγούμενου στρώματος (είτε το στρώμα εισόδου είτε άλλο κρυφό στρώμα) και παράγει μια έξοδο που περνά στο επόμενο επίπεδο. Ο αριθμός των κρυφών επιπέδων και ο αριθμός των νευρώνων σε κάθε κρυφό στρώμα είναι υπερπαράμετροι που πρέπει να προσδιοριστούν κατά τη φάση σχεδιασμού του μοντέλου. Αυτό το επίπεδο αποτελείται από νευρώνες που παράγουν την τελική έξοδο του δικτύου. Ο αριθμός των νευρώνων στο επίπεδο εξόδου εξαρτάται από τη φύση της εργασίας. Στη δυαδική ταξινόμηση, μπορεί να υπάρχουν είτε ένας είτε δύο νευρώνες ανάλογα με τη συνάρτηση ενεργοποίησης και αντιπροσωπεύουν την πιθανότητα να ανήκουν σε μία κατηγορία. ενώ σε εργασίες ταξινόμησης πολλαπλών κλάσεων, μπορεί να υπάρχουν

πολλοί νευρώνες στο επίπεδο εξόδου. Οι νευρώνες σε γειτονικά στρώματα είναι πλήρως συνδεδεμένοι μεταξύ τους. Κάθε σύνδεση έχει ένα σχετικό βάρος, το οποίο καθορίζει την αντοχή της σύνδεσης. Αυτά τα βάρη μαθαίνονται κατά τη διάρκεια της διαδικασίας εκπαίδευσης. Εκτός από τους νευρώνες εισόδου και τους κρυφούς νευρώνες, κάθε στρώμα (εκτός από το στρώμα εισόδου) συνήθως περιλαμβάνει έναν νευρώνα πόλωσης που παρέχει μια σταθερή είσοδο στους νευρώνες στο επόμενο στρώμα. Οι νευρώνες πόλωσης έχουν το δικό τους βάρος που σχετίζεται με κάθε σύνδεση, το οποίο μαθαίνεται επίσης κατά τη διαδικασία εκπαίδευσης. Τυπικά, κάθε νευρώνας στα κρυφά επίπεδα και στο επίπεδο εξόδου εφαρμόζει μια συνάρτηση ενεργοποίησης στο σταθμισμένο άθροισμα των εισόδων του. Οι κοινές συναρτήσεις ενεργοποίησης περιλαμβάνουν σιγμοειδή (sigmoid), εφαπτομένη (tanh), συνάρτηση ReLU (Rectified Linear Unit) και softmax. Αυτές οι συναρτήσεις εισάγουν μη γραμμικότητα στο δίκτυο, επιτρέποντάς του να μάθει πολύπλοκα μοτίβα στα δεδομένα. Τα MLP εκπαιδεύονται χρησιμοποιώντας τον αλγόριθμο backpropagation, ο οποίος υπολογίζει τις διαβαθμίσεις μιας συνάρτησης απώλειας σε σχέση με τις παραμέτρους του μοντέλου και ενημερώνει τις παραμέτρους επαναληπτικά για να ελαχιστοποιήσει την απώλεια. Για κάθε νευρώνα σε ένα κρυφό στρώμα ή στο επίπεδο εξόδου, υπολογίζεται το σταθμισμένο άθροισμα των εισόδων του. Αυτό περιλαμβάνει τον πολλαπλασιασμό κάθε εισόδου με το αντίστοιχο βάρος της, και την πρόσθεση της πόλωσης, όπως φαίνεται στον επόμενο τύπο.

$$\text{Σταθμισμένο Άθροισμα} = \sum_{i=1}^{n} w_i x_i + b \tag{2.1}$$

όπου n είναι ο συνολικός αριθμός των συνδέσεων εισόδου, wi είναι το βάρος για την i-οστή είσοδο και xi είναι η i-οστή τιμή εισόδου [1].

Η δομή ενός πολυεπίπεδου perceptron απεικονίζεται στο σχήμα 2.1.



Σχήμα 2.1: Παράδειγμα ενός πολυεπίπεδου perceptron (Multilayer Perceptron, MLP) με δύο κρυμμένα επίπεδα [1]

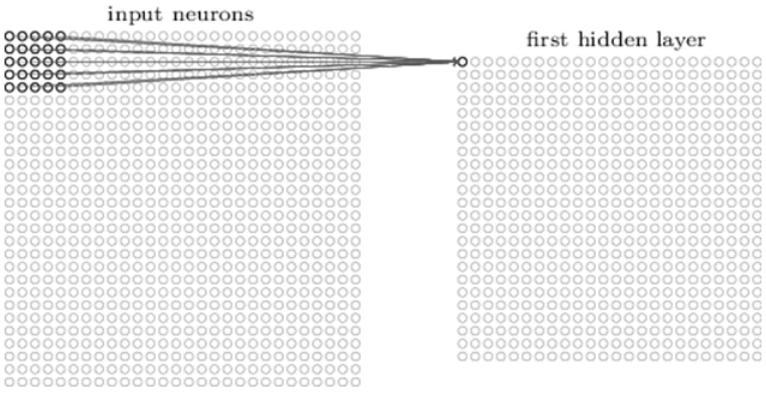
Στη συνέχεια, αναφέρεται ένα παράδειγμα ταξινόνησης εικόνας μέσω νευρωνικών δικτύων. Ένας υπολογιστής όταν δέχεται για είσοδο μια εικόνα, βλέπει έναν πίνακα από εικονοστοιχεία. Σκοπός είναι να υπολογιστεί μία πιθανότητα η οποία θα ταξινομήσει την εικόνα σε μία κατηγορία, ανάλογα με την ετικέτα που θα της δώσει ο υπολογιστής. Τα νευρωνικά δίκτυα επιτρέπουν στον υπολογιστή να διαφοροποιήσει τις εικόνες που δέχεται και να ανιχνεύσει τα χαρακτηριστικά τα οποία απεικονίζονται. Για παράδειγμα, ο υπολογιστής πρέπει να είναι σε θέση να αναγνωρίσει τα χαρακτηριστικά μιας εικόνας που απεικονίζει ένα σκύλο, έναντι μιας εικόνας που απεικονίζει μία γάτα. Έτσι, ο υπολογιστής, αναζητά χαμηλού επιπέδου χαρακτηριστικά, όπως καμπύλες και ακμές και στη συνέχεια, μέσω των νευρωνικών δικτύων δημιουργούνται αφηρημένες έννοιες. Συνοπτικά, οι εικόνες περνάνε μέσα από συνελικτικά, μη γραμμικά, ομαδοποιημένα και στενά συνδεδεμένα επίπεδα και προκύπτει το τελικό αποτέλεσμα, το οποίο μπορεί να είναι είτε μία μονή κλάση, είτε μία πιθανότητα κλάσεων που περιγράφει καλύτερα την εικόνα [2].

2.3 Βαθιά Μάθηση (Deep Learning)

Η βαθιά μάθηση αποτελεί έναν κλάδο της μηχανικής μάθησης, η οποία βασίζεται σε νευρωνικά δίκτυα με πολλά επίπεδα, τα οποία ονομάζονται βαθιά νευρωνικά δίκτυα. Η διαφορά της μηχανικής μάθησης και της βαθιάς μάθησης βρίσκεται στη δομή των νευρωνικών δικτύων. Στη μηχανική μάθηση τα νευρωνικά είναι απλά και αποτελούνται από ένα ή δύο επίπεδα, ενώ στη βαθιά μάθηση τα επίπεδα είναι πολλά περισσότερα προκειμένου να εκπαιδεύσουν το μοντέλο.Τα βαθιά μοντέλα μάθησης χρησιμοποιούν επιβλεπόμενη μάθηση. Με αυτό τον τρόπο, μπορούν να εξάγουν χαρακτηριστικά και σχέσεις που χρειάζονται, προκειμένου να διεξαχθούν ακριβή αποτελέσματα, από μη δομημένα δεδομένα. Τα νευρωνικά δίκτυα μιμούνται τον ανθρώπινο εγκέφαλο, μέσα από έναν συνδυασμό δεδομένων και βαρών. Τα βαθιά νευρωνικά δίκτυα αποτελούνται από πολλά επίπεδα συνδεδεμένων κόμβων. Τα στρώματα εισόδου και εξόδου στα βαθιά νευρωνικά δίκτυα ονομάζονται ορατά στρώματα. Το στρώμα εισόδου είναι αυτό στο οποίο τα βαθιά μοντέλα δέχονται τα δεδομένα και για επεξεργασία και το στρώμα εξόδου είναι αυτό όπου γίνεται η τελική πρόβλεψη ή ταξινόμηση. Υπάρχουν διάφορες κατηγορίες βαθιάς μηχανικής μάθησης. Για παράδειγμα, CNNs, ή αλλιώς συνελικτικά νευρωνικά δίκτυα, RNNs, ή αλλιώς επαναλαμβανόμενα νευρωνικά δίκτυα, autoencoders, GANs, ή αλλιώς ανταγωνιστικά νευρωνικά δίκτυα και άλλα πολλά [12].

2.4 Συνελικτικό Επίπεδο

Το συνελικτικό επίπεδο αποτελεί το πρώτο επίπεδο σε ένα συνελικτικό νευρωνικό δίκτυο. Προκειμένου να κατανοηθεί καλύτερα το συνελικτικό επίπεδο, φανταζόμαστε μία λάμψη η οποία ξεκινάει από το πάνω αριστερό μέρος της εικόνας και διαχέεται σε ολόκληρη την εικόνα. Στη μηχανική μάθηση, αυτή η λάμψη καλείται φίλτρο ή πυρήνας και η περιοχή την οποία διαπερνά κάθε φορά το φίλτρο ονομάζεται receptive filed (δεκτική περιοχή). Αυτό το φίλτρο είναι ένας πίνακας με αριθμούς , οι οποίοι ονομάζονται βάρη ή παράμετροι. Αξίζει να σημειωθεί ότι το βάθος του φίλτρου πρέπει να είναι ίδιο με το βάθος της περιοχής που αρχικά έχει καλύψει το φίλτρο, έτσι ώστε να σιγουρευτούμε ότι το μαθηματικό κομμάτι δουλεύει καλά. Για παράδειγμα, εάν στην είσοδο που δέχεται το νευρωνικό, οι διαστάσεις του φίλτρου είναι 5*5, τότε πρέπει οπωσδήποτε και στην τελευταία περιοχή που θα καλύψει το

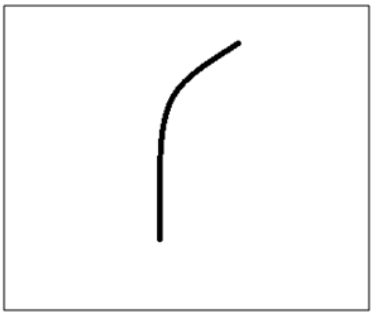


Σχήμα 2.2: Απεικόνιση της διάχυσης ενός 5*5 Φίλτρου Και κατασκευή ενός χάρτη ενεργοποίησης [2]

φίλτρο, οι διαστάσεις να είναι 5*5*3, αν πρόκειται για έγχρωμη εικόνα. Η πρώτη θέση του φίλτρου είναι η επάνω αριστερή μεριά σε μία εικόνα. Καθώς το φίλτρο διαχέεται κατά μήκος της εικόνας, πολλαπλασιάζει τις τιμές του φίλτρου με τις τιμές των πρωτότυπων pixels (element wise multiplications). Όλοι αυτοί οι υπολογισμοί συνοψίζονται στο τέλος, έτσι ώστε να δημιουργηθεί ένας μονός αριθμός, ο οποίος αντιπροσωπεύει μόνο την αρχική θέση του φίλτρου, δηλαδή επάνω αριστερά. Η διαδικασία αυτή επαναλαμβάνεται για κάθε τοποθεσία που επισκέπτεται το φίλτρο, το οποίο μετακινείται προς τα δεξιά κατά μία μονάδα, μετά πάλι δεξιά κατά μια μονάδα, μέχρι τέλους. Αφού το φίλτρο καλύψει όλες τις περιοχές, το φίλτρο θα βρίσκεται πάλι αριστερά και θα έχει δημιουργηθεί ένας χάρτης χαρακτηριστικών ή χάρτης ενεργοποίησης (activation map ή feature map) [2].

Με άλλα λόγια, κάθε ένα από αυτά τα φίλτρα χρησιμεύουν ως ανιχνευτές χαρακτηριστικών. Ως χαρακτηριστικά, εννοούνται ευθείες ακμές, απλά χρώματα και καμπύλες. Για παράδειγμα, ένα φίλτρο το οποίο έχει τοποθετηθεί στην επάνω αριστερή γωνία στην εικόνα, αναγνωρίζει μία καμπύλη στα pixels της εικόνας που έχουν μεγαλύτερες αριθμητικές τιμές, όπως φαίνεται στο σχήμα 2.3 παρακάτω [2].

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0



Σχήμα 2.3: Αναπαράσταση των εικονοστοιχείων του φίλτρου και απεικόνιση της καμπύλης που ανιχνεύει το φίλτρο [2]

Πίσω στο μαθηματικό κομμάτι, όπως αναφέρθηκε και παραπάνω, πολλαπλασιάζονται οι τιμές του φίλτρου με τις πρωτότυπες τιμές των εικονοστοιχείων της εικόνας, όπως φαίνεται στην παρακάτω συνάρτηση: [2]

$$(50 \times 30) + (50 \times 30) + (50 \times 30) + (20 \times 30) + (50 \times 30) = 6600 \tag{2.2}$$

Μία τέτοια αναπαράσταση απεικονίζεται στο σχήμα 2.4.

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

$*$

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Σχήμα 2.4: Αναπαράσταση των εικονοστοιχείων του φίλτρου και πολλαπλασιασμός μεταξύ των αριθμητικών τιμών [2]

Γενικά, εάν υπάρχει ένα σχήμα που μοιάζει με την καμπύλη που αντιπροσωπεύει αυτό το φίλτρο, τότε όλοι οι πολλαπλασιασμοί που αθροίζονται μαζί θα έχουν ως αποτέλεσμα μια μεγάλη τιμή. Όταν το φίλτρο μετακινείται , για παράδειγμα όταν φτάσει στην επάνω δεξιά μεριά της εικόνας, τότε παρατηρείται μηδενική τιμή. Το φίλτρο και η μετακίνησή του απεικονίζονται στο σχήμα 2.5 [2].

0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

$*$

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

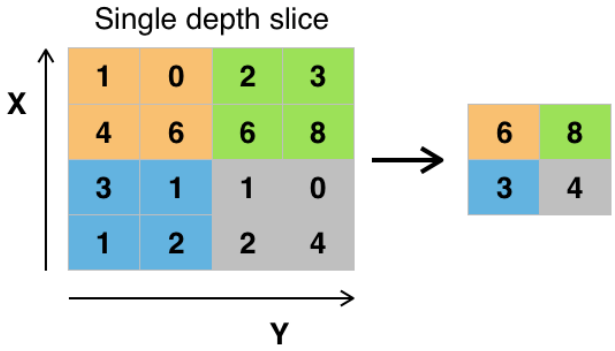
Σχήμα 2.5: Απεικόνιση του Φίλτρου στην εικόνα και αναπαράσταση των εικονοστοιχείων [2]

2.5 Συνάρτηση Ενεργοποίησης

Αμέσως μετά το συνελικτικό επίπεδο, εφαρμόζεται ένα μη γραμμικό επίπεδο, ή αλλιώς επίπεδο ενεργοποίησης. Στόχος του συγκεκριμένου επιπέδου είναι να εισάγει τη μη γραμμικότητα σε ένα σύστημα το οποίο έχει υπολογίσει γραμμικές πράξεις στο συνελικτικό επίπεδο(δηλαδή πολλαπλασιασμούς και αθροίσματα). Για το λόγο αυτό παλιότερα χρησιμοποιούνταν μη γραμμικές συναρτήσεις, όπως η εφαπτομένη (tanx) και η σιγμοειδής (sigm), ωστόσο ανακαλύφθηκε τα τελευταία χρόνια ότι η συνάρτηση Relu δουλεύει καλύτερα, γιατί είναι ικανή να εκπαιδεύσει γρηγορότερα το μοντέλο, χωρίς αξιοσημείωτη διαφορά στην ακρίβεια. Η συνάρτηση Relu εφαρμόζει τη συνάρτηση f(x)=max(0,x) σε όλες τις τιμές στην εικόνα εισόδου και ουσιαστικά μετατρέπει όλες τις αρνητικές τιμές σε 0. Αυτό το επίπεδο αυξάνει τις μη γραμμικές ιδιότητες του μοντέλου και του συνολικού δικτύου χωρίς να επηρεάζει τα φίλτρα του συνελικτικού επιπέδου [3].

2.6 Μέγιστη Συγκέντρωση (Maxpooling)

Μετά τα στρώματα που περιλαμβάνουν τη συνάρτηση ενεργοποίησης, πολλοί επιστήμονες στον κλάδο του προγραμματισμού επιλέγουν να εφαρμόσουν ένα επίπεδο συγκέντρωσης (pooling), γνωστό επίσης ως downsampling επίπεδο. Επιλέγεται ένα φίλτρο (συνήθως μεγέθους 2*2) και ένα βήμα του ίδιου μήκους. Στη συνέχεια, εφαρμόζεται στην είσοδο και εξάγει τον μέγιστο αριθμό σε κάθε υποπεριοχή γύρω από την οποία περιστρέφεται το φίλτρο. Άλλες επιλογές για poolinglayers είναι τα average pooling και L2-pooling. Γενικά, ο συλλογισμός γύρω από αυτό το επίπεδο είναι ότι όταν γνωρίζουμε ότι ένα χαρακτηριστικό βρίσκεται στην αρχή της εισόδου (όπου υπάρχει υψηλή τιμή ενεργοποίησης) , δεν ενδιαφέρει τόσο η ακριβής θέση του, όσο η σχετική του θέση με τα υπόλοιπα χαρακτηριστικά. Ένα παράδειγμα απεικόνισης μέγιστης συγκέντρωσης ενός 2*2 φίλτρου [3] απεικονίζεται στο σχήμα 2.6.



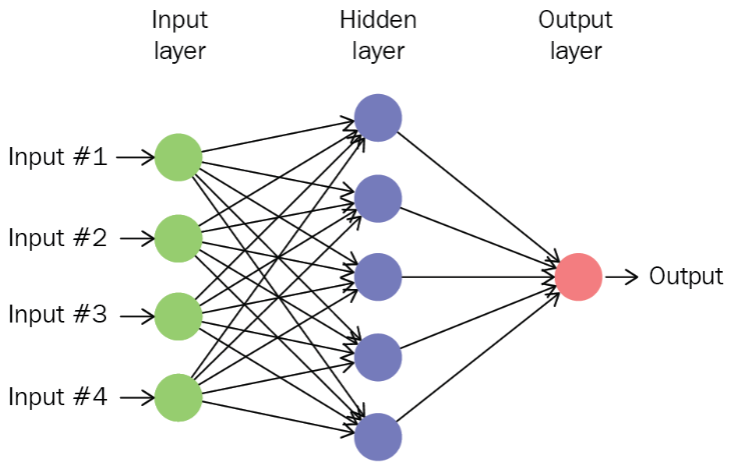
Σχήμα 2.6: Παράδειγμα maxpool 2*2 φίλτρου με δρασκελισμό 2 [3]

2.7 Επίπεδο Απομάκρυνσης (Dropout layer)

Αυτά τα επίπεδα λειτουργούν πολύ συγκεκριμένα στα νευρωνικά δίκτυα. Πολλές φορές εμφανίζεται το πρόβλημα της υπερεκπαίδευσης (overfitting) , όταν τα βάρη του δικτύου είναι πολύ συντονισμένα στα παραδείγματα εκπαίδευσης που δίνονται, έτσι ώστε να μην μπορεί το μοντέλο να εκπαιδευτεί καλά στα νέα παραδείγματα. Το dropout επίπεδο βοηθάει στη μείωση της υπερεκπαίδευσης. Ουσιαστικά, απομακρύνει ένα σύνολο συναρτήσεων από το νευρωνικό δίκτυο, θέτοντάς το μηδέν. Αξίζει να σημειωθεί ότι τα συσκεκιμένα επίπεδα χρησιμοποιούνται μόνο κατά τη διάρκεια της εκπαίδευσης, και όχι κατά τη διάρκεια του ελέγχου (test set) [3].

2.8 Ισχυρά συνδεδεμένα επίπεδα (Fully connected layers)

Το τελευταίο στρώμα στο νευρωνικό δίκτυο ονομάζεται ισχυρά συνδεδεμένο (fully connected layer). Το συγκεκριμένο επίπεδο δέχεται σαν είσοδο μία εικόνα και επιστρέφει στην έξοδο ένα N-διάστατο διάνυσμα, όπου N είναι ο αριθμός των κλάσεων, ανάμεσα στις οποίες πρέπει να επιλέξει το πρόγραμμα. Εάν για παράδειγμα πρόκειται για μία ψηφιακή ταξινόμηση, το διάνυσμα θα είχε 10 διαστάσεις, εφόσον υπάρχουν 10 ψηφία. Κάθε αριθμός στο διάνυσμα αντιπροσωπεύει την πιθανότητα να ανήκει σε συγκεκριμένη κλάση. Για παράδειγμα, εάν το διάνυσμα εξόδου ήταν : [0 .1 .1 .75 0 0 0 0 0 .05], τότε υπάρχει 10% πιθανότητα να ανήκει στην κατηγορία 1, 10% να ανήκει την κατηγορία 2, 75% να ανήκει στην κατηγορία 3 και 5% πιθα-



Σχήμα 2.7: Παράδειγμα ενός ισχυρά συνδεδεμένου επιπέδου [4]

νότητα να ανήκει στην κατηγορία 9. Το τελευταίο στρώμα του δικτύου κοιτάζει τα χαρακτηριστικά που έχουν αναγνωριστεί στο προηγούμενο επίπεδο και αποφασίζει ποια χαρακτηριστικά ταιριάζουν περισσότερο σε κάθε κατηγορία. Εάν για παράδειγμα, σε μία εικόνα απεικονίζεται ένας σκύλος, θα πρέπει να υπάρχουν υψηλές τιμές στις θέσεις που απεικονίζονται υψηλού επιπέδου χαρακτηριστικά, όπως τέσσερα πόδια, ουρά και άλλα χαρακτηριστικά. Επίσης, εάν η εικόνα απεικονίζει ένα πουλί, θα υπάρχουν υψηλές τιμές στις θέσεις που απεικονίζονται χαρακτηριστικά, όπως φτερά, ράμφος και άλλα παρόμοια χαρακτηριστικά. Συνοπτικά, ένα επίπεδο, όπως είναι το fully connected layer, συσχετίζει τα χαρακτηριστικά που έχουν ανιχνευτεί στο προηγούμενο στρώμα με συγκεκριμένη κατηγορία και με βάση συγκεκριμένα βάρη, υπολογίζονται οι σωστές πιθανότητες για διαφορετικές κατηγορίες. Ένα παράδειγμα ισχυρά συνδεδεμένου επιπέδου απεικονίζεται στο σχήμα 2.7 [2].

2.9 Μέθοδος Οπισθοδιάδοσης Σφάλματος (Backpropagation)

Γεννιούνται πολλά ερωτήματα σχετικά με τη διαδικασία που ακολουθείται στα νευρωνικά δίκτυα, δηλαδή πώς ξέρουν τα φίλτρα να αναγνωρίσουν ακμές ή καμπύλες, ή πώς ξέρουν τα φίλτρα τι τιμές πρέπει να πάρουν και άλλα παρόμοια ερωτήματα. Ο τρόπος με τον οποίο ο υπολογιστής καταφέρνει να προσαρμόσει τις τιμές των φίλτρων (βάρη), γίνεται μέσω μιας διαδικασίας εκπαίδευσης που ονομάζεται Μέθοδος Οπισθοδιάδοσης Σφάλματος (Backpropagation). Η αλήθεια είναι ο υπολογιστής δε μπορεί να ξεχωρίσει τα χαρακτηριστικά μιας εικόνας από μόνος του,

ούτε τα φίλτρα μπορούν να αναγνωρίσουν ακμές ή καμπύλες. Η βασική ιδέα είναι να δώσουμε στον υπολογιστή μία εικόνα και μία ετικέτα. Υπάρχει μία διαδικασία εκπαίδευσης, όπου ο υπολογιστής μαθαίνει μέσα από χιλιάδες εικόνες και τις αντίστοιχες ετικέτες, οι οποίες περνάνε μέσα από το νευρωνικό δίκτυο. Η μέθοδος Οπισθοδιάδοσης σφάλματος αποτελείται από 4 μέρη, forwardpass (πέρασμα προς τα εμπρός), loss function (συνάρτηση απώλειας), backwardpass (πέρασμα προς τα πίσω) και weightupdate (ενημέρωση βαρών). Κατά τη διάρκεια του forwardpass, μία εικόνα εκπαίδευσης περνάει μέσα από το νευρωνικό δίκτυο. Στο πρώτο παράδειγμα εκπαίδευσης, οι τιμές των βαρών ή των φίλτρων αρχικοποιούνται τυχαία. Το αποτέλεσμα πιθανώς να είναι [.1 .1 .1 .1 .1 .1 .1 .1 .1], χωρίς να δίνει κάποια προτίμηση σε κανέναν αριθμό συγκεκριμένα. Με αυτά τα βάρη, το δίκτυο δεν είναι ικανό να αναγνωρίσει χαρακτηριστικά ή να ταξινομήσει τις εικόνες σε κατηγορίες. Σε αυτό χρησιμεύει η συνάρτηση απώλειας του Backpropagation. Η πιο κοινή συνάρτηση απώλειας είναι η MSE (μέσο τετραγωνικό σφάλμα) , η οποία ορίζεται από τον εξής τύπο : [2]

$$E_{total} = \sum \frac{1}{2} (target - output)^2 \tag{2.3}$$

όπου target είναι η πραγματική τιμή (ground truth) και output είναι η εκτιμώμενη τιμή που παρήγαγε το μοντέλο (η έξοδος του νευρώνα ή του δικτύου).

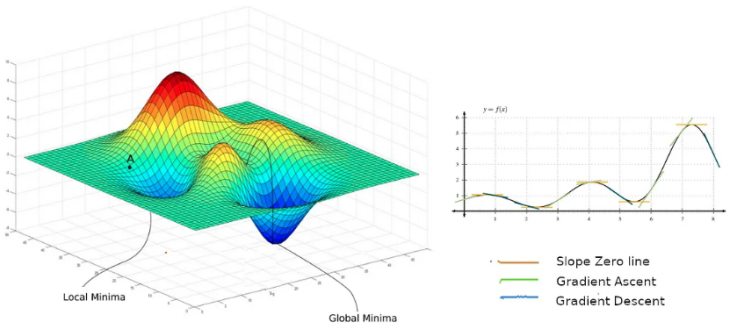
2.10 Βελτιστοποιητές

Κατά την εκπαίδευση ενός μοντέλου, στο πρώτο ζευγάρι εικόνων που εκπαιδεύεται θα υπάρχει υψηλό σφάλμα, το οποίο μειώνεται κατά τη διάρκεια της εκπαίδευσης. Στόχος είναι να βρεθεί το σημείο στο οποίο η ετικέτα πρόβλεψης, δηλαδή το αποτέλεσμα που θα προβλέψει το νευρωνικό δίκτυο, να είναι ίδια με την ετικέτα εκπαίδευσης , πράγμα το οποίο σημαίνει ότι έχει γίνει σωστή πρόβλεψη. Προκειμένου να συμβεί αυτό, πρέπει να μειωθεί αρκετά η τιμή του σφάλματος. Το συγκεκριμένο πρόβλημα αποτελεί ένα πρόβλημα βελτιστοποίησης, το οποίο μπορεί να λυθεί παρατηρώντας τα βάρη που προκαλούν μεγαλύτερη μείωση σφάλματος. Ένας από τους πιο δημοφιλείς βελτιστοποιητές είναι η κλίση κατάβασης (gradient descent), η οποία λειτουργεί προσαρμόζοντας επαναληπτικά τις παραμέτρους του μοντέλου προς την κατεύθυνση που ελαχιστοποιεί τη λειτουργία απώλειας. Μία παραλλαγή αποτελεί

η στοχαστική κλίση κατάβασης (stochastic gradient descent), η οποία ενημερώνει τις παραμέτρους του μοντέλου μετά από κάθε παράδειγμα εκπαίδευσης, καθιστώντας το πιο αποτελεσματικό για μεγάλα σύνολα δεδομένων σε σύγκριση με την παραδοσιακή κλίση κατάβασης (gradient descent), η οποία χρησιμοποιεί ολόκληρο το σύνολο δεδομένων για κάθε ενημέρωση. Τέλος, ο αλγόριθμος προσαρμοστικής κλίσης (Adaptive Gradient Algorithm, AdaGrad) και η μέση τετραγωνική διάδοση ρίζας (Root Mean Square Propagation, RMSProp) αποτελούν προηγμένους βελτιστοποιητές. Ο AdaGrad προσαρμόζει τον ρυθμό εκμάθησης για κάθε παράμετρο με βάση τις πληροφορίες κλίσης, ενώ ο RMSProp (Root Mean Square Propagation) βελτιώνει τον AdaGrad εισάγοντας έναν παράγοντα αποσύνθεσης για να αποτρέψει την πολύ γρήγορη μείωση του ποσοστού μάθησης [13].

2.10.1 Κλίση κατάβασης (Gradient Descent)

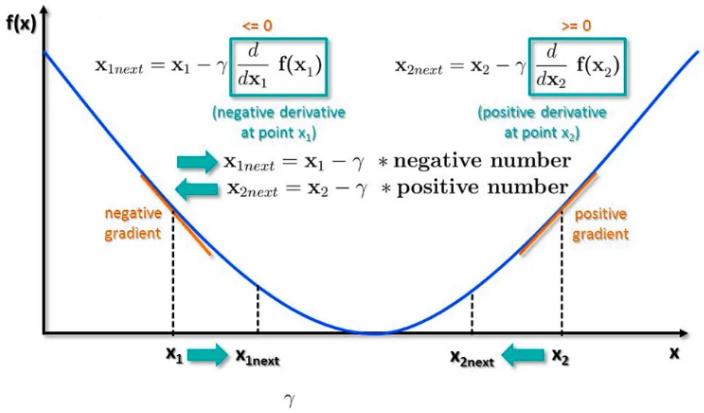
Η κλίση κατάβασης (gradient descent) είναι μία επαναληπτική διαδικασία ,κατά την οποία βελτιστοποιούνται οι παράμετροι των συνελικτικών νευρωνικών δικτύων, επομένως βελτιστοποιείται η συνάρτηση απώλειας. Η κλίση σε μία συνεχή συνάρτηση ορίζεται ως το διάνυσμα που περιέχει τις μερικές παραγώγους που έχουν υπολογιστεί σε ένα συγκεκριμένο σημείο. Η κλίση είναι πεπερασμένη και ορίζεται εάν και μόνο εάν όλες οι μερικές παράγωγοι είναι επίσης καθορισμένες και πεπερασμένες. Με άλλα λόγια, η κλίση κατάβασης βοηθάει κατά τη διαδικασία βελτιστοποίησης, στον καθορισμό της μετακίνησης των διαφόρων παραμέτρων, προκειμένου να ελαχιστοποιηθεί το σφάλμα. Στο παρακάτω σχήμα 2.8 απεικονίζονται οι κλίσεις της συνάρτησης απώλειας σε 3D και 2D [5].



Σχήμα 2.8: Απεικόνιση της κλίσης καθόδου σε 3D και 2D [5]

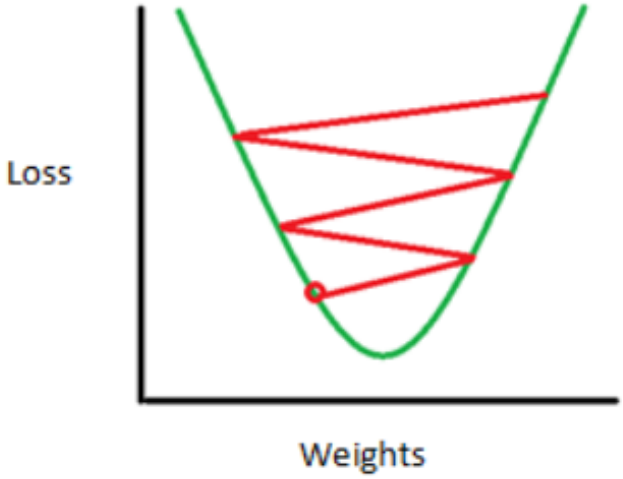
Η παράγωγος μιας συνάρτησης ορίζεται ως ο στιγμιαίος ρυθμός μεταβολής μιας

συνάρτησης σε ένα συγκεκριμένο σημείο. Η παράγωγος δίνει την ακριβή κλίση κατά μήκος της καμπύλης σε ένα συγκεκριμένο σημείο. Ένα παράδειγμα απεικόνισης της κατεύθυνσης της κλίσης απεικονίζεται στο σχήμα 2.9 [5].



Σχήμα 2.9: Απεικόνιση της αρνητικής και θετικής κλίσης καθόδου [5]

Ο ρυθμός μάθησης (learning rate), ο οποίος συμβολίζεται ως γ στην παραπάνω εικόνα, καθορίζεται από τον προγραμματιστή. Ένας υψηλός ρυθμός μάθησης σημαίνει μεγάλα βήματα, προκειμένου να αναβαθμιστούν τα βάρη, επομένως απαιτείται λιγότερος χρόνος για να φτάσουμε στα βέλτιστα βάρη. Ωστόσο, ένα υπερβολικά υψηλό ποσοστό εκμάθησης οδηγεί σε τεράστια βήματα και δεν μπορεί να φτάσει με ακρίβεια το βέλτιστο σημείο. Στο επόμενο σχήμα 2.10 παρατηρούνται τεράστια βήματα, με αποτέλεσμα να μη μπορεί να ελαχιστοποιηθεί το σφάλμα [2].



Σχήμα 2.10: Αποτέλεσμα ενός υψηλού ποσοστού εκμάθησης όπου τα βήματα είναι υπερβολικά μεγάλα, χωρίς ελαχιστοποίηση του σφάλματος [2]

Η διαδικασία του περάσματος προς τα εμπρός, της συνάρτησης απώλειας, του περάσματος προς τα πίσω και της ενημέρωσης παραμέτρων είναι μία επανάληψη εκπαίδευσης. Το πρόγραμμα θα επαναλάβει αυτή τη διαδικασία για έναν σταθερό αριθμό επαναλήψεων για κάθε σύνολο εικόνων εκπαίδευσης (κοινώς ονομάζεται παρτίδα). Μόλις ολοκληρωθεί η ενημέρωση παραμέτρων στο τελευταίο παράδειγμα εκπαίδευσης, ελπίζουμε ότι το δίκτυο θα πρέπει να έχει εκπαιδευτεί αρκετά καλά, ώστε τα βάρη των επιπέδων να συντονιστούν σωστά [2].

2.10.2 Κλίση Καθόδου Κατά Παρτίδα (Batch Gradient Descent)

Μία παραλλαγή της κλίσης καθόδου (gradient descent), η οποία είναι ευρέως χρησιμοποιούμενη σε γραμμικά μοντέλα και σε νευρωνικά δίκτυα είναι η κλίση καθόοδυ κατά παρτίδα (Batch Gradient Descent). Αποτελεί έναν αλγόριθμο βελτιστοποίησης, προκειμένου να ελαχιστοποιηθεί το σφάλμα. Η Batch Gradient Descent κάνει υπολογισμούς στο σύνολο των δεδομένων εκπαίδευσης σε κάθε βήμα, με αποτέλεσμα να είναι πολύ αργός αλγόριθμος σε πολύ μεγάλα δεδομένα εκπαίδευσης. Αποτελεί έναν ντετερμινιστικό αλγόριθμο, ο οποίος δίνει τη βέλτιστη λύση με επαρκή χρόνο για σύγκλιση, δεν απαιτείται τυχαίο ανακάτεμα σημείων, δε μπορεί να ξεφύγει εύκολα από τα ρηχά τοπικά ελάχιστα και η σύγκλιση είναι αργή [14].

2.10.3 Στοχαστική Κλίση Καθόδου (Stochastic Gradient Descent)

Η Στοχαστική κλίση καθόδου (Stochastic Gradient Descent) αποτελεί επίσης έναν αλγόριθμο βελτιστοποίησης , ο οποίος εφαρμόζεται προκειμένου να ελαχιστοποιηθεί το σφάλμα. Χρησιμοποιείται προκειμένου να λύσει το κύριο πρόβλημα της Batch Gradient Descent, δηλαδή τη χρήση ολόκληρου του συνόλου δεδομένων εκπαίδευσης, προκειμένου να υπολογιστεί η κλίση σε κάθε βήμα. Η Stochastic Gradient Descent είναι στοχαστική, καθώς μαζεύει κάθε φορά ένα στιγμιότυπο των δεδομένων εκπαίδευσης σε κάθε βήμα και μετά υπολογίζει την κλίση με πιο γρήγορο ρυθμό, αφού υπάρχουν πολύ λιγότερα δεδομένα για να διαχειριστεί κάθε φορά, σε αντίθεση με την Batch Gradient Descent. Το μειονέκτημα είναι ότι όταν φτάσει κοντά στην ελάχιστη τιμή δεν σταματά, αλλά αντιθέτως αναπηδά, δίνοντας καλές τιμές για τις παραμέτρους, αλλά όχι βέλτιστες. Το αναπήδημα μπορεί να βελτιωθεί, μειώνοντας το ρυθμό εκπαίδευσης σε κάθε βήμα και με αυτό τον τρόπο, η stochastic Gradient Descent σταματά στο ελάχιστο μετά από κάποιο χρονικό διάστημα. Τέλος, μπορεί να αποφυγεί η υπερεκπαίδευση, ενημερώνοντας τις παραμέτρους του μοντέλου πιο συχνά [15] [14].

2.10.4 Βελτιστοποιητής Adam

Ο Adam είναι ένας αλγόριθμος βελτιστοποίησης , ο οποίος χρησιμοποιείται αντί για την κλασική στοχαστική κλίση καθόδου (stochastic Gradient Descent), για να αναβαθμίσει τα βάρη του δικτύου επαναληπτικά με βάση τα δεδομένα εκπαίδευσης. Η μέθοδος αυτή είναι αποτελεσματική όταν υπάρχει μεγάλος αριθμός δεδομένων ή πολλές παράμετροι, καθώς απαιτείται λιγότερη μνήμη σε σχέση με άλλες μεθόδους. Αποτελεί συνδυασμό του αλγορίθμου στοχαστική κλίση κατάβασης με ορμή (stochastic gradient descent with momentum,SGD with Momentum) [16] και της μέσης τετραγωνικής διάδοσης ρίζας (Root Mean Square Propagation, RMSP) [17]. Συγκεκριμένα, ο πρώτος αλγόριθμος χρησιμοποιείται για την επιτάχυνση του αλγόριθμου gradient descent λαμβάνοντας υπόψη τον «εκθετικά σταθμισμένο μέσο όρο» των κλίσεων. Η χρήση μέσων όρων κάνει τον αλγόριθμο να συγκλίνει προς τα ελάχιστα με ταχύτερο ρυθμό [14] [18].

$$w_{t+1} = w_t - \alpha m_t \tag{2.4}$$

όπου :

$$m_t = \beta m_{t+1} + (1 - \beta) \left[\frac{\partial L}{\partial w_t} \right] \tag{2.5}$$

$m_t =$ σύνολο των παραγώγων σε μία χρονική στιγμή t

$m_{t-1} =$ σύνολο των παραγώγων την προηγούμενη χρονική στιγμή $t-1$

$w_t =$ βάρη τη χρονική στιγμή t

$w_{t+1} =$ βάρη τη χρονική στιγμή $t+1$

$\alpha_t =$ ποσοστό εκμάθησης τη χρονική στιγμή t

$\partial L =$ παράγωγος της συνάρτησης απώλειας

$\partial w_t =$ παράγωγος του βάρους τη χρονική στιγμή t

$\beta =$ παράμετρος μέσου όρου[6]

Η Διάδοση της Τετραγωνικής Μέσης Τιμής (Root mean square prop ή RMS prop) είναι ένας προσαρμοστικός αλγόριθμος εκμάθησης που προσπαθεί να βελτιώσει τον αλγόριθμο προσαρμοστικής κλίσης (Adaptive Gradient Algorithm, AdaGrad) [19]. Αντί να παίρνει το συνολικό άθροισμα των τετραγωνικών κλίσεων όπως στο Adagrad, παίρνει τον "εκθετικό κινούμενο μέσο όρο". Ορισμένα από τα πλεονεκτήματα του συγκεκριμένου αλγορίθμου είναι η αποτελεσματικότητα σε μη κυρτά προβλήματα, το σταθερό ποσοστό εκμάθησης και η εμπειρικά αποδεδειγμένη απόδοση [20].

$$w_{t+1} = w_t - \frac{\alpha_t}{(v_t + \epsilon)^{1/2}} * \left[\frac{\partial L}{\partial w_t} \right] \tag{2.6}$$

$$v_t = \beta v_{t-1} + (1 - \beta) * \left[\frac{\partial L}{\partial w_t} \right]^2 \tag{2.7}$$

w_t = βάρη τη χρονική στιγμή t

w_{t+1} = βάρη τη χρονική στιγμή $t+1$

α_t = ποσοστό εκμάθησης τη χρονική στιγμή t

∂L = παράγωγος της συνάρτησης απώλειας

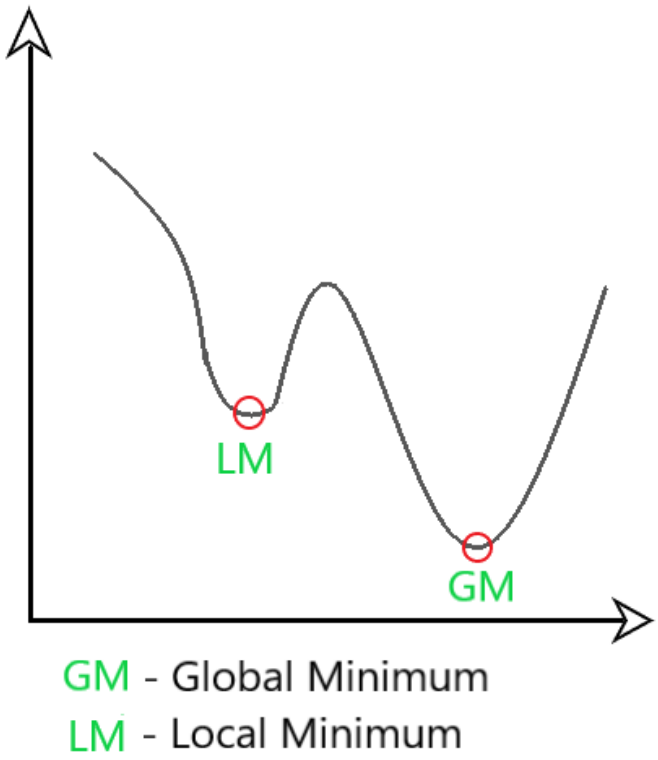
∂W_t = παράγωγος των βαρών τη χρονική στιγμή t

V_t = άθροισμα των τετραγώνων των περασμένων κλίσεων

β = παράμετρος κινούμενου μέσου όρου

ϵ = μία μικρή θετική σταθερά(10^{-8})[6]

Ο βελτιστοποιητής Adam κληρονομεί τα πλεονεκτήματα των παραπάνω δύο μεθόδων και βασίζεται σε αυτά για να δώσει μια πιο βελτιστοποιημένη κλίση κατάβασης. Στο σχήμα 2.11 απεικονίζονται οι θέσεις τοπικού και ολικού ελαχίστου.



Σχήμα 2.11: Απεικόνιση ολικού και τοπικού ελαχίστου [6]

Ο ρυθμός κλίσης καθόδου ελέγχεται με τέτοιο τρόπο, έτσι ώστε να υπάρχει ελάχιστη ταλάντωση όταν φτάσει στο ολικό ελάχιστο, ενώ γίνονται μεγάλα βήματα,

προκειμένου να ξεπεραστούν τα τοπικά ελάχιστα.

2.11 Συναρτήσεις απώλειας

Οι συναρτήσεις απώλειας αποτελούν ένα από τα πιο σημαντικά κομμάτια στα νευρωνικά δίκτυα, καθώς είναι υπέυθυνες για το ταίριασμα του μοντέλου με τα δοθέντα δεδομένα εκπαίδευσης. Η συνάρτηση απώλειας έιναι μία συνάρτηση η οποία συγκρίνει την πραγματική τιμή με την τιμή που προκύπτει από την πρόβλεψη. Ουσιαστικά, μετράει πόσο καλά εκπαιδεύεται το μοντέλο στα δεδομένα εκπαίδευσης. Στόχος είναι να ελαχιστοποιηθεί το σφάλμα μεταξύ πραγματικής και προβλεπόμενης τιμής [21]. Οι υπερπαράμετροι προσαρμόζονται με σκοπό την ελαχιστοποίηση του μέσου σφάλματος, βρίσκοντας τα κατάλληλα βάρη (w) και σταθερούς όρους (biases, b). Σκοπός είναι να ελαχιστοποιηθεί η συνάρτηση:

$$J(w^T, b) = \frac{1}{m} \sum_{i=1}^{m} L(\hat{y}^{(i)}, y^{(i)}) \tag{2.8}$$

όπου m ο αριθμός των παραδειγμάτων στο σύνολο εκπαίδευσης, $y^{(i)}$ είναι η πραγματική τιμή ετικέτας (ground truth) και $\hat{y}^{(i)}$είναι η πρόβλεψη τουν μοντέλου και υπολογίζεται από τον τύπο :

$$\hat{y}^{(i)} = \sigma\left(w^T x^{(i)} + b\right). \tag{2.9}$$

όπου w διάνυσμα βαρών, x^i το διάνυσμα που δέχεται σαν είσοδο το δίκτυο με δείκτη i και σ η σειγμοειδής συνάρτηση.

Στην κατ'επίβλεψη μάθηση, υπάρχουν δύο κύρια είδη συναρτήσεων απώλειας: παλινδρόμησης και ταξινόμησης. Οι συναρτήσεις απώλειας παλινδρόμησης χρησιμοποιούνται σε νευρωνικά δίκτυα παλινδρόμησης. Δοθέντος μίας τιμής εισόδου, το μοντέλο προβλέπει την αντίστοιχη τιμή αποτελέσματος. Παραδείγματα τέτοιων συναρτήσεων είναι το μέσο τετραγωνικό σφάλμα και το μέσο απόλυτο σφάλμα. Οι συναρτήσεις απώλειας ταξινόμησης χρησιμοποιούνται σε νευρωνικά δίκτυα ταξινόμησης. Δοθέντος μίας τιμής εισόδου, το νευρωνικό δίκτυο παράγει ένα διάνυσμα πιθανοτήτων με το αντικείμενο εισόδου να ανήκει σε πολλές κατηγορίες, επιλέγοντας την κατηγορία με την υψηλότερη πιθανότητα να ανήκει στη συγκεκριμένη

κατηγορία. Παραδείγματα τέτοιων συναρτήσεων είναι Binary cross entropy και categorical cross entropy. Μία από τις πιο δημοφιλείς συναρτήσεις απώλειας είναι το σφάλμα ελαχίστου τετραγώνου (MSE) , το οποίο υπολογίζει τη μέση τετραγωνική διαφορά μεταξύ της πραγματικής και της προβλεπόμενης τιμής. Ορίζεται από τον τύπο :

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y^{(i)} - \hat{y}^{(i)}\right)^2 \tag{2.10}$$

όπου n ο συνολικός αριθμός των δειγμάτων (ή παρατηρήσεων) στο σετ εκπαίδευσης, i ο δείκτης που παίρνει τιμές από 1 μέχρι n, $y^{(i)}$ η πραγματική (ground-truth) τιμή για το δείγμα i και $\hat{y}^{(i)}$ η προβλεπόμενη τιμή που δίνει το μοντέλο για το δείγμα i.

Η διαφορά είναι τετραγωνική, δηλαδή δεν έχει σημασία εάν η προβλεπόμενη τιμή είναι πάνω ή κάτω από την τιμή-στόχο. Ωστόσο, τιμές με μεγάλο σφάλμα τιμωρούνται. Το μέσο τετραγωνικό σφάλμα επίσης είναι μία κυρτή συνάρτηση με προκαθορισμένο ολικό ελάχιστο, με αποτέλεσμα να είναι πιο εύκολη η βελτιστοποίηση της κατάβασης κλίσης προκειμένου να οριστούν οι τιμές των βαρών. Το μειονέκτημα αυτής της συνάρτησης είναι ότι είναι πολύ ευαίσθητη σε ακραίες τιμές. Με λίγα λόγια, εάν η τιμή της πρόβλεψης είναι σημαντικά μεγαλύτερη ή μικρότερη από την τιμή στόχο της, θα αυξήσει πολύ την απώλεια. Άλλη μία σημαντική συνάρτηση απώλειας είναι το μέσο απόλυτο σφάλμα (MAE), το οποίο υπολογίζει τη μέση απόλυτη διαφορά ανάμεσα στην τιμή-στόχο και στην τιμή πρόβλεψης και δίνεται από τον τύπο:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y^{(i)} - \hat{y}^{(i)}\right| \tag{2.11}$$

Η συγκεκριμένη συνάρτηση χρησιμοποιείται ως εναλλακτική για το μέσο τετραγωνικό σφάλμα στις περισσότερες περιπτώσεις. Το απόλυτο μέσο σφάλμα είναι εξαιρετικά ευαίσθητο σε ακραίες τιμές, επηρεάζοντας δραματικά το σφάλμα επειδή η απόσταση είναι τετραγωνική. Η συνάρτηση MAE χρησιμοποιείται σε περιπτώσεις όπου τα δεδομένα εκπαίδευσης έχουν ένα μεγάλο αριθμό από ακραίες τιμές για να μετριάσουν. Ωστόσο, η συνάρτηση απώλειας MAE έχει ορισμένα μειονεκτήματα. Καθώς η μέση απόσταση προσεγγίζει το 0, η βελτιστοποίηση της κλίσης κατάβασης δε θα λειτουργήσει, καθώς η παράγωγος της συνάρτησης στο 0 δεν ορίζεται(θα οδηγήσει σε σφάλμα, καθώς δε μπορεί να γίνει διαίρεση με το 0). Εξαιτίας αυτού

του γεγονότος, γεννιέται μία καινούρια συνάρτηση, η οποία αποτελεί συνδυασμό των συναρτήσεων μέσου απόλυτου σφάλματος (MAE) και σφάλματος ελαχίστου τετραγώνου (MSE) και ονομάζεται Huber Loss [22], όπου ο τύπος είναι:

$$\frac{1}{n}\sum_{i=1}^{n}\begin{cases}(y^{(i)}-\hat{y}^{(i)})^2, & \text{αν} |y^{(i)}-\hat{y}^{(i)}| \leq \delta \\ \delta(|y^{(i)}-\hat{y}^{(i)}|-\frac{1}{2}\delta), & \text{αν } |y^{(i)}-\hat{y}^{(i)}| > \delta\end{cases} \tag{2.12}$$

Με λίγα λόγια, η συνάρτηση τετραγωνικού σφάλματος εφαρμόζεται όταν η διαφορά μεταξύ πραγματικής και τιμής πρόβλεψης είναι μικρότερη ή ίση με την τιμή δ του κατωφλίου. Αλλιώς, εφαρμόζεται η συνάρτηση απόλυτου σφάλματος.
Μία άλλη συνάρτηση απώλειας που χρησιμοποιείται σε δυαδικά μοντέλα ταξινόμησης είναι η Binary-cross Entropy (Δυαδική διασταυρούμενη εντροπία), όπου το μοντέλο δέχεται σαν είσοδο μία εικόνα και πρέπει στην έξοδο να την ταξινομήσει σε μία από τις δύο κατηγορίες. Η συνάρτηση αυτή δίνεται από τον τύπο:

$$\frac{1}{n}\sum_{i=1}^{N} -\left(y_i \cdot \log(p_i) + (1-y_i)\cdot \log(1-p_i)\right) \tag{2.13}$$

όπου n ο συνολικός αριθμός δειγμάτων (παρατηρήσεων) στο σετ εκπαίδευσης, i ο δείκτης που παίρνει τιμές από 1 έως n, y_i η πραγματική (ground-truth) ετικέτα για το δείγμα i (0 ή 1) και τέλος, p_i η προβλεπόμενη πιθανότητα ότι το δείγμα i ανήκει στην κλάση 1.
Τα νευρωνικά δίκτυα ταξινόμησης δίνουν σαν αποτέλεσμα ένα διάνυσμα πιθανοτήτων, με την πιθανότητα η εικόνα εισόδου να ανήκει σε μία από τις υπάρχουσες κατηγορίες, επιλέγοντας την μεγαλύτερη πιθανότητα σαν τελικό αποτέλεσμα. Στη δυαδική ταξινόμηση, υπάρχουν μόνο δύο πιθανές πραγματικές τιμές , 0 ή 1.Επομένως, για να καθοριστεί ακριβώς το σφάλμα μεταξύ της πραγματικής και της τιμής πρόβλεψης, χρειάζεται να συγκριθεί η πραγματική τιμή (0 ή 1) με την πιθανότητα ότι η εικόνα εισόδου ανήκει σε αυτή την κατηγορία. Τέλος, εάν ο αριθμός των κατηγοριών είναι μεγαλύτερος από δύο, χρησιμοποιείται η κατηγορική διασταυρούμενη εντροπία, όπου ο τύπος δίνεται από:

$$-\frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\cdot\log(p_{ij}) \tag{2.14}$$

Η δυαδική διασταυρούμενη εντροπία (Binary cross entropy) αποτελεί μία ειδική περίπτωση κατηγορικής διασταυρούμενης εντροπίας (categorical cross entropy), όπου M=2, δηλαδή ο αριθμός των κατηγοριών είναι 2 [22].

2.12 Μετρικές

Οι μετρικές απόδοσης είναι αριθμητικές τιμές οι οποίες μετρούν πόσο καλά το μοντέλο πετυχαίνει τους στόχους του. Αρχικά, ως μετρικές μπορούν να χρησιμοποιηθούν όλες οι συναρτήσεις απώλειας. Ορισμένες μετρικές ακριβείας είναι οι εξής: accuracy class, binary accuracy class, categorical accuracy class, κ.ο.κ. Ορισμένες πιθανοτικές μετρικές είναι οι εξής: binary cross entropy class, categorical cross entropy class, poisson class. Υπάρχουν και μετρικές παλινδρόμησης, όπως: μέσο τετραγωνικό σφάλμα, μέσο απόλυτο σφάλμα. Επιπλέον, υπάρχουν μετρικές ταξινόμησης με βάση τα αληθινά/ψευδή θετικά και αρνητικά, όπως: AUC class, precision class, recall class, F1 score και άλλες. Τέλος, όσον αφορά την κατάτμηση εικόνας, τις πιο δημοφιλείς μετρικές αποτελούν οι εξής: IOU, binary IOU, mean IOU [23].

2.13 Εκπαίδευση

Ξεκινώντας στα νευρωνικά δίκτυα, αρχικοποιούμε τα βάρη τυχαία. Κατά τη διάρκεια της εκπαίδευσης, θέλουμε να ξεκινήσουμε με ένα νευρωνικό δίκτυο με χαμηλή απόδοση και σιγά σιγά να βελτιώνεται η απόδοση, έτσι ώστε να φτάσει σε υψηλή απόδοση. Στο τέλος της εκπαίδευσης, στόχος είναι να ελαχιστοποιηθεί κατά πολύ η συνάρτηση απώλειας. Προσαρμόζοντας τα βάρη, μπορούμε να βελτιώσουμε την απόδοση του νευρωνικού δικτύου. Το πρόβλημα της εκπαίδευσης συνεπάγεται την ελαχιστοποίηση της συνάρτησης απώλειας. Υπάρχουν πολλοί αλγόριθμοι ο οποίοι βελτιστοποιούν τη λειτουργία του νευρωνικού δικτύου [24].

Κεφαλαιο 3

Περιγραφη μοντελων

Στο κεφάλαιο 3 παρουσιάζεται η δομή των μοντέλων τα οποία χρησιμοποιήθηκαν στη συγκεκριμένη αναφορά και έχουν παρόμοια δομή. Παράλληλα, γίνεται επεξήγηση των διαφορών τους, παρουσιάζονται οι παραάμετροι και οι συναρτήσεις απώλειας ξεχωριστά για κάθε μοντέλο.

3.1 Δομή Stardist

Το μοντέλο StarDist [25] μπορεί να προβλέψει ένα πολύγωνο σε σχήμα αστεριού με R κόμβους για κάθε εικονοστοιχείο σε μία εικόνα που απεικονίζει ένα ή περισσότερα βιολογικά αντικείμενα. Συγκεκριμένα, για κάθε εικονοστοιχείο (i,j) στην εικόνα, ο StarDist προβλέπει ένα σύνολο ακτινικών αποστάσεων $D_{ij} = \{d_{ij}^k\}_{k=0,\dots,R-1}$ μέχρι το περίγραμμα του αντικειμένου, δημιουργώντας προκαθορισμένες γωνίες ίσου μήκους $\frac{2\pi}{R}$ rad. Επειδή οι ακτινικές αποστάσεις δεν ορίζονται για στοιχεία που ανήκουν στο φόντο, υπολογίζεται για κάθε pixel μία πιθανότητα p_{ij} , αγνοώντας τις μικρές πιθανότητες των αντικειμένων. Για κάθε εικονοστοιχείο στην εικόνα, το σφάλμα του StarDist βρίσκεται από τον τύπο:

$$L_{\text{StarDist}}(p_{ij}, \hat{p}_{ij}, D_{ij}, \hat{D}_{ij}) = L_{\text{BCE}}(p_{ij}, \hat{p}_{ij}) + \lambda_1 L_{\text{radii}}(p_{ij}, D_{ij}, \hat{D}_{ij}) \tag{3.1}$$

όπου P_{ij} και $\hat{P}_{ij}$ είναι η πραγματική πιθανότητα και η προβλεπόμενη πιθανότητα, και D_{ij}, $\hat{D}_{ij}$ το πραγματικό και αντίστοιχα το προβλεπόμενο σύνολο ακτινικών αποστάσεων ως το περίγραμμα του αντικειμένου. Ο πρώτος όρος L_{BCE} είναι το binary cross entropy loss, όπου εκτιμάται η τιμή της πιθανότητας μεταξύ 0 και 1, με αυτή την τιμή να αντιστοιχίζεται στην πιθανότητα να ανήκει το δείγμα στην κλάση ή κατηγορία, και λ_1 είναι ένας παράγοντας τακτοποίησης (μερικές φορές ονομάζεται υπερπαράμετρος βάρους ή αντιστάθμισης) που κλιμακώνει τη συμβολή του δεύτερου όρου L_{radii}, σε σχέση με τον πρώτο όρο L_{BCE}. Εάν αυξηθεί αυτή η υπερπαράμετρος, το δίκτυο θα προβλέψει πιο σωστά τις ακτίνες, άρα και τα περιγράμματα του αντικειμένου, ενώ αν μειωθεί, το δίκτυο θα επικεντρωθεί περισσότερο στη σωστή ταξινόμηση των εικονοστοιχείων. Η συνάρτηση απώλειας δίνει τη διαφορά μεταξύ προβλεπόμενης και πραγματικής τιμής. Η εντροπία υπολογίζει το βαθμό τυχαιότητας ή αποδιάταξης στο σύστημα. Ο παραπάνω τύπος βασίζεται στον τύπο που αφορά το binary classification, όπου η συνάρτηση απώλειας ορίζεται ως:

$$L(y, f(x)) = -\left[y \log f(x) + (1-y) \log(1-f(x))\right] \tag{3.2}$$

όπου L είναι η συνάρτηση binary cross entropy loss, y είναι το true binary label (0 ή 1), και $f(x)$ είναι η προβλεπόμενη πιθανότητα της θετικής κλάσης (με τιμές μεταξύ 0 και 1).

Ο δεύτερος όρος L_{radii} αντιπροσωπεύει το μέσο απόλυτο σφάλμα, σταθμισμένο με βάση τις ground truth πιθανότητες.

$$\mathcal{L}_{\text{radii}}(p_{ij}, D_{ij}, \hat{D}_{ij}) = p_{ij} \cdot \mathbf{1}_{p_{ij}>0} \cdot \frac{1}{R} \sum_{k=0}^{R-1} \left| d_{ij}^{k} - \hat{d}_{ij}^{k} \right| + \lambda_2 \cdot \mathbf{1}_{p_{ij}=0} \cdot \frac{1}{R} \sum_{k=0}^{R-1} \left| \hat{d}_{ij}^{k} \right|, \tag{3.3}$$

όπου p_{ij} η ground-truth πιθανότητα ότι το εικονοστοιχείο (i,j) ανήκει σε αντικείμενο (1 εάν ανήκει στο αντικείμενο, 0 για background), $\mathbf{1}_{\{p_{ij}>0\}}$ ο δείκτης ισούται με 1, μόνο αν τα εικονοστοιχεία ανήκουν στο αντικείμενο, $\mathbf{1}_{\{p_{ij}=0\}}$ ο δείκτης ισούται με 1, μόνο αν τα εικονοστοιχεία ανήκουν στο φόντο. Επίσης, D_{ij} είναι το σύνολο των πραγματικών ακτινικών αποστάσεων από το εικονοστοιχείο (i,j) προς το όριο του αντικειμένου, $\widehat{D}_{ij}$ το σύνολο των προβλεπόμενων ακτινικών αποστάσεων, R ο αριθμός των ακτινών γύρω από κάθε εικονοστοιχείο που χρησιμοποιούνται για να προσεγγίσουμε το σχήμα και λ_2 είναι ένας παράγοντας τακτοποίησης. Με μεγάλο λ_2 το δίκτυο θα θέσει στις προβλεπόμενες ακτίνες που υπάρχουν στο φόντο την τιμή 0, αλλιώς μικρή τιμή του λ_2 θέτει μικρή τιμή στις ακτίνες ακόμη και στο φόντο.

Μόλις έχουν δημιουργηθεί τα πολύγωνα για σχήματα αστεριού, τα επικαλυπτόμενα υποψήφια σημεία φιλτράρονται με μια διαδικασία που ονομάζεται Non maximum suppression (NMS), λαμβάνοντας υπόψη τις αντίστοιχες πιθανότητες. Το τελικό σύνολο πολυγώνων που δημιουργείται αντιστοιχεί σε στιγμιότυπα αντικειμένων μέσα στην εικόνα [7].

3.2 Δομή Splinedist

Ο Splinedist είναι ένα νευρωνικό δίκτυο, το οποίο αποτελεί εξέλιξη του Stardist και σχηματίζει δισδιάστατες παραμετρικές καμπύλες, οι οποίες ορίζονται ως εξής:

$$\mathbf{s}(t) = \begin{bmatrix} s_x(t) \\ s_y(t) \end{bmatrix} = \sum_{k=0}^{M-1} c[k]\varphi_M(t-k), \tag{3.4}$$

όπου c[k] είναι δισδιάστατα σημεία ελέγχου (control points), τα οποία αποτελούν τις παραμέτρους του μοντέλου. Τα control points μπορούν να τοποθετηθούν οπουδήποτε μέσα στην εικόνα. Η συνάρτηση $\varphi_M(t) = \sum_{m\in\mathbb{Z}} \varphi(t - Mm)$ αποτελεί τη βάση φ της παραμετρικής καμπύλης.

Το μοντέλο του Splinedist είναι γενικό και ενοποιητικό. Μπορεί να χρησιμοποιηθεί οποιαδήποτε βάση spline για τη συνάρτηση ϕ, τυπικά πολυωνυμικές B-splines οποιασδήποτε τάξης. Παρακάτω, ορίζεται ως $\text{SplineDist}_{\beta^n}$ η έκδοση του Splinedist που βασίζεται σε $\phi = \beta^n$, δηλαδή η πολυωνυμική βάση B-spline του βαθμού n. Με αυτό τον τρόπο, ο Splinedist προβλέπει μία πιθανότητα για κάθε εικονοστοιχείο (i,j) στην εικόνα καθώς και ένα σύνολο M δισδιάστατων σημείων ελέγχου $c_{ij}[k]$, τα οποία ορίζουν την παραμετρική καμπύλη. Τα σημεία ελέγχου $c_{ij}[k]$ εκφράζονται σε πολική μορφή. Πιο συγκεκριμένα, ο Splinedist προβλέπει M γωνίες $a_i j^k$ και M ακτινικές αποστάσεις d_{ij}^k τέτοιες ώστε να ισχύει :

$$c_{ij}^{[k]} = \left(d_{ij}^k \cos(\alpha_{ij}^k), d_{ij}^k \sin(\alpha_{ij}^k)\right) \tag{3.5}$$

Η κατασκευή αυτή αποτελεί ένα πιο ευέλικτο μοντέλο, σε σχέση με τα πολύγωνα που κατασκευάζονται στο μοντέλο Stardist, καθώς οι γωνίες δεν είναι προκαθορισμένες και τα σημεία ελέγχου μπορούν να τοποθετηθούν σε οποιαδήποτε διεύθυνση, σε σχέση με τα εικονοστοιχεία (i,j). Ορίζεται ως:

$$S_{ij} = \left\{ s_{ij}^n \right\}_{n=0,\ldots,N-1} \tag{3.6}$$

όπου N το σύνολο των διακεκριμένων σημείων που παράγονται από μία ομοιόμορφη δειγματοληψία $S_{ij}(t)$ ως:

$$s_{ij}^n = s_{ij}(t)\Big|_{t=\frac{nM}{N}}. \tag{3.7}$$

Αρκετά σύνολα σημείων ελέγχου μπορεί να σχεδιάσουν το ίδιο περίγραμμα, για αυτό το λόγο το σφάλμα του αλγορίθμου SplineDist χρησιμοποιείται για να αξιολογήσει την ομοιότητα ανάμεσα στα S_{ij}, δηλαδή τη διακεκριμένη γραμμή περιγράμματος του αντικειμένου που δημιουργείται από τα σημεία ελέγχου $c_{ij}[k]_{k=0,\ldots,M-1}$ και στην πραγματική γραμμή περιγράμματος $P_{ij} = \{p_{ij}^n\}_{n=0,\ldots,N-1}$ που εξάγεται από ένα στιγμιότυπο μάσκας.

Συνοπτικά, το δίκτυο U-Net του Spline προβλέπει τις παραμέτρους της συνεχούς παραμετρικής καμπύλης, οι οποίες όταν ομαδοποιηθούν απεικονίζουν την πραγματική γραμμή περιγράμματος του αντικειμένου.

Το σφάλμα του SplineDist εκφράζεται ως:

$$L_{\text{SplineDist}}(p_{ij}, \hat{p}_{ij}, P_{ij}, S_{ij}) = L_{\text{BCE}}(p_{ij}, \hat{p}_{ij}) + \lambda_1 L_{\text{contour}}(p_{ij}, P_{ij}, S_{ij}) \tag{3.8}$$

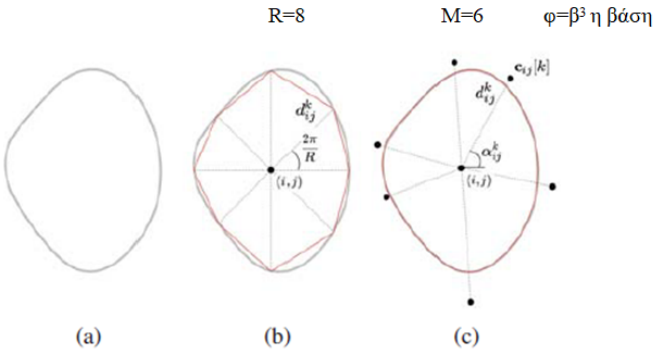
όπου

$$L_{\text{contour}}(p_{ij}, P_{ij}, S_{ij}) = \left(p_{ij} \cdot \mathbf{1}_{p_{ij}>0} \cdot \frac{1}{N} \sum_{n=0}^{N-1} \left| p_{ij}^n - s_{ij}^n \right| \right) + \lambda_2 \cdot \mathbf{1}_{p_{ij}=0} \cdot \frac{1}{R} \sum_{n=0}^{N-1} |S_{ij}^n| \tag{3.9}$$

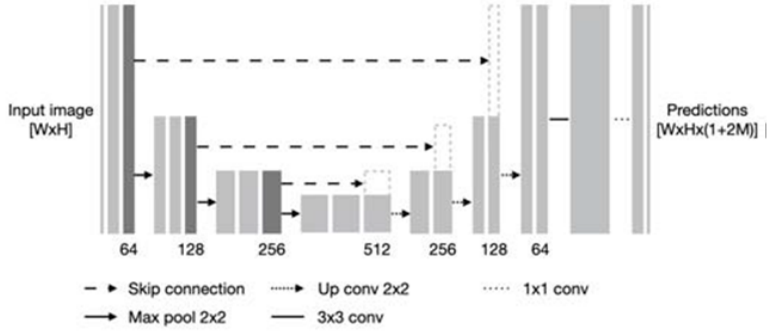
όπου p_{ij} είναι η πραγματική πιθανότητα και $\hat{p}_{ij}$ είναι η προβλεπόμενη πιθανότητα.

Η διαφορά με τον Stardist είναι ότι σε αυτή την περίπτωση δεν απαιτείται κάθε μεμονωμένο pixel(i,j) μέσα στην εικόνα να έχει τη δική του πραγματική τιμή (ground truth). Όλα τα εικονοστοιχεία (i, j) που περικλείονται στο ίδιο αντικείμενο O μπορεί πράγματι να μοιράζονται ένα ενιαίο περίγραμμα ($p_{ij} = p_O$ για όλα τα $(i, j) \in O$ εκφρασμένο σε απόλυτες συντεταγμένες. Το περίγραμμα Sij που προβλέπει κάθε pixel μπορεί επίσης να εκφραστεί σε απόλυτες συντεταγμένες, μετατοπίζοντας την καμπύλη spline Sij γύρω από τα σημεία (i,j). Έτσι, είναι εφικτή η αποθήκευση των συντεταγμένων περιγράφοντας την πραγματική τιμή (ground truth) για κάθε pixel. Η υπόλοιπη αρχιτεκτονική του μοντέλου Splinedist είναι παρόμοια με αυτή του

Stardist, με τη μόνη διαφορά να εντοπίζεται στο μέγεθος του τελευταίου επιπέδου εξόδου [7]. Στο σχήμα 3.1 απεικονίζονται οι εφαρμογές των αλγορίθμων Stardist και Splinedist σε ένα στιγμιότυπο αντικειμένου. Στο σχήμα 3.2 απεικονίζεται η αρχιτεκτονική του Splinedist και τέλος, στο σχήμα 3.3 απεικονίζεται η δομή του δικτύου U-Net.

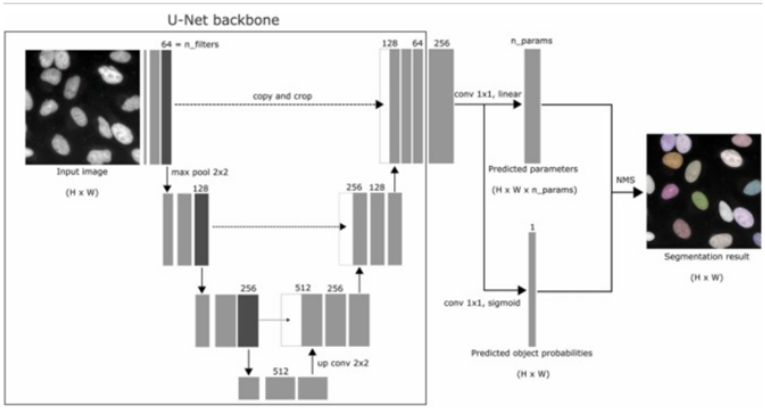


Σχήμα 3.1:
(α) Στιγμιότυπο αντικειμένου
(β) Ο Stardist δημιουργεί για κάθε εικονοστοιχείο (i, j) ένα πολύγωνο σχήματος αστεριού στο περίγραμμα του αντικειμένου, υπολογίζοντας τις ακτινικές αποστάσεις d_{ijk}, με προκαθορισμένες γωνίες ίσου μήκους $\frac{2\pi}{R}$.
(γ) Ο Splinedist δημιουργεί μία παραμετρική καμπύλη η οποία παράγεται από M σημεία ελέγχου (control points) $c_{ij}[k] = (d_{ijk}\cos(a_{ijk}), d_{ijk}\sin(a_{ijk}))$ [7].



Σχήμα 3.2: Αρχιτεκτονική Splinedist

Ο Splinedist δέχεται μία εικόνα και προβλέπει για κάθε εικονοστοιχείο(i, j) μία πιθανότητα p_{ij}και ένα σύνολο M δισδιάστατων σημείων ελέγχου $\{c_{ij}[k]\}_{k=0,\ldots,M-1}$. Η συνολική δομή του Splinedist είναι ίδια με του δικτύου του Stardist, με τη μόνη διαφορά να είναι στις διαστάσεις του τελευταίου στρώματος του νευρωνικού δικτύου, όπου $W \times H \times (1 + R)$ για τον Stardist και $W \times H \times (1 + 2M)$ για τον Splinedist [7].



Σχήμα 3.3: Απεικόνιση της δομής U-Net Backbone [8].

Κεφαλαιο 4

Συνολο δεδομενων SIPAKMED

Για την εκτέλεση των πειραμάτων χρησιμοποιήθηκε η βάση δεδομένων SIPakMed (Cervical Cancer Largest dataset), η οποία αποτελείται από 4049 εικόνες απομονωμένων κυττάρων που έχουν περικοπεί χειροκίνητα από 966 εικόνες κυττάρων cluster διαφανειών. Οι εικόνες αυτές αποκτήθηκαν μέσω μίας κάμερας CCD προσαρμοσμένης σε ένα οπτικό μικροσκόπιο. Οι εικόνες κυττάρων διακρίνονται σε πέντε κατηγορίες κυττάρων, οι οποίες περιλαμβάνουν φυσιολογικά, μη φυσιολογικά και καλοήθη κύτταρα. Τα φυσιολογικά κύτταρα αποτελούν την πλειοψηφία κυττάρων στο τεστ ΠΑΠ. Αποτελούνται από δύο κατηγορίες κυττάρων, τα επιφανειακά-ενδιάμεσα και τα παραβασικά. Στη συνέχεια, τα μη φυσιολογικά αλλά όχι κακοήθη κύτταρα αποτελούνται και αυτά από δύο κατηγορίες, τα κοιλοκύτταρα και τα δυσκερατοτικά. Τέλος, τα καλοήθη κύτταρα ονομάζονται αλλιώς και μεταπλαστικά και αποτελούν τη ζώνη μετασχηματισμού, δηλαδή την περιοχή στην οποία σχεδόν όλα τα προκαρκινικά και καρκινικά κύτταρα αναπτύσσονται. Οι πέντε κατηγορίες κυττάρων απεικονίζονται στο σχήμα 4.1.

Σχήμα 4.1: Εικόνες κυττάρων πέντε κατηγοριών :(α) Επιφανειακά-Ενδιάμεσα (β) Παραβασικά (γ) Κοιλοκύτταρα (δ) Δυσκερατωσικά ε) Μεταπλαστικά [9]

4.1 Φυσιολογικά κύτταρα

Τα φυσιολογικά κύτταρα αποτελούν πλακώδη επιθηλιακά και ο τύπος τους ορίζεται ανάλογα με τη θέση τους στα επιθηλιακά στρώματα και το βαθμό ωρίμανσής τους.

4.1.1 Επιφανειακά-ενδιάμεσα κύτταρα

Τα επιφανειακά-ενδιάμεσα κύτταρα αποτελούν την πλειονότητα των κυττάρων σε ένα τεστ ΠΑΠ. Συνήθως είναι επίπεδα με στρογγυλό, οβάλ ή πολυγωνικό σχήμα. Περιέχουν ένα κεντρικό πυκνωτικό πυρήνα. Έχουν καλά καθορισμένο, μεγάλο πολυγωνικό κυτταρόπλασμα και εύκολα αναγνωρίσιμα πυρηνικά όρια (μικρά πυκνωτικά σε επιφανειακούς και φυσαλιώδεις πυρήνες στα ενδιάμεσα κύτταρα. Αυτοί οι τύποι κυττάρων παρουσιάζουν τις μορφολογικές αλλαγές λόγω πιο σοβαρών βλαβών.

4.1.2 Παραβασικά κύτταρα

Τα παραβασικά κύτταρα είναι ανώριμα πλακώδη κύτταρα και είναι τα μικρότερα επιθηλιακά κύτταρα που φαίνονται σε ένα τυπικό κολπικό επίχρισμα. Το κυτταρόπλασμα είναι κυανόφιλο και συνήθως περιέχει ένα μεγάλο φυσαλιώδη πυρήνα. Αξίζει να σημειωθεί ότι τα παραβασικά κύτταρα έχουν παρόμοια μορφολογικά χαρακτηριστικά με τα κύτταρα που αναγνωρίζονται ως μεταπλαστικά και δύσκολα διακρίνονται από αυτά.

4.2 Μη φυσιολογικά κύτταρα

Τα μη φυσιολογικά κύτταρα χαρακτηρίζονται από μορφολογικές αλλαγές στη δομή τους και αναδεικνύουν την ύπαρξη παθολογικών καταστάσεων.

4.2.1 Κοιλοκύτταρα

Τα κοιλοκύτταρα αποτελούν συχνότερα ώριμα πλακώδη κύτταρα (ενδιάμεσα και επιφανειακά) και μερικές φορές μεταπλαστικά κοιλοκυτταρικά κύτταρα. Συνήθως είναι πολύ ελαφρά χρωματισμένα και χαρακτηρίζονται από μεγάλη περιπυρηνική κοιλότητα. Ωστόσο, η περιφέρεια του κυτταροπλάσματος είναι πολύ πυκνά χρωματισμένη. Οι πυρήνες των συγκεκριμένων κυττάρων είναι συνήθως διάσπαρτοι, υπερχρωματικοί και παρουσιάζουν ανωμαλία του περιγράμματος της πυρηνικής μεμβράνης. Υπάρχουν περιπτώσεις κυττάρων όπου ανιχνεύονται δύο πυρήνες ή και περισσότεροι. Μία συγκεκριμένη μόλυνση από τον ιό HPV μπορεί να ανιχνευθεί στα κοιλοκύτταρα, καθώς ανιχνεύεται στον πυρήνα μία μορφή εκφυλισμού, ανάλογα με το στάδιο της μόλυνσης.

4.2.2 Δυσκερατωσικά κύτταρα

Τα δυσκερατωσικά κύτταρα αποτελούν πλακώδη κύτταρα που υπέστησαν πρόωρα μη φυσιολογική κερατινοποίηση εντός μεμονωμένων κυττάρων ή σε τρισδιάστατα συμπλέγματα. Οι πυρήνες των συγκεκριμένων κυττάρων περιέχουν φυσαλίδες και είναι πανομοιότυποι με τους πυρήνες των κοιλοκυκυττάρων. Αποτελούν χαρακτηριστικό της μόλυνσης από τον ιό HPV και πολλές φορές είναι μία παθογνωμονική ένδειξη. Τέλος, συνήθως είναι τρισδιάστατες, παχιές στοιβάδες, που είναι αρκετά δύσκολο να ξεχωρίσει κανείς είτε τον πυρήνα, είτε τα όρια του κυτταροπλάσματος.

4.3 Καλοήθη κύτταρα

Τα καλοήθη κύτταρα αποτελούν μία ζώνη μετασχηματισμού, καθώς σε αυτά αναπτύσσονται όλα τα προκαρκινικά και καρκινικά κύτταρα.

4.3.1 Μεταπλαστικά κύτταρα

Τα μεταπλαστικά κύτταρα είναι μικρά ή μεγάλα κύτταρα παραβασικού τύπου με εμφανή κυτταρικά όρια, που συχνά εμφανίζουν έκκεντρους πυρήνεςκαι μερικές φορές περιέχουν ένα μεγάλο ενδοκυτταρικό κενοτόπιο. Η χρώση στο κεντρικό τμήμα είναι συνήθως ανοιχτό καφέ και συχνά διαφέρει από αυτό στο οριακό τμήμα. Επίσης, υπάρχει ένα πιο σκουρόχρωμο κυτταρόπλασμα και παρουσιάζουν μεγάλη

ομοιομορφία μεγέθους και σχήματος σε σύγκριση με τα παραβασικά κύτταρα. Η παρουσία τους στο τεστ Παπανικολάου σχετίζεται με υψηλά ποσοστά ανίχνευσης προκαρκινικών βλάβες (HSIL) [9].

Κεφαλαιο 5

Εκτελεση πειραματων

Σκοπός των πειραμάτων είναι η ανίχνευση του πυρήνα των κυττάρων μοντελοποιώντας τον ως παραμετρική καμπύλη, χρησιμοποιώντας το μοντέλο του Splinedist,το οποίο εφαρμόζεται στη συγκεκριμένη βάση κώδικα `https://github.com/uhlmanngroup/splinedist`. Η αξιολόγηση της κατάτμησης γίνεται στο σύνολο δεδομένων SIPaKMed. Το συγκεκριμένο σύνολο δεδομένων, αποτελείται από 4049 εικόνες συνολικά, όπου 3239 είναι το σύνολο των εικόνων που χρησιμοποιείται για εκπαίδευση (training set) και 810 είναι το σύνολο των εικόνων που χρησιμοποιείται για έλεγχο (test set). Χρησιμοποιούνται τυχαίες υπερπαράμετροι, παρόμοιες με τον Stardist, εφόσον αποτελούνται από την ίδια δομή δικτύου U-Net κι εφαρμόζεται η τεχνική επαύξησης δεδομένων (data augmentation), όπου οι εικόνες περιστρέφονται, κλιμακώνονται, αναστρέφονται ή περικόπτονται. Επιπλέον, οι εικόνες εκπαίδευσης χωρίζονται σε 2753 εικόνες εκπαίδευσης και 486 εικόνες αξιολόγησης (validation set). Οι εικόνες αρχικά είχαν μέγεθος (256 ,256) και ήταν έγχρωμες, ωστόσο για για τις ανάγκες της συγκεκριμένης εργασίας, πραγματοποιήθηκε μείωση των διαστάσεων των εικόνων σε (128, 128) και μετατροπή σε γκρι κλίμακα. Επιπροσθέτως, χρησιμοποιούνται προεπιλεγμένες ρυθμίσεις εκπαίδευσης, όπως για παράδειγμα, ρυθμός εκπαίδευσης ίσος με 3×10^{-4}, μέγεθος παρτίδας ίσος με 4 και όσον αφορά τις επαναλήψεις, το μοντέλο εκπαιδεύτηκε για 200 εποχές, με ακτίνες M=8, M=16 και M=32, αντίστοιχα. Η επαναληπτική διαδικασία της εκπαίδευσης προϋποθέτει τη συνέχιση των εποχών διατηρώντας τα βάρη της τελευταίας εποχής. Στη συνέχεια, δοκιμάστηκε η μέθοδος του πρόωρου τερματισμού (early stopping) προκειμένου να γίνει μία σύγκριση και να διερευνηθεί εάν υπάρχει βελτίωση στα αποτελέσματα ή

όχι. Λόγω μειωμένης χωρητικότητας και για τις ανάγκες της εργασίας, οι εικόνες φορτώνονταν σε παρτίδες, ανά 100 για κάθε εποχή ξεχωριστά. Τέλος, χρησιμοποιήθηκε το προεπιλεγμένο NMS (Non maximum suppresion) του StarDist και κατώφλια πιθανότητας αντικειμένου.Το NMS ορίζει μία τιμή, όπου οι προβλέψεις οι οποίες έχουν τιμή μικρότερη από αυτή, απορρίπτονται. Επίσης, το κατώφλι πιθανότητας ορίζει μία τιμή, όπου όταν δύο προβλέψιμα πολύγωνα επικαλύπτονται περισσότερο από αυτό το κατώφλι, τότε αυτό με τη μικρότερη τιμή καταστέλλεται. Για την αξιολόγηση της τεχνικής της αναπαράστασης των κυττάρων, χρησιμοποιήθηκε η μετρική IoU (Intersection over Union), η οποία αλλιώς ονομάζεται και δείκτης Jaccard. Η μετρική αυτή ορίζεται από τον τύπο:

$$\text{IoU} = \frac{\text{Περιοχή επικάλυψης}}{\text{Περιοχή ένωσης}}, \tag{5.1}$$

όπου η περιοχή επικάλυψης είναι ο αριθμός των κοινών θετικών εικονοστοιχείων και στην μάσκα που προκύπτει μετά από την πρόβλεψη και στη μάσκα αληθείας, ενώ περιοχή ένωσης είναι ο συνολικός αριθμός των εικονοστοιχείων που είναι θετικά είτε στην μάσκα που προκύπτει από την πρόβλεψη, είτε στη μάσκα αληθείας. Διαφορετικά, για δυαδική ταξινόμηση, η μετρική δίνεται από τον τύπο:

$$\text{Intersection over Union (IoU)} = \frac{\text{TP}}{\text{TP} + \text{FN} + \text{FP}} \tag{5.2}$$

όπου

$$\begin{aligned} \text{TP} &= \text{True Positive (Αληθώς Θετικά)} \\ \text{FN} &= \text{False Negative (Ψευδώς Αρνητικά)} \\ \text{FP} &= \text{False Positive (Ψευδώς Θετικά)} \end{aligned}$$

Τιμές κοντά στο 1 δηλώνουν ότι η πρόβλεψη ταιριάζει ακριβώς στην πραγματική εικόνα, ενώ τιμές από 0.7 και άνω, θεωρούνται ισχυρές [26].

Κεφαλαιο 6

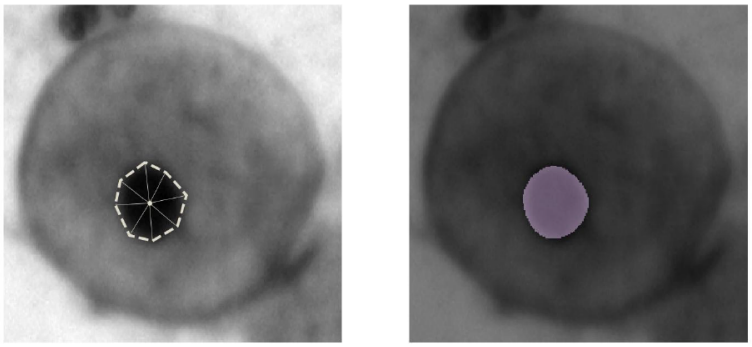
Αποτελεσματα πειραματων

Σε αυτό το κεφάλαιο παρουσιάζονται τα αποτελέσματα των πειραμάτων που εκτελέστηκαν, για αριθμό ακτινών 8, 16 και 32 αντίστοιχα.

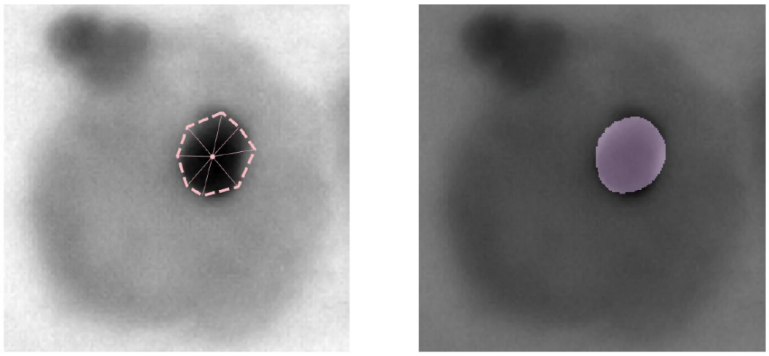
6.1 Αποτελέσματα πειραμάτων για 200 εποχές

6.1.1 Αποτελέσματα πειραμάτων για αριθμό ακτινών M=8

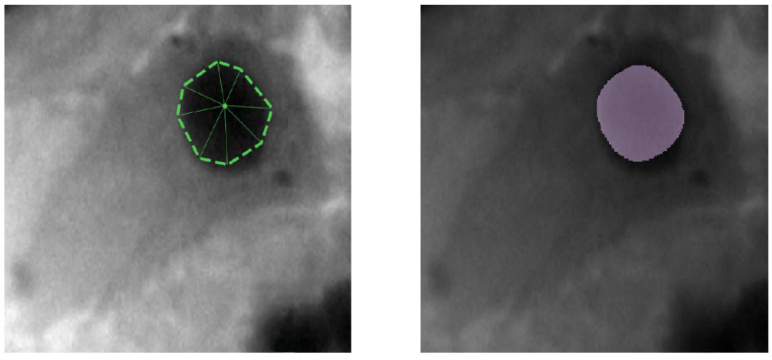
Στη συγκεκριμένη περίπτωση, ορίστηκε τιμή πιθανότητας ίση με 0.66 και τιμή NMS ίση με 0.3. Παρακάτω παρουσιάζονται ορισμένα παραδείγματα εικόνων κυττάρων, στις οποίες έχει εφαρμοστεί το μοντέλο Splinedist.



Σχήμα 6.1: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



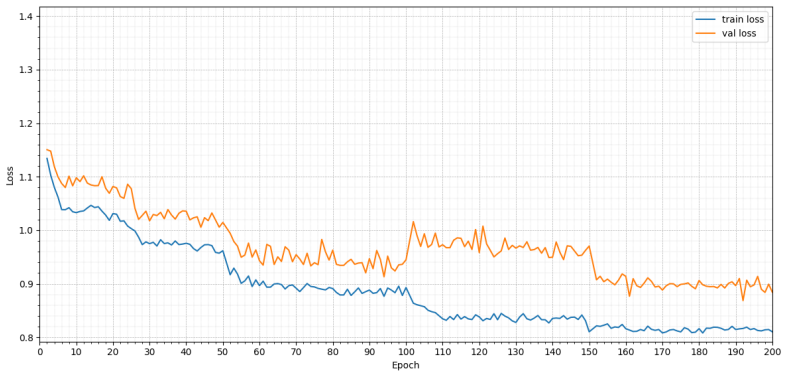
Σχήμα 6.2: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



Σχήμα 6.3: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

Κατά την εκπαίδευση του μοντέλου, το συνολικό μέσο σφάλμα εκπαίδευσης (training loss) υπολογίστηκε σε 0.8104. Παράλληλα, αναλύθηκαν επιμέρους δείκτες απόδοσης, όπως το σφάλμα πρόβλεψης πιθανοτήτων και αποστάσεων, τα αποτελέσματα των οποίων παρατίθενται στο Παράρτημα Α. Τα επιμέρους αυτά μεγέθη

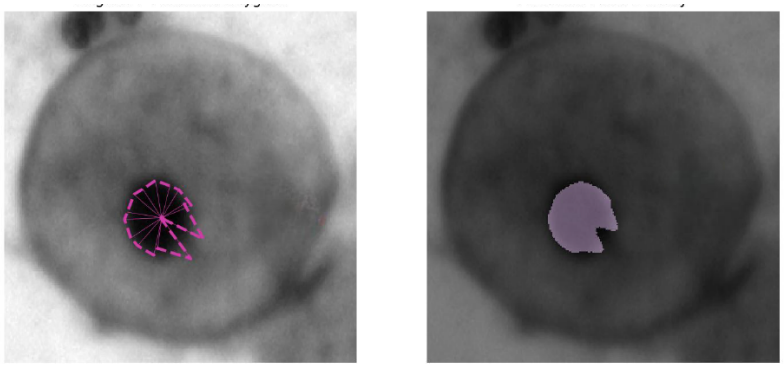
επιβεβαιώνουν τη συνολική καλή απόδοση του μοντέλου σε επίπεδο εκπαίδευσης. Το μοντέλο αξιολογήθηκε βάσει του Δείκτη Επικάλυψης (IoU) στο σύνολο ελέγχου (test set), ο οποίος αποτελεί βασικό μέτρο για την ποσοτικοποίηση της ακρίβειας εντοπισμού περιοχών ενδιαφέροντος. Η μέση τιμή IoU και η τυπική απόκλιση στο συγκεκριμένο δείγμα ανήλθε σε (0.71 ± 0.25). Το μοντέλο με 8 ακτίνες μαθαίνει γρήγορα να εντοπίζει αντικείμενα (χαμηλές απώλειες) αλλά παλεύει με τα ακριβή όρια (μεγαλύτερη απώλεια απόστασης) και προσαρμόζεται ελαφρώς. Στο σχήμα 6.4 απεικονίζεται η πορεία του μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 8 ακτίνες, όπου το μοντέλο εκπαιδεύτηκε για 200 εποχές. Στο σχήμα παρατηρείται μείωση και των δύο σφαλμάτων, με το validation loss να μη διαφέρει σημαντικά από το training loss, επομένως παρατηρείται μία καλή γενίκευση του μοντέλου.



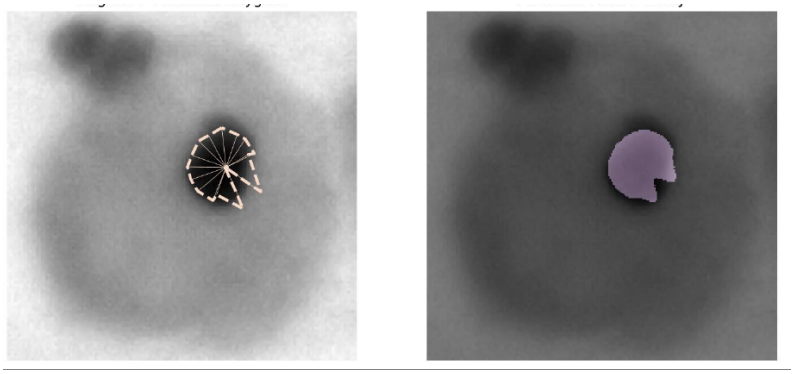
Σχήμα 6.4: Διάγραμμα απεικόνισης μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 8 ακτίνες

6.1.2 Αποτελέσματα πειραμάτων για αριθμό ακτινών M=16

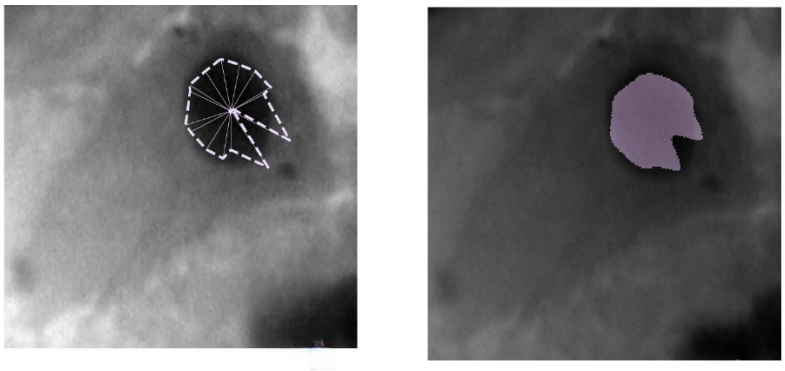
Στη συγκεκριμένη περίπτωση, ορίστηκε τιμή πιθανότητας ίση με 0.58 και τιμή NMS ίση με 0.3. Στη συνέχεια, παρουσιάζονται παραδείγματα εικόνων κυττάρων, όπου σχηματίζονται παραμετρικές καμπύλες μετά από εφαρμογή του μοντέλου Splinedist, με αριθμό ακτινών 16.



Σχήμα 6.5: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



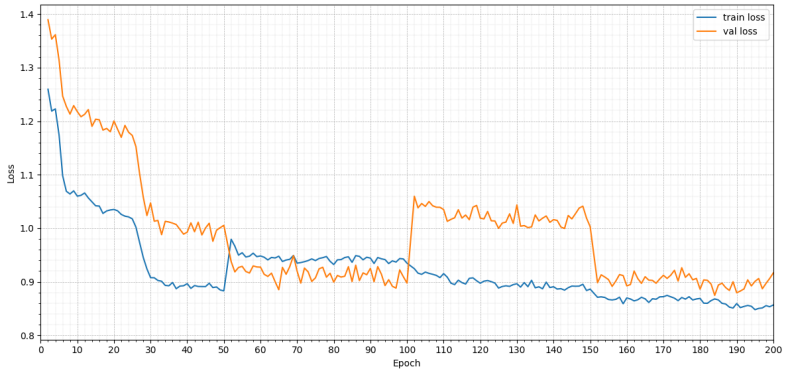
Σχήμα 6.6: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



Σχήμα 6.7: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

Κατά τη διάρκεια της εκπαίδευσης του μοντέλου, το συνολικό μέσο σφάλμα εκπαίδευσης (training loss) υπολογίστηκε σε 0.8567. Το μοντέλο γενικεύει ικανοποιητικά και δεν παρουσιάζει έντονες ενδείξεις υπερπροσαρμογής (overfitting). Υπολο-

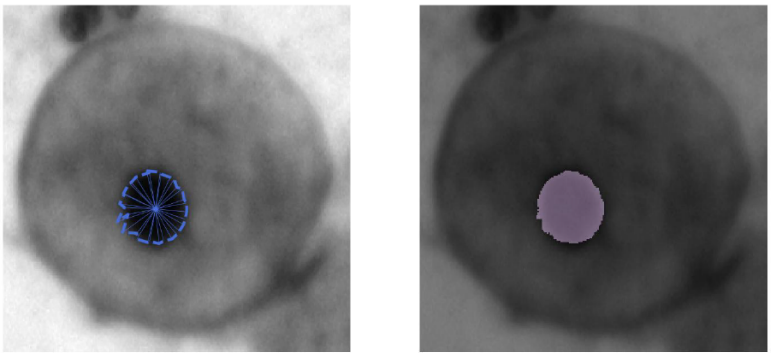
γίστηκαν και επιμέρους δείκτες απόδοσης, όπως το probability loss και το distance loss, οι οποίοι δείχνουν επίσης ικανοποιητικά αποτελέσματα και παρατίθενται αναλυτικά στο παράρτημα Α. Η μέση τιμή IoU και η τυπική απόκλιση στο συγκεκριμένο δείγμα ανήλθε σε (0.72 ± 0.19). Κατά μέσο όρο, η κατάτμηση έχει $\sim 72\%$ επικάλυψη με την πραγματικότητα, με σχετικά χαμηλή διακύμανση από αντικείμενο σε αντικείμενο. Στο σχήμα 6.8 απεικονίζεται η πορεία του μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 16 ακτίνες, όπου το μοντέλο εκπαιδεύτηκε για 200 εποχές. Στο σχήμα παρατηρείται σταθερή μείωση του training loss, ωστόσο το παρατηρείται αύξηση του validation loss στο διάστημα των 100-150 εποχών, ενώ το training loss συνεχίζει να μειώνεται, γεγονός που οδηγεί σε μέτριο overfitting.



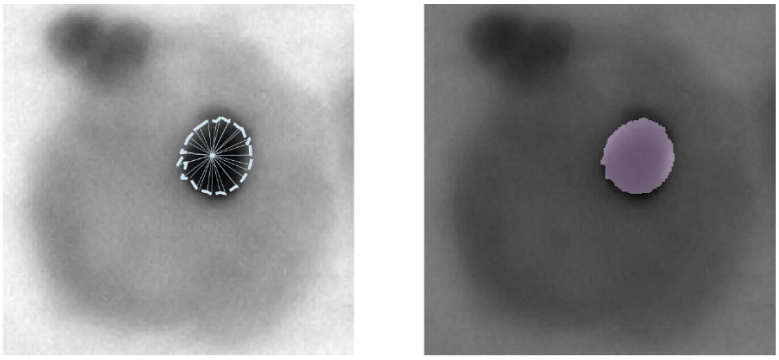
Σχήμα 6.8: Διάγραμμα απεικόνισης μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 16 ακτίνες

6.1.3 Αποτελέσματα πειραμάτων για αριθμό ακτινών M=32

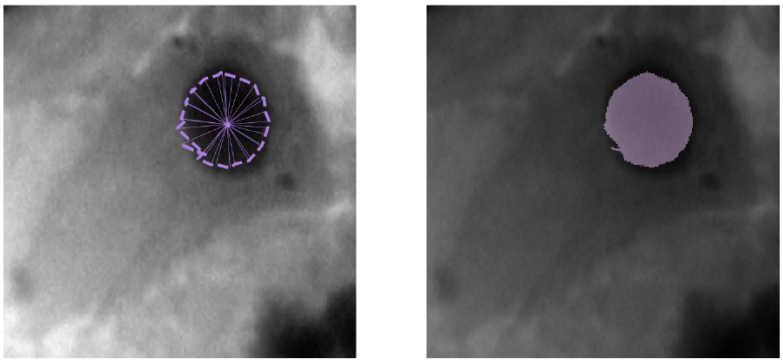
Τέλος, σε αυτή την περίπτωση ορίστηκε κατώφλι πιθανότητας ίσο με 0.45, και τιμή NMS ίση με 0.3. Παρακάτω παρουσιάζονται παραδείγματα εικόνων κυττάρων, όπου έχει εφαρμοστεί το μοντέλο Splinedist, για αριθμό ακτινών 32.



Σχήμα 6.9: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



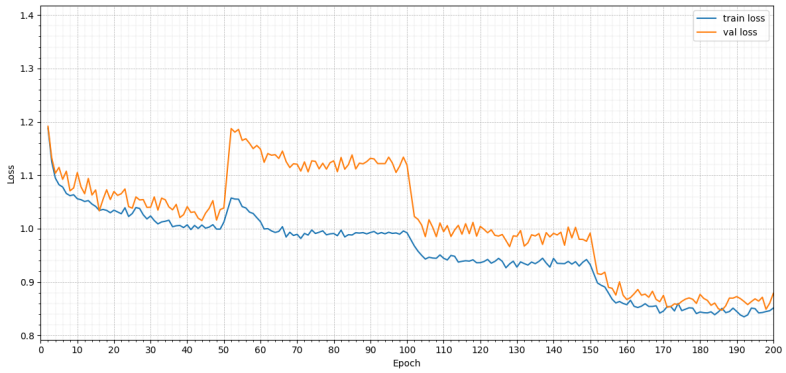
Σχήμα 6.10: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



Σχήμα 6.11: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

Κατά τη διάρκεια της εκπαίδευσης του μοντέλου, το συνολικό μέσο σφάλμα εκπαίδευσης (training loss) υπολογίστηκε σε 0.8475. Το μοντέλο έχει κάποια περιορισμένη υπερπροσαρμογή, ωστόσο υποδεικνύεται καλή γενίκευση. Ωστόσο, ορι-

σμένες μετρικές , οι οποίες παρατίθενται στο παράρτημα Α, όπως το distance loss και το MSE δείχνουν ότι το μοντέλο μπορεί να βελτιωθεί στην πρόβλεψη νέων δεδομένων. Η μέση τιμή IoU και η τυπική απόκλιση στο συγκεκριμένο δείγμα ανήλθε σε (0.74 ± 0.16). Στο σχήμα 6.12 απεικονίζεται η πορεία του μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 32 ακτίνες, όπου το μοντέλο εκπαιδεύτηκε για 200 εποχές. Στο σχήμα παρατηρείται συνολική μείωση του σφάλματος εκπαίδευσης και επικύρωσης, ωστόσο μετά τις 50 εποχές το σφάλμα επικύρωσης αρχίζει να αυξάνεται σημαντικά, πριν μειωθεί ξανά, γεγονός που οδηγεί σε μέτριο overfitting περίπου στις 50 εποχές.



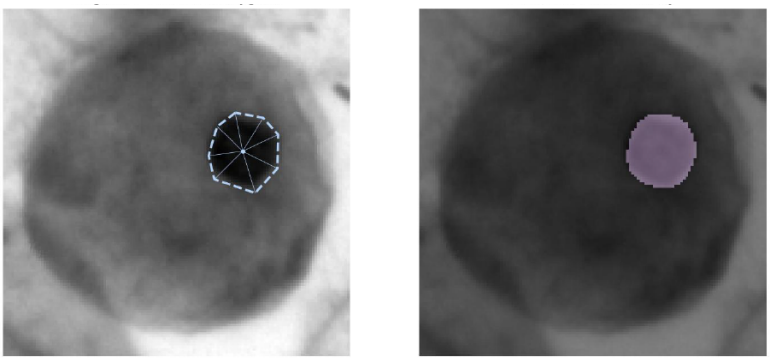
Σχήμα 6.12: Διάγραμμα απεικόνισης μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 32 ακτίνες

6.2 Αποτελέσματα πειραμάτων με τη μέθοδο πρόωρου τερματισμού (early stopping)

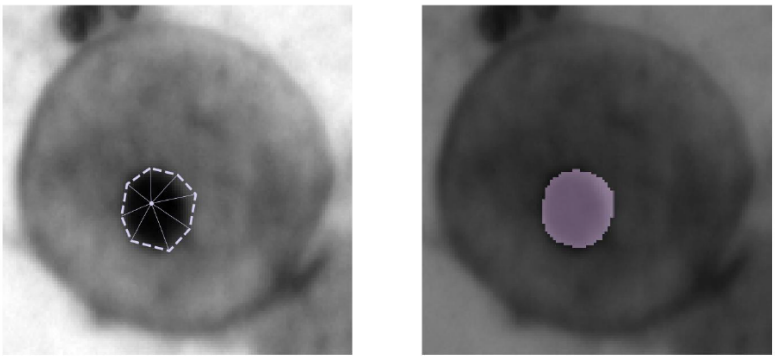
Το μοντέλο εκπαιδεύτηκε για 50 εποχές με τη μέθοδο πρόωρου τερματισμού, με κριτήριο τερματισμού τις 10 εποχές εάν δεν υπάρχει βελτίωση στα δεδομένα επικύρωσης (validation set), προκειμένου να αποφευχθεί η υπερπροσαρμογή (overfitting). Παρουσιάζονται τα αποτελέσματα που προκύπτουν μετά από εφαρμογή διαφορετικού αριθμού ακτινών, 8,16 και 32 αντίστοιχα, προκειμένου να παρατηρηθεί η επίδρασή τους.

6.2.1 Αποτελέσματα για αριθμό ακτινών ίσο με 8

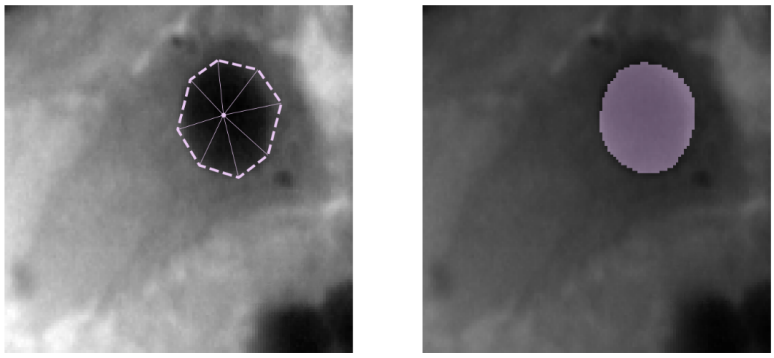
Στη συγκεκριμένη περίπτωση, η εκπαίδευση τερματίστηκε στις 32 εποχές, ενώ ορίστηκε τιμή πιθανότητας ίση με 0.46 και τιμή NMS ίση με 0.3. Στη συνέχεια, παρουσιάζονται παραδείγματα εικόνων κυττάρων, όπου σχηματίζονται παραμετρικές καμπύλες μετά από εφαρμογή του μοντέλου Splinedist, με αριθμό ακτινών 8.



Σχήμα 6.13: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

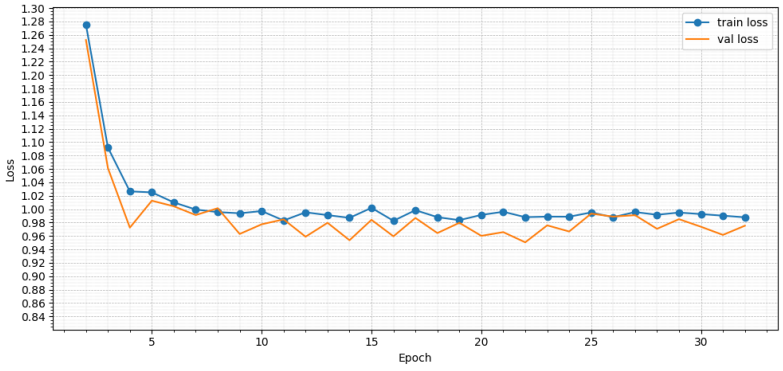


Σχήμα 6.14: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



Σχήμα 6.15: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

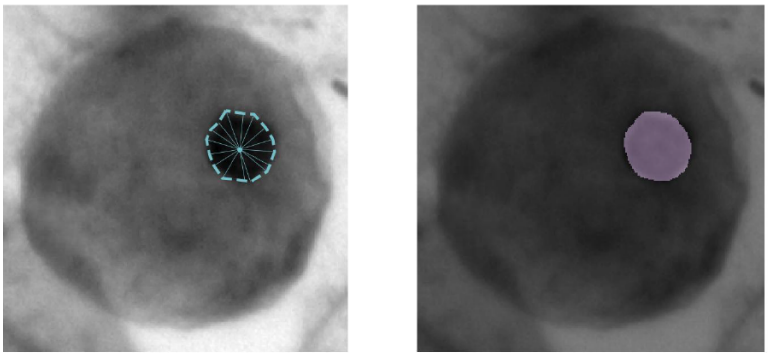
Κατά την εκπαίδευση του μοντέλου, το συνολικό καλύτερο μέσο σφάλμα εκπαίδευσης (training loss) εμφανίστηκε στις 22 εποχές και υπολογίστηκε σε 0.9887. Παράλληλα, αναλύθηκαν επιμέρους δείκτες απόδοσης, όπως το σφάλμα πρόβλεψης πιθανοτήτων και αποστάσεων, τα αποτελέσματα των οποίων παρατίθενται στο Παράρτημα Β. Τα επιμέρους αυτά μεγέθη επιβεβαιώνουν τη συνολική καλή απόδοση του μοντέλου σε επίπεδο εκπαίδευσης. Το μοντέλο αξιολογήθηκε βάσει του Δείκτη Επικάλυψης (IoU) στο σύνολο ελέγχου (test set), ο οποίος αποτελεί βασικό μέτρο για την ποσοτικοποίηση της ακρίβειας εντοπισμού περιοχών ενδιαφέροντος. Η μέση τιμή IoU και η τυπική απόκλιση στο συγκεκριμένο δείγμα ανήλθε σε (0.71 ± 0.17). Στο σχήμα 6.16 απεικονίζεται η πορεία του μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss), όπου η εκπαίδευση τερματίστηκε μετά από 32 εποχές. Στο σχήμα παρατηρείται εξαιρετική γενίκευση, δεν υπάρχουν σημάδια overfitting και υπάρχει σταθερή μείωση των σφαλμάτων.



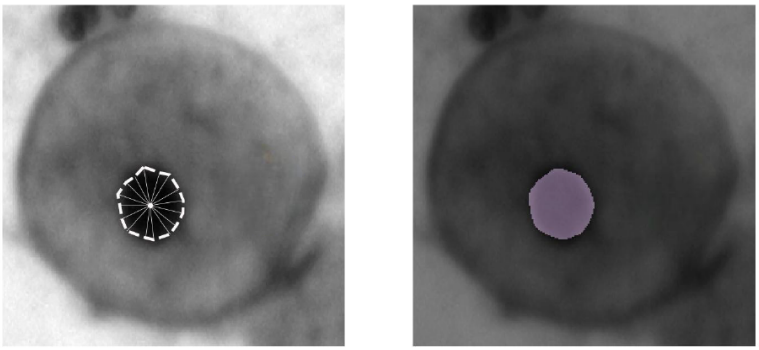
Σχήμα 6.16: Διάγραμμα απεικόνισης μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 8 ακτίνες

6.2.2 Αποτελέσματα για αριθμό ακτινών ίσο με 16

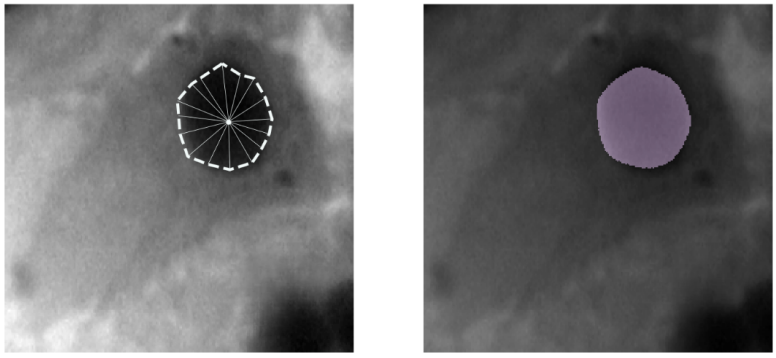
Στη συγκεκριμένη περίπτωση, η εκπαίδευση τερματίστηκε στις 23 εποχές, ενώ ορίστηκε τιμή πιθανότητας ίση με 0.58 και τιμή NMS ίση με 0.3. Στη συνέχεια, παρουσιάζονται παραδείγματα εικόνων κυττάρων, όπου σχηματίζονται παραμετρικές καμπύλες μετά από εφαρμογή του μοντέλου Splinedist, με αριθμό ακτινών 16.



Σχήμα 6.17: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

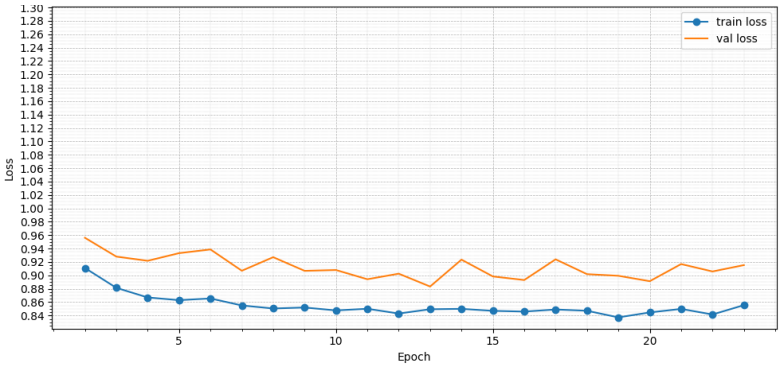


Σχήμα 6.18: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



Σχήμα 6.19: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

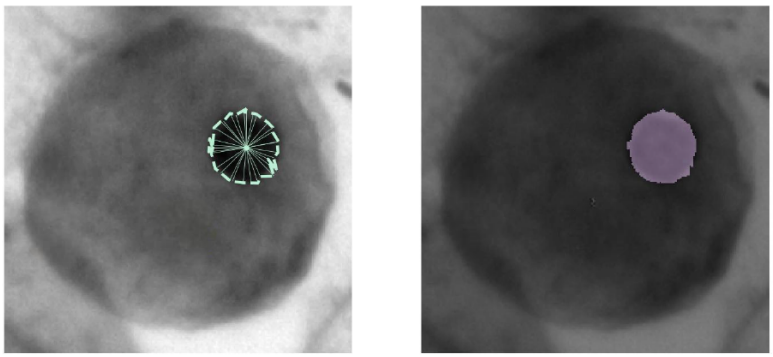
Κατά την εκπαίδευση του μοντέλου, το συνολικό καλύτερο μέσο σφάλμα εκπαίδευσης (training loss) εμφανίστηκε στις 13 εποχές και υπολογίστηκε σε 0.8508. Παράλληλα, αναλύθηκαν επιμέρους δείκτες απόδοσης, όπως το σφάλμα πρόβλεψης πιθανοτήτων και αποστάσεων, τα αποτελέσματα των οποίων παρατίθενται στο Παράρτημα Β. Τα επιμέρους αυτά μεγέθη επιβεβαιώνουν τη συνολική καλή απόδοση του μοντέλου σε επίπεδο εκπαίδευσης. Το μοντέλο αξιολογήθηκε βάσει του Δείκτη Επικάλυψης (IoU) στο σύνολο ελέγχου (test set), ο οποίος αποτελεί βασικό μέτρο για την ποσοτικοποίηση της ακρίβειας εντοπισμού περιοχών ενδιαφέροντος. Η μέση τιμή IoU και η τυπική απόκλιση στο συγκεκριμένο δείγμα ανήλθε σε (0.74 ± 0.18). Στο σχήμα 6.20 απεικονίζεται η πορεία του μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss), όπου η εκπαίδευση τερματίστηκε μετά από 23 εποχές. Στο σχήμα παρατηρείται ομαλή μείωση του training loss, με τη διαφορά των δύο σφαλμάτων να παραμένει μικρή και σταθερή.



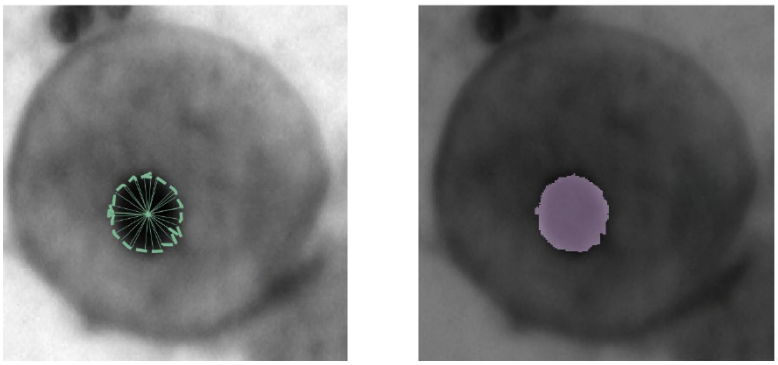
Σχήμα 6.20: Διάγραμμα απεικόνισης μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 16 ακτίνες

6.2.3 Αποτελέσματα για αριθμό ακτινών ίσο με 32

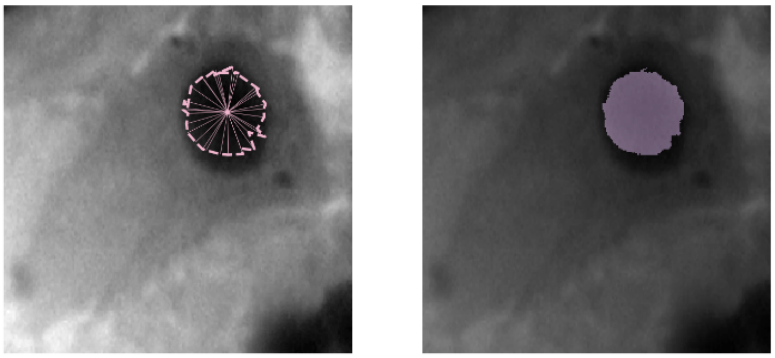
Στη συγκεκριμένη περίπτωση, η εκπαίδευση τερματίστηκε στις 49 εποχές, ενώ ορίστηκε τιμή πιθανότητας ίση με 0.44 και τιμή NMS ίση με 0.3. Στη συνέχεια, παρουσιάζονται παραδείγματα εικόνων κυττάρων, όπου σχηματίζονται παραμετρικές καμπύλες μετά από εφαρμογή του μοντέλου Splinedist, με αριθμό ακτινών 32.



Σχήμα 6.21: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

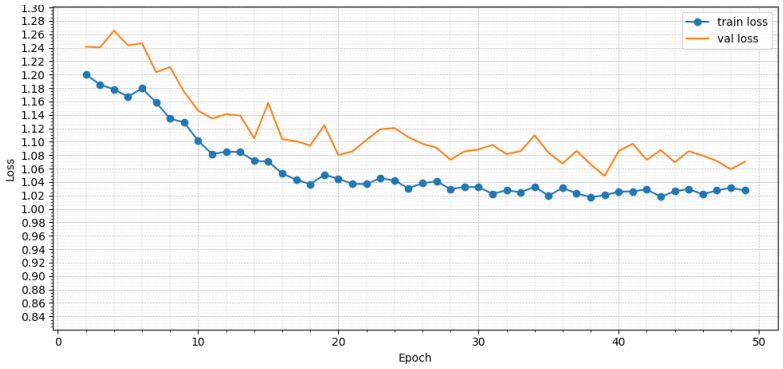


Σχήμα 6.22: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης



Σχήμα 6.23: (α) Σχηματισμός παραμετρικής καμπύλης στην πρωτότυπη εικόνα (β) Μάσκα πρόβλεψης

Κατά την εκπαίδευση του μοντέλου, το συνολικό καλύτερο μέσο σφάλμα εκπαίδευσης (training loss) εμφανίστηκε στις 39 εποχές και υπολογίστηκε σε 1.0331. Παράλληλα, αναλύθηκαν επιμέρους δείκτες απόδοσης, όπως το σφάλμα πρόβλεψης πιθανοτήτων και αποστάσεων, τα αποτελέσματα των οποίων παρατίθενται στο Παράρτημα Β. Το μοντέλο (test set) αξιολογήθηκε βάσει του Δείκτη Επικάλυψης (IoU), ο οποίος αποτελεί βασικό μέτρο για την ποσοτικοποίηση της ακρίβειας εντοπισμού περιοχών ενδιαφέροντος. Η μέση τιμή IoU και η τυπική απόκλιση στο συγκεκριμένο δείγμα ανήλθε σε (0.70 ± 0.19). Στο σχήμα 6.24 απεικονίζεται η πορεία του μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss), όπου η εκπαίδευση τερματίστηκε μετά από 49 εποχές. Στο σχήμα παρατηρείται σταθερή μείωση και των δύο σφαλμάτων, το validation loss να είναι σταθερά πιο υψηλό, ωστόσο η διαφορά δεν είναι σημαντικά μεγάλη.



Σχήμα 6.24: Διάγραμμα απεικόνισης μέσου σφάλματος εκπαίδευσης (training loss) και επικύρωσης (validation loss) για 32 ακτίνες

Κεφαλαιο 7

Συμπερασματα και Μελλοντικη Εργασια

Αρχικά, όσον αφορά τα πειράματα όπου το μοντέλο εκπαιδεύτηκε για 200 εποχές συνολικά, παρατηρήθηκε ότι, αυξάνοντας τον αριθμό παραμέτρων M από 8 σε 16, παρατηρείται αύξηση του δείκτη αξιολόγησης IoU (Intersection over Union) στο test set. Το μοντέλο με M=8 είχε μέσο όρο περίπου 0.71 (με τυπική απόκλιση 0.25), ενώ για M=16 ο μέσος όρος ανέβηκε στο 0.72 (με τυπική απόκλιση 0.19). Όμως, το μοντέλο με M=32 παρουσίασε καλύτερο μέσο όρο IoU (0.74) και μικρότερη τυπική απόκλιση (0.16), υποδεικνύοντας καλύτερη ταύτιση με τις πραγματικές περιοχές ενδιαφέροντος. Συνολικά, η αύξηση του αριθμού ακτινών οδηγεί ενδεχομένως σε καλύτερη απόδοση στο IoU, αλλά υπάρχει ο κίνδυνος υπερπροσαρμογής. Το M=32 φαίνεται να προσφέρει καλύτερη γενίκευση όσον αφορά την κατάτμηση. Αυτό σημαίνει ότι περισσότερες ακτίνες δίνουν μια πιο πλούσια αναπαράσταση σχήματος, αποδίδοντας στενότερη επικάλυψη με τις πραγματικές τιμές αντικειμένων. Παρακάτω, επισυνάπτεται ένας πίνακας, όπου γίνεται σύγκριση μετρικών απόδοσης για διαφορετικό αριθμό παραμέτρων M. Όσον αφορά τα πειράματα, όπου εφαρμόστηκε η τεχνική του early stopping, παρατηρήθηκε αύξηση του δείκτη αξιολόγησης IoU (Intersection over Union) στο test set, αυξάνοντας τον αριθμό παραμέτρων από 8 σε 16. Στην περίπτωση όπου ο αριθμός των παραμέτρων είναι ίσος με 16, φαίνεται να υπάρχει καλύτερη ταύτιση της πρόβλεψης με τις περιοχές ενδιαφέροντος, όπου ο μέσος όρος του δείκτη αξιολόγησης ανέρχεται στην τιμή 0.74 (με τυπική απόκλιση 0.18). Παρακάτω, επισυνάπτονται δύο πίνακες (πίνακες 7.1 και 7.2), όπου γίνεται σύγκριση μετρικών απόδοσης για διαφορετικό αριθμό παραμέτρων M. Συνολικά, η

τεχνική του early stopping είναι αποτελεσματική και οδηγεί σε καλύτερα αποτελέσματα, ειδικά στην περίπτωση με αριθμό παραμέτρων ίσο με 16, όπου βελτιώνεται σημαντικά το σχήμα της παραμετρικής καμπύλης, η οποία γίνεται πιο ομαλή, μοντελοποιώντας σωστά τον πυρήνα.

Πίνακας 7.1: Σύγκριση δεικτών απόδοσης για διαφορετικές τιμές M μετά από 200 εποχές

	Training	Testing	
		IoU	
M	Loss	Mean	Std
8	0.8104	0.71	0.25
16	0.8567	0.72	0.19
32	0.8475	0.74	0.16

Πίνακας 7.2: Σύγκριση δεικτών απόδοσης για διαφορετικές τιμές M με τη μέθοδο πρόωρου τερματισμού (early stopping)

	Training	Testing	
		IoU	
M	Loss	Mean	Std
8	0.9887	0.71	0.17
16	0.8508	0.74	0.18
32	1.0331	0.70	0.19

Προτείνονται τα παρακάτω, για μελλοντική εργασία:
Αρχικά, συστήνεται εφαρμογή πιο ενισχυμένων τεχνικών regularization (dropout, weight decay, L2-regularization) ώστε να μειωθεί η απόκλιση μεταξύ training/validation loss, ειδικά για M=32. Επιπλέον, προτείνεται η υιοθέτηση k-fold cross-validation για πιο αξιόπιστη εκτίμηση της γενίκευσης, και εντοπισμός συγκεκριμένων συνθηκών όπου το IoU πέφτει χαμηλά. Επιπροσθέτως, προτείνεται περαιτέρω εκπαίδευση για την περίπτωση των 32 παραμέτρων στην περίπτωση

του early stopping και εφαρμογή convex hull για την περίπτωση των 16 παραμέτρων, όπου παρουσιάζονται κάποια σημεία με κυρτότητα. Τέλος, προκειμένου να βελτιωθεί η σταθερότητα του μοντέλου, προτείνεται να δοκιμαστούν πιο σύνθετα μοντέλα αρχιτεκτονικής για καλύτερη γενίκευση.

Βιβλιογραφια

[1] DataCamp. (2024) Multilayer perceptrons in machine learning. [Online]. Available: https://www.datacamp.com/tutorial/multilayer-perceptrons-in-machine-learning?dc_referrer=https%3A%2F%2Fwww.google.com%2F

[2] A. Deshpande. (2017) A beginner's guide to understanding convolutional neural networks. [Online]. Available: https://adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/

[3] ——. (2017) A beginner's guide to understanding convolutional neural networks part 2. [Online]. Available: https://adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks-Part-2/

[4] Sarvesh Bhatt. (2019) What is artificial neural network and how does it work? [Online]. Available: https://techtalkwithbhatt.com/2019/01/23/what-is-artificial-neural-network-and-how-does-it-work/

[5] C. Prabha. (2024) Using gradient descent in cnn implementation. [Online]. Available: https://chithraprabhap.medium.com/using-gradient-descent-in-cnn-implementation-3c94029ffed8

[6] BimAnt. (2023) Adam optimizer intuition. [Online]. Available: http://www.bimant.com/blog/adam-optimizer-intuition/

[7] S. Mandal and V. Uhlmann, "Splinedist: Automated cell segmentation with spline curves," *bioRxiv*, 2020. [Online]. Available: https://www.biorxiv.org/content/10.1101/2020.10.27.357640v2

[8] D. Hirling and P. Horvath, "Cell segmentation and representation with shape priors," *Computational and Structural Biotechnology Journal*, vol. 21,

pp. 742–750, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2001037022005888

[9] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, and A. Charchanti, "Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images," in *Proceedings of the 25th IEEE International Conference on Image Processing (ICIP)*, Athens, Greece, 2018, pp. 3144–3148. [Online]. Available: https://www.cse.uoi.gr/~cnikou/Publications/C072%20-%20Plissiti%20-%20icip%202018%20-%20Athens.pdf

[10] S. Mandal and V. Uhlmann, "A learning-based formulation of parametric curve fitting for bioimage analysis," in *Proceedings of Numerical Mathematics and Advanced Applications – ENUMATH'19*, Egmond aan Zee, The Netherlands, Sep. 2021, pp. 1031–1038, conference held September 30–October 4, 2019. [Online]. Available: https://doi.org/10.1007/978-3-030-55874-1_102

[11] IBM. (2025) What is machine learning (ml)? [Online]. Available: https://www.ibm.com/think/topics/machine-learning

[12] ——. (2025) What is deep learning? [Online]. Available: https://www.ibm.com/think/topics/deep-learning

[13] GeeksforGeeks. (2025) Optimization rule in deep neural networks. [Online]. Available: https://www.geeksforgeeks.org/optimization-rule-in-deep-neural-networks/

[14] GeeksforGeeks. (2020) Ml | stochastic gradient descent (sgd). [Online]. Available: https://www.geeksforgeeks.org/ml-stochastic-gradient-descent-sgd/

[15] GeeksforGeeks. (2025) Difference between batch gradient descent and stochastic gradient descent. [Online]. Available: https://www.geeksforgeeks.org/difference-between-batch-gradient-descent-and-stochastic-gradient-descent/

[16] Piyush Kashyap. (2024) Understanding sgd with momentum in deep learning. [Online]. Available: https://medium.com/@piyushkashyap045/understanding-sgd-with-momentum-in-deep-learning-a-beginner-friendly-guide\-0252ede605b4

[17] GeeksforGeeks. (2025) Rmsprop optimizer in deep learning. [Online]. Available: https://www.geeksforgeeks.org/rmsprop-optimizer-in-deep-learning/

[18] J. Brownlee. (2017) Adam optimization algorithm for deep learning. [Online]. Available: https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/

[19] GeeksforGeeks. (2025) Adagrad optimizer in deep learning. [Online]. Available: https://www.geeksforgeeks.org/intuition-behind-adagrad-optimizer/

[20] P. Kashyap. (2023) Understanding rmsprop: A simple guide to one of deep learning's powerful optimizers. [Online]. Available: https://medium.com/@piyushkashyap045/understanding-rmsprop-a-simple-guide-to-one-of-deep-learnings-powerful\optimizers-403baeed9922

[21] S. Koli. (2023) Loss functions and their use in neural networks. [Online]. Available: https://medium.com/@MrBam44/loss-functions-and-their-use-in-neural-networks-5ca908d0e8fc

[22] V. Yathish. (2022) Loss functions and their use in neural networks. [Online]. Available: https://towardsdatascience.com/loss-functions-and-their-use-in-neural-networks-a470e703f1e9/

[23] Keras. (2025) Metrics. [Online]. Available: https://keras.io/api/metrics/

[24] V. Bushaev. (2017) How do we 'train' neural networks? [Online]. Available: https://medium.com/towards-data-science/how-do-we-train-neural-networks-edd985562b73

[25] stardist. (2025) Stardist: Object detection with star-convex shapes. [Online]. Available: https://github.com/stardist/stardist

[26] D. Shah. (2023) Intersection over union (iou): Definition, calculation, code. [Online]. Available: https://www.v7labs.com/blog/intersection-over-union-guide

Παραρτημα Α

Παραρτημα: Αναλυτικοι Δεικτες Αποδοσης Μοντελου

Απώλεια πιθανοτήτων (Probability loss) : Μετράει την απώλεια μεταξύ της πρόβλεψης και της αλήθειας, για κάθε pixel αν είναι μέρος ενός αντικειμένου (1) ή φόντου (0).

Απώλεια απόστασης [(Distance loss) : Το μοντέλο προσπαθεί να προβλέψει πόσο μακριά είναι το όριο του αντικειμένου από κάθε pixel. Η απώλεια απόστασης συγκρίνει αυτές τις προβλέψεις με τις πραγματικές αποστάσεις.

KL-απόκλιση πιθανοτήτων (Probability KL) : Κανονιστής (regularizer) που επιβάλλει την KL-απόκλιση (Kullback–Leibler divergence) ανάμεσα στη διανομή πιθανοτήτων που παράγει το μοντέλο και σε μια προκαθορισμένη κατανομή-στόχο.Η KL-απόκλιση μετράει πόσο η κατανομή που προβλέπει το μοντέλο "ξεφεύγει" από την ιδανική. Είναι σαν ποινή αν το μοντέλο διασκορπίζει τις πιθανότητές του πιο "άτακτα" από το αναμενόμενο.

Σχετικό MAE απόστασης (Distance relevant MAE) : Το μέσο απόλυτο σφάλμα (Mean Absolute Error, MAE) των προβλέψεων αποστάσεων, υπολογιζόμενο μόνο για τα εικονοστοιχεία που ανήκουν πραγματικά σε κάποιο αντικείμενο ($\text{mask} > 0$).

Σχετικό MSE απόστασης (Distance relevant MSE) : Το μέσο τετραγωνικό

σφάλμα (Mean Squared Error, MSE) των αποστάσεων και πάλι μόνο για εικονοστοιχεία εντός αντικειμένων.

Ρυθμός μάθησης (Learning rate) : Το βήμα (step size) που χρησιμοποιεί ο βελτιστοποιητής σε κάθε ενημέρωση βαρών. Δεν είναι δείκτης απόδοσης, αλλά καταγράφεται για την παρακολούθηση της διαδικασίας εκπαίδευσης. Ρυθμίζει "πόσο μεγάλα βήματα" κάνει το μοντέλο κάθε φορά που μαθαίνει από τα λάθη του. Αν είναι πολύ μεγάλο, μπορεί να ξεφύγει από τη βέλτιστη λύση, ενώ αν είναι πολύ μικρό, θα χρειάζεται πάρα πολλές επαναλήψεις.

Πίνακας Α.1: Αναλυτικοί δείκτες απόδοσης μοντέλου μετά από 200 εποχές

	Εκπαίδευση		
Δείκτης	**M=8**	**M=16**	**M=32**
Probability loss	0.048	0.050	0.052
Distance loss	3.606	3.606	3.729
Probability KLD	0.010	0.008	0.007
Distance relevant MAE	3.599	3.600	3.723
Distance relevant MSE	23.901	23.657	25.516
Learning rate	3×10^{-4}	3×10^{-4}	3×10^{-4}

Πίνακας Α.2: Αναλυτικοί δείκτες απόδοσης μοντέλου με την τεχνική του early stopping

	Εκπαίδευση		
Δείκτης	**M=8**	**M=16**	**M=32**
Probability loss	0.051	0.046	0.056
Distance loss	4.069	3.630	5.316
Probability KLD	0.009	0.009	0.013
Distance relevant MAE	4.063	3.624	5.309
Distance relevant MSE	29.398	24.266	56.174
Learning rate	3×10^{-4}	3×10^{-4}	3×10^{-4}

Σύντομο Βιογραφικο

Ονομάζομαι Πλακούτση Φωτεινή , γεννήθηκα στις 26 Αυγούστου 1997 και κατάγομαι από την Άρτα. Είμαι απόφοιτη του τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων, με ειδίκευση στον τομέα της στατιστικής και επιχειρισιακής έρευνας και συνέχισα τις σπουδές μου σε μεταπτυχιακό επίπεδο στο τμήμα Μηχανικών Η/Υ και Πληροφορικής του Πανεπιστημίου Ιωαννίνων, επιλέγοντας την κατεύθυνση της επεξεργασίας και ανάλυσης δεδομένων.