



Πολυτεχνική Σχολή
Τμήμα Μηχανικών Η/Υ & Πληροφορικής

ΕΙΣΑΓΩΓΗ ΣΤΗ ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ

**Μετασχηματίζοντας την Ανίχνευση του
Καρκίνου με Καινοτόμα Αρχιτεκτονική
Multicloud για Εξατομικευμένη Φροντίδα**

Φίλιππος-Παρασκευάς Ζυγούρης

A.M. 1084660

Πάτρα, 2024

Περιεχόμενα

1. Εισαγωγή.....	4
A. Τοποθέτηση Προβλήματος.....	4
B. Σπουδαιότητα.....	5
Γ. Συνεισφορά.....	6
2. Θεμελιώδεις Έννοιες και Τάσεις.....	7
1. Καρκίνος του πνεύμονα	7
2. Καρκίνος του μαστού	8
3. Καρκίνος του δέρματος (μελάνωμα).....	8
A. Τρέχουσες Λύσεις.....	9
Τεχνητή Νοημοσύνη	9
Μηχανική Μάθηση.....	10
Βαθιά Μάθηση.....	10
Τι είναι τα νευρωνικά δίκτυα ;.....	10
ANN	11
CNN	11
B. Πρακτικές.....	12
ANN - Καρκίνος Μαστού.....	12
CNN - Καρκίνος του Πνεύμονα.....	12
Γ. State-of-the-art.....	12
3. Προτεινόμενες τεχνικές.....	15
A. Θεωρητική περιγραφή	15
I. Παρουσίαση.....	15
Ψηφιακή Ιατρική Απεικόνιση	15
Computer Vision	15
Radiomics	15
Ψηφιακά Αρχεία Υγείας.....	16
Natural Language Processing (NLP).....	16
II. Περιγραφή	17
B. Υλοποίηση-Πειράματα.....	17
I. Μελέτη Περίπτωσης.....	17
Radiomics + Computer Vision	18

Εντοπισμός Σαγματικού Σημείου.....	18
Σχήμα 1.1: Σαγματικό Σημείο	18
LoG	19
Σχήμα 1.2: Laplacian of the Gaussian.....	19
Σχήμα 1.3: Εύρεση Μηδενικών Διασταυρώσεων.....	20
Σχήμα 1.4: Αναγνώριση Blob.....	21
DoG	21
Gaussian Pyramid.....	22
Laplacian Pyramid	22
Hessian-LoG	22
Σχήμα 1.5: Ορίζουσα Hessian	23
Ανάλυση Οριοθετημένων Περιοχών	23
GLCM.....	23
NLP	24
II. Πρακτική Εφαρμογή	24
CNN	24
Μοντέλο NLP	25
Ενοποιημένες Τοπικές Βάσεις Δεδομένων.....	25
Πολλαπλά Σύννεφα	26
Σχήμα 1.6: Εντοπισμός Σαγματικού Σημείου και Οριοθέτηση Περιοχών Ενδιαφέροντος.....	27
Σχήμα 1.7: Αγωγός Ανάπτυξης Μοντέλου NLP	28
Σχήμα 1.8: Ροή Εργασίας Ενοποιημένων Τοπικών Βάσεων Δεδομένων.....	29
Σχήμα 1.9: Ροή Εργασίας Πολλαπλών Σύννεφων	30
4. Μελλοντικές Βελτιώσεις και Επόμενα Βήματα	32
Α. Αντιμετώπιση ανησυχιών για AI στην Υγεία.....	32
Β. Μετρήσεις σε Πραγματικό Χρόνο	32
Γ. Καρκίνος, DNA και Εξατομικευμένη Θεραπεία	33
5. Σύνοψη και Συμπέρασμα	35
Σύνοψη	35
Συμπέρασμα	36
6. Βιβλιογραφία	38

1. Εισαγωγή

A. Τοποθέτηση Προβλήματος

Τα τελευταία χρόνια ιδιαίτερα με την ραγδαία ανάπτυξη της τεχνητής νοημοσύνης, τα δεδομένα, έχουν γίνει ένας άρρηκτος φυσικός πόρος που ολοένα αυξάνεται η αξία τους. Ο άνθρωπος μετριέται όλο και περισσότερο, χαρτογραφείται και καταγράφεται σε ψηφιακά κομμάτια (bits) για την αιωνιότητα. Επομένως, έχει γίνει επιτακτική η ανάγκη για σωστή ανάκτηση και χρήση της πληροφορίας που εξάγεται από τα δεδομένα, για να βελτιωθεί το πιο πολύτιμο αγαθό της ανθρώπινης ύπαρξης, η υγεία.

Η κύρια αιτία νοσηρότητας και θνησιμότητας παγκοσμίως είναι ο καρκίνος. Καθ' όλη την διάρκεια της ανθρώπινης ζωής, τα άτομα αιμελούν ότι υπάρχει πολύ σοβαρό ενδεχόμενο να κινδυνέψουν από την ανάπτυξη της νόσου. Εάν είναι τυχεροί, οι γιατροί μπορεί να ανακαλύψουν την μάζα και να την χειρουργήσουν πριν εξαπλωθεί. Αν δεν υπάρχει αυτό το τυχαίο εύρημα, ο καρκίνος πιθανότατα δεν θα εντοπιστεί και το ταξίδι τους στο κόσμο αυτό, θα φτάσει στο τέλος του. Συνεπώς, η διάγνωση του σε πρώιμο στάδιο αυξάνει την πιθανότητα αποτελεσματικής θεραπείας του.

Ωστόσο, ακόμη και με τακτικό έλεγχο, το πιο εξειδικευμένο ιατρικό προσωπικό δεν μπορεί να εντοπίσει τα πάντα, καθώς δεν διαθέτει τα κατάλληλα τεχνολογικά εργαλεία στην φαρέτρα του. Οι περισσότεροι άνθρωποι που διαγιγνώσκονται με καρκίνο παρουσιάζουν δυστυχώς μεταστατική νόσο, επειδή έχει εξαπλωθεί σε ένα μέρος που τους προκαλεί πόνο ή σύμπτωμα, οδηγώντας τους να επισκεφτούν τον ιατρό τους. Ως αποτέλεσμα, οι περισσότερες περιπτώσεις καρκίνου σε πρώιμο στάδιο δεν ανιχνεύονται, καθώς οι άνθρωποι δεν εμφανίζουν συμπτώματα και, επομένως, δεν προβαίνουν σε διαγνωστικό έλεγχο.

Συνεπώς, οι ιατροί αναζητούν εξογκώματα ή μάζες, όταν η νόσος ήδη βρίσκεται σε εξελικτικό στάδιο. Λόγω αυτής της αδυναμίας, που έχει κοστίσει εκατομμύρια ζωές, η ανθρωπότητα επικεντρώνεται στην ανάπτυξη κατάλληλων εργαλείων και μεθόδων για την έγκαιρη διάγνωσή της, ακόμη και πριν σχηματιστεί όγκος, ή παρατηρηθούν επώδυνα συμπτώματα.

B. Σπουδαιότητα

Η έγκαιρη ανίχνευση του καρκίνου επιτυγχάνεται μέσω εργαλείων όπως η τεχνητή νοημοσύνη και η μηχανική μάθηση. Η τεχνητή νοημοσύνη συμβάλλει σε αυτό, με την πρόβλεψη των πτυχών της νόσου και με την λήψη κατάλληλων αποφάσεων για το επίπεδο κινδύνου της. Αυτά πραγματοποιούνται μέσω του υποσυνόλου της, την μηχανική μάθηση, που επίσης διαδραματίζει πολύ σημαντικό ρόλο, καθώς επικεντρώνεται στη δημιουργία των αλγορίθμων που μαθαίνουν από τα περίπλοκα μοτίβα αυτών των βιολογικών δεδομένων και ενισχύουν την αποτελεσματικότητά τους μέσω της εμπειρίας τους πάνω σε αυτά, ώστε να κάνουν στοχευμένες προβλέψεις για το ποιος μπορεί να διατρέχει κίνδυνο μελλοντικά.

Συνεπώς, ενισχύουν τη διαγνωστική ακρίβεια εξετάζοντας πολύπλοκα ιατρικά δεδομένα, όπως ιατρικές εικόνες και αρχεία υγείας, εντοπίζοντας έτσι μικρές λεπτομέρειες που μπορεί να διαφέύγουν από το ανθρώπινο μάτι. Επιπλέον, εντοπίζουν ασθένειες στα εκκολαπτόμενα στάδια τους, φέρνοντας έτσι επανάσταση στην προληπτική υγειονομική περίθαλψη. Η έγκαιρη ανίχνευση είναι ζωτικής σημασίας, καθώς συχνά οδηγεί σε πιο αποτελεσματικές θεραπείες. Επίσης, ενισχύουν την επιλογή θεραπείας, διευκολύνοντας την ανάπτυξη εξατομικευμένων σχεδίων θεραπείας για κάθε ασθενή.

Η ενσωμάτωση της τεχνητής νοημοσύνης και της μηχανικής μάθησης στην ανίχνευση καρκίνου όχι μόνο ενισχύει την αποτελεσματικότητα, αλλά ανοίγει το δρόμο για ιατρική ακριβείας, μεταμορφώνοντας τελικά το τοπίο της ογκολογίας και βελτιώνοντας την ποιότητα της περίθαλψης των ασθενών παγκοσμίως. Αυτό δεν σημαίνει ότι θα αντικατασταθεί ο ανθρώπινος παράγοντας. Με την βοήθεια αυτών των τεχνικών το ιατρικό προσωπικό κάνει και βλέπει περισσότερα, και διαχειρίζεται καλύτερα την ασθένεια σε σύγκριση με την απουσία τους. Άρα κάθε δυνατό όπλο είναι απαραίτητο για την καταπολέμηση του καρκίνου, καθώς συμβάλλει σημαντικά στην διάσωση ζωών.

Γ. Συνεισφορά

Σκοπός της ερευνητικής αναφοράς είναι η παροχή δημιουργικών ιδεών, βελτιώσεων και συγχωνεύσεων στρατηγικών, που βασίζονται στη διεξοδική ανάλυση των διθέντων και περαιτέρω δημοσιεύσεων και του βιβλίου, με στόχο την εξατομικευμένη και βασισμένη σε δεδομένα διάγνωση. Κάθε ασθενής είναι μοναδικός, και ως εκ τούτου, οι διαγνωστικές διαδικασίες πρέπει να προσαρμόζονται στις ατομικές ανάγκες του. Για το σκοπό αυτό, προτείνονται αρκετές καινοτόμες ιδέες που περιστρέφονται γύρω από την προσαρμογή των διαγνωστικών διαδικασιών. Αυτές οι ιδέες πηγάζουν από τις πιο πρόσφατες εξελίξεις στην τεχνητή νοημοσύνη και την μηχανική μάθηση και έχουν σχεδιαστεί για να διασφαλίζουν ότι η διάγνωση δεν είναι μόνο ακριβής αλλά και ευθυγραμμισμένη με το συγκεκριμένο ιατρικό ιστορικό και τη γενετική σύνθεση του ασθενούς.

Επιπλέον, επισημαίνονται πολλές βελτιώσεις στις υπάρχουσες διαγνωστικές στρατηγικές. Αυτές οι βελτιώσεις στοχεύουν στην αντιμετώπιση των περιορισμών και των κενών στις τρέχουσες μεθοδολογίες, ώστε να βελτιωθεί η εγκυρότητα και η προγνωστική τους ισχύ. Αυτό, με τη σειρά του, μπορεί να μειώσει σημαντικά τον χρόνο που απαιτείται για τη διάγνωση, επιτρέποντας την έγκαιρη ανίχνευση και θεραπεία του καρκίνου. Επίσης, εκτός από τις μεμονωμένες ιδέες και βελτιώσεις, εξετάζονται οι δυνατότητες συγχώνευσης διαφορετικών στρατηγικών. Πιο συγκεκριμένα, συνδυάζονται τα δυνατά στοιχεία διαφόρων μεθοδολογιών, ώστε αυτές οι συγχωνεύσεις να οδηγήσουν σε μια πιο ολοκληρωμένη προσέγγιση για τη διάγνωση του καρκίνου. Αυτή η προσέγγιση όχι μόνο αυξάνει την πιθανότητα έγκαιρης ανίχνευσης, αλλά παρέχει επίσης μια πληρέστερη κατανόηση της νόσου, η οποία είναι ζωτικής σημασίας για την κατάρτιση αποτελεσματικών σχεδίων θεραπείας.

Συμπερασματικά, η υπάρχουσα μελέτη συμβάλλει στον τομέα της έγκαιρης διάγνωσης του καρκίνου παρέχοντας μια νέα προοπτική και καινοτόμες λύσεις. Αξιοποιώντας τη δύναμη της τεχνητής νοημοσύνης και της μηχανικής μάθησης, κύριος στόχος είναι η μετατροπή της διάγνωσης της νόσου από μια γενική, σε μια πιο εξατομικευμένη και βασισμένη σε δεδομένα διαδικασία.

2. Θεμελιώδεις Έννοιες και Τάσεις

Τα τελευταία δεδομένα από το Παγκόσμιο Παρατηρητήριο Καρκίνου δίνουν μια απογοητευτική εικόνα της εκτεταμένης επίδρασης της ασθένειας. Ο εκτιμώμενος αριθμός νέων περιπτώσεων εμφάνισης και θανάτων που προκλήθηκαν από καρκίνο απεικονίζουν την εμβέλεια αυτής της ασθένειας σε παγκόσμια κλίμακα. Μόνο το 2020, κάθε 60 δευτερόλεπτα, 37 άτομα παγκοσμίως λάμβαναν διάγνωση της νόσου που άλλαζε τη ζωή τους και περισσότερες από 19 ζωές χάνονταν. Αυτοί οι αριθμοί υπογραμμίζουν την επείγουσα ανάγκη για την καταπολέμηση του καρκίνου μέσω ενισχυμένης πρόληψης, έγκαιρης ανίχνευσης και βελτιωμένων θεραπειών σε παγκόσμιο επίπεδο.

Όπως δείχνουν τα δεδομένα, η ελαχιστοποίηση των καταστροφικών επιπτώσεων του καρκίνου απαιτεί μια συντονισμένη απάντηση από τη διεθνή κοινότητα. Η αποτελεσματική καταπολέμηση της νόσου απαιτεί τον εντοπισμό της ρίζα του προβλήματος, δηλαδή των παραγόντων που προκαλούν αυτή την καταστροφική ασθένεια. Ήτοι, η εστίασή έγκειται στην ανίχνευση αναγνωρισμένων καρκινογόνων ουσιών και στη διερεύνηση προληπτικών μέτρων που μπορούν να μετριάσουν τον κίνδυνο ανάπτυξης διαφόρων μορφών καρκίνου. Διασφαγνίζοντας τόσο τους παράγοντες ενεργοποίησης όσο και τα προστατευτικά μέτρα, μπορούν να κατανοθούν καλύτερα οι τρέχουσες λύσεις και οι πρακτικές αυτής της περίπλοκης πρόκλησης για την υγεία.

Η συγκεκριμένη έρευνα, επικεντρώνεται αποκλειστικά στους ακόλουθους τρεις τύπους καρκίνου:

1. Καρκίνος του πνεύμονα

Ο καρκίνος του πνεύμονα ξεχώρισε ως ο πιο θανατηφόρος το 2020. Ο απολογισμός του ήταν συγκλονιστικός, καθώς 1,8 εκατομμύρια ζωές χάθηκαν, που είναι περίπου το ένα πέμπτο όλων των θανάτων που σχετίζονται με καρκίνο εκείνο το έτος (18%). Η κυριαρχία του δεν περιορίζεται μόνο στα ποσοστά θνησιμότητας. Κατατάχθηκε ως ο δεύτερος πιο συχνά εμφανιζόμενος καρκίνος που διαγνώστηκε, με εντυπωσιακά 2,2 εκατομμύρια νέες περιπτώσεις να έχουν εντοπιστεί παγκοσμίως, που αποτελούν το 11,4% των συνολικών νέων περιπτώσεων καρκίνου για το έτος. Οι συγκλονιστικοί αυτοί αριθμοί υπογραμμίζουν γιατί ο καρκίνος του πνεύμονα πρέπει να είναι το επίκεντρο των ερευνητικών προσπαθειών, δεδομένου ότι πολλαπλασιάζεται ταχέως, ενώ ταυτόχρονα αποδεικνύεται εξαιρετικά θανατηφόρος.

Καρκινογόνες Ουσίες:

- Κάπνισμα:** Αυτή είναι η πιο κοινή αιτία καρκίνου του πνεύμονα. Περιλαμβάνει τόσο το ενεργό κάπνισμα όσο και την έκθεση στο παθητικό κάπνισμα.

- b) **Σκόνες αμιάντου, αρσενικού, καδμίου και πυριτίου:** Πρόκειται για επιβλαβείς ουσίες στις οποίες μπορεί να εκτεθούν οι άνθρωποι κατά την εργασία, ειδικά σε ορισμένες βιομηχανίες όπως οι κατασκευές και τα ορυχεία.
- c) **Εξατμίσεις κινητήρων ντίζελ, καύση άνθρακα, ρύπανση του εξωτερικού αέρα:** Αυτές είναι πηγές ατμοσφαιρικής ρύπανσης που μπορεί να αυξήσουν τον κίνδυνο καρκίνου του πνεύμονα.

Προληπτικά Μέτρα:

- a) **Διακοπή καπνίσματος:** Ο καλύτερος τρόπος για να μειωθεί ο κίνδυνος καρκίνου του πνεύμονα.
- b) **Μείωση της έκθεσης σε επιβλαβείς ουσίες:** Η χρήση προστατευτικού εξοπλισμού αποκτά πρωταρχική σημασία όταν χρησιμοποιείται σε βιομηχανίες με πιθανή έκθεση σε καρκινογόνες ουσίες όπως αυτές που αναφέρθηκαν.
- c) **Μείωση της ατμοσφαιρικής ρύπανσης:** Βρίσκεται πέρα από τον άμεσο έλεγχο ενός ατόμου, ωστόσο η υποστήριξη πολιτικών μαζί με ευρύτερες πρωτοβουλίες που επικεντρώνονται στη μείωση της ατμοσφαιρικής ρύπανσης μπορούν να αποφέρουν οφέλη.

2. Καρκίνος του μαστού

Ο καρκίνος του μαστού αναδείχθηκε ως ο πιο συχνά διαγνωσμένος καρκίνος παγκοσμίως για το 2020 και εντοπίστηκαν 2,26 εκατομμύρια νέες περιπτώσεις, που αποτελούν το 11,7% όλων των νέων περιπτώσεων καρκίνου εκείνο το έτος. Κατατάχθηκε επίσης μεταξύ των πιο θανατηφόρων μορφών, υπεύθυνος για πάνω από 685.000 χαμένες ζωές. Αυτός ο ανατριχιαστικός αριθμός θανάτων αντιπροσώπευε το 6,9% της συνολικής θνησιμότητας εκείνο το έτος. Αυτός ο ανησυχητικός συνδυασμός υψηλών ποσοστών θανάτου και συχνής διάγνωσης έχει καταστροφικές συνέπειες, ανατρέποντας αμέτρητες ζωές ασθενών.

Κύρια Καρκινογόνα Ουσία: Κατανάλωση αλκοόλ.

Προληπτικά Μέτρα: Αποφυγή ή μείωση της κατανάλωσης αλκοόλ.

3. Καρκίνος του δέρματος (μελάνωμα)

Η εκτίμηση της συχνότητας του καρκίνου του δέρματος αποτελεί σημαντική πρόκληση λόγω των πολυάριθμων υποτύπων καρκίνου του δέρματος. Για παράδειγμα, ο καρκίνος του δέρματος μη μελανώματος συχνά δεν

καταγράφεται, καθώς τα περισσότερα περιστατικά αντιμετωπίζονται επιτυχώς με χειρουργική επέμβαση ή αφαίρεση. Λόγω αυτών των παραγόντων, είναι πιθανό ότι η αναφερόμενη παγκόσμια συχνότητα εμφάνισης καρκίνου του δέρματος είναι υποεκτίμηση. Το 2020, υπήρχαν περίπου 1,2 εκατομμύρια νέες περιπτώσεις μη μελανωματικού καρκίνου του δέρματος και περίπου 325.000 νέες περιπτώσεις μελανώματος παγκοσμίως. Οι θάνατοι που σχετίζονται με το μελάνωμα εκείνη τη χρονιά ήταν περίπου 57.000, ενώ ο μη μελανωματικός καρκίνος του δέρματος αντιπροσώπευε 63.731 θανάτους.

Κύρια Καρκινογόνα Ουσία: Υπεριώδης ακτινοβολία.

Προληπτικά Μέτρα: Αποφυγή ή μείωση της έκθεσης στην υπεριώδη ακτινοβολία μέσω αντηλιακής προστασίας, όπως η χρήση αντηλιακού ή η χρήση προστατευτικού ρουχισμού.

A. Τρέχουσες Λύσεις

Τεχνητή Νοημοσύνη

Η τεχνητή νοημοσύνη αναδεικνύεται ως ένας ισχυρός διαγνωστικός σύμμαχος, ο οποίος περιλαμβάνει ένα σύνολο αλγορίθμων μηχανικής μάθησης που έχουν σχεδιαστεί για την προσομοίωση της ανθρώπινης νοημοσύνης στην εξέταση, παρουσίαση και κατανόηση των σύνθετων δεδομένων υγείας, όπως ηλεκτρονικά αρχεία υγείας και ιατρικές εικόνες. Πιο συγκεκριμένα, η τεχνητή νοημοσύνη αναλύει αυτά τα δεδομένα εισόδου με πρωταρχικό στόχο να εξετάσει τις συσχετίσεις μεταξύ των αποτελεσμάτων των ασθενών και των κλινικών διαδικασιών, που αντιμετωπίζουν την κατάσταση της υγείας του ασθενούς και ιδανικά οδηγούν σε θετικό αποτέλεσμα. Μέσω αυτής της συσχέτισης, γίνεται επίγνωση του πόσο αποτελεσματικές είναι ορισμένες κλινικές διαδικασίες για την επίτευξη των επιθυμητών αποτελεσμάτων του ασθενούς. Για παράδειγμα, εάν ένα συγκεκριμένο σχέδιο θεραπείας (κλινική διαδικασία) οδηγεί σταθερά σε υψηλά ποσοστά ανάκαμψης του ασθενούς, υπάρχει θετική συσχέτιση μεταξύ αυτής της διαδικασίας και του αποτελέσματος.

Παράλληλα, η τεχνητή νοημοσύνη χρησιμοποιεί προκαθορισμένες οδηγίες για να κάνει προβλέψεις σχετικά με διάφορες πτυχές του καρκίνου. Σε αντίθεση με τις παραδοσιακές μεθόδους στατιστικής ανάλυσης, έχει τη δυνατότητα να συλλέγει δεδομένα, να τα επεξεργάζεται, να τα ερμηνεύει και να παρέχει συγκεκριμένα αποτελέσματα, καθώς ερμηνεύει καλύτερα σύνθετες, μη γραμμικές σχέσεις μεταξύ διαφορετικών μεταβλητών για να συλλάβει την πολυπλοκότητα του πραγματικού κόσμου. Αντί να αντικαταστήσει την ανθρώπινη τεχνογνωσία, η τεχνητή νοημοσύνη στοχεύει να αυξήσει τις ικανότητες των κλινικών ιατρών παρέχοντας πρόσθετη αναλυτική δύναμη που ενσωματώνει πολλαπλά σημεία δεδομένων σε ένα γενικό προφίλ ασθενών, ώστε να γίνει πιο συνειδητή η λήψη αποφάσεων.

Μηχανική Μάθηση

Σε μια εποχή όπου τα δεδομένα κατέχουν το κλειδί, οι ιατρικές επιστήμες επωφελούνται πάρα πολύ από τη δύναμη της μηχανικής μάθησης. Αυτός ο καινοτόμος τομέας, που βρίσκεται στην καρδιά της τεχνητής νοημοσύνης, εστιάζει στην ανάπτυξη αλγορίθμων με ένα άνευ προηγουμένου επίπεδο αυτοματισμού, ενδυναμώνοντας τα ευφυή συστήματα να αντιμετωπίζουν πολύπλοκα προβλήματα με ελάχιστη έως καθόλου ανθρώπινη καθοδήγηση. Μαθαίνοντας από προηγούμενες ανόμοιες ροές δεδομένων και παρατηρήσεων, με την πάροδο του χρόνου, αυτοί οι αλγόριθμοι κάνουν προβλέψεις με βάση τη συσσωρευμένη εμπειρία τους, μοντελοποιώντας τα δυναμικά φαινόμενα του πραγματικού κόσμου και αποκαλύπτοντας μοτίβα, που είναι αόρατα με γυμνό μάτι.

Βαθιά Μάθηση

Η βαθιά μάθηση, ένα υποσύνολο της μηχανικής μάθησης, μεταμορφώνει την ταξινόμηση των κακοηθών καρκίνων, χρησιμοποιώντας μοντέλα που βασίζονται σε νευρωνικά δίκτυα, για να μιμηθεί την ικανότητα του ανθρώπινου εγκεφάλου να αναλύει και να εξάγει πληροφορίες από τεράστια, πολύπλοκα σύνολα δεδομένων.

Τι είναι τα νευρωνικά δίκτυα ;

Τα νευρωνικά δίκτυα, τα οποία είναι εμπνευσμένα από τα βιολογικά νευρωνικά δίκτυα στον ανθρώπινο εγκέφαλο, αποτελούνται από διασυνδεδεμένες μονάδες επεξεργασίας πληροφοριών που ονομάζονται τεχνητοί νευρώνες. Αυτές οι μονάδες λειτουργούν σαν ομάδα, ανταλλάσσοντας πληροφορίες παρόμοιες με το πώς επικοινωνούν οι νευρώνες μέσα στον εγκέφαλο. Ενώ μιμούνται τη δομή του εγκεφάλου, οι τεχνητοί νευρώνες είναι πολύ πιο απλοί, δεν πλησιάζουν καθόλου την πολυπλοκότητα των νευρικών κυττάρων και δεν έχουν την ικανότητα να σκέφτονται σαν άνθρωποι, δηλαδή να αναλύουν στιγμιαία χιλιάδες λεπτομέρειες για να αποκωδικοποιήσουν νέα αντικείμενα και να τα τοποθετήσουν στο ίδιο πλαίσιο αναφοράς.

Για να λειτουργήσουν αποτελεσματικά τα νευρωνικά δίκτυα, πρέπει να υποβληθούν σε μια διαδικασία εκπαίδευσης. Αυτή η διαδικασία εκμάθησης περιλαμβάνει την έκθεση του δικτύου σε ένα προσεκτικά επιλεγμένο σύνολο δεδομένων εισόδου, προσαρμόζοντας σταδιακά τα βάρη και τις παραμέτρους των κρυφών επιπέδων του, ελαχιστοποιώντας την συνάρτηση κόστους, δηλαδή την διαφορά μεταξύ των εξόδων του δικτύου και των επιθυμητών αποτελεσμάτων, ώστε το νευρωνικό δίκτυο να είναι αποδοτικό στον πραγματικό κόσμο.

Μια βασική διάκριση μεταξύ νευρωνικών δικτύων και κλασικών αλγορίθμων έγκειται στη θεμελιώδη προσέγγισή τους, ότι οι κλασικοί αλγόριθμοι απαιτούν σχολαστική σχεδίαση κάθε βήματος, ενώ τα νευρωνικά δίκτυα διαθέτουν την αξιοσημείωτη ικανότητα να μαθαίνουν και να εκτελούν εργασίες σωστά από

μόνα τους. Οι περίπλοκες υπολογιστικές διαδικασίες εξαγωγής συμπερασμάτων μέσα σε πολύπλοκα νευρωνικά δίκτυα παραμένουν κάπως αδιαφανείς, ακόμη και για τους δημιουργούς τους, γι' αυτό συνήθως γίνεται αξιοποίηση προεκπαιδευμένων νευρωνικών δίκτυων που είναι προετοιμασμένα για συγκεκριμένους τύπους δεδομένων. Αυτή η προσέγγιση είναι παρόμοια με την εκπαίδευση των ατόμων σε συγκεκριμένες δραστηριότητες χωρίς την πλήρη κατανόηση των υποκείμενων νευρολογικών διεργασιών, πιστεύοντας ότι η εκπαίδευση διαδικασία θα είναι αποτελεσματική εάν εκτελεστεί σωστά.

ANN

Μια λεπτομερή περιγραφή των δομών των νευρωνικών δίκτυων είναι τα τεχνητά νευρωνικά δίκτυα (ANN), τα οποία δημιουργούν προβλέψεις σε πραγματικό χρόνο. Αυτά τα υπολογιστικά μοντέλα μπορούν να αξιοποιήσουν διάφορες κλινικές παραμέτρους, όπως την ηλικία, τις συνήθειες κατανάλωσης καπνίσματος και αλκοόλ και την εθνικότητα ενός ατόμου, για να κάνουν τεκμηριωμένες προβλέψεις σχετικά με την ευαισθησία του στον καρκίνο. Ένα ANN είναι δομημένο σε επίπεδα: το επίπεδο εισόδου λαμβάνει τα αρχικά τροφοδοτούμενα δεδομένα, το κρυφό στρώμα επεξεργάζεται αυτά τα δεδομένα μέσω κόμβων που επαναληπτικά προσαρμόζουν τα βάρη και τις τιμές μεροληψίας για να βελτιώσουν σταδιακά την ακρίβεια των προβλεψεων και το επίπεδο εξόδου, που λαμβάνει το σταθμισμένο άθροισμα, εφαρμόζει μια συνάρτηση ενεργοποίησης, ώστε να μετατρέψει την έξοδο σε μια τιμή πιθανότητας που υποδεικνύει την ευαισθησία στον καρκίνο. Τα μοντέλα βαθιάς μάθησης είναι μια πιο εξελιγμένη μορφή ANN, που διαθέτουν πολλαπλά κρυφά επίπεδα, επιτρέποντάς τους να ανιχνεύουν και να επεξεργάζονται πολύπλοκα μοτίβα μέσα στα δεδομένα, ενισχύοντας έτσι την ακρίβεια πρόβλεψης.

CNN

Τα συνελικτικά νευρωνικά δίκτυα (CNN) είναι μια εξειδικευμένη μορφή αλγορίθμων βαθιάς μάθησης, που είναι ειδικά προσαρμοσμένα για την επεξεργασία δεδομένων που μοιάζουν με πλέγμα, όπως εικόνες και βίντεο, που απαιτούν αναγνώριση αντικειμένων, όπως ταξινόμηση, ανίχνευση και τμηματοποίηση εικόνων. Τα CNN είναι εμπνευσμένα από την περίπλοκη πολυεπίπεδη αρχιτεκτονική του ανθρώπινου οπτικού φλοιού, χρησιμοποιούν μια ιεραρχική δομή χτίζοντας σταδιακά σύνθετες αναπαραστάσεις χαρακτηριστικών πάνω από απλούστερες που εξάγονται σε προηγούμενα στρώματα, για την δημιουργία χαρτών ανεξάρτητων από τη θέση και τον προσανατολισμό χαρακτηριστικών. Παράλληλα αξιοποιούν την τοπική συνδεσιμότητα, δηλαδή κάθε νευρώνας συνδέεται μόνο σε μια μικρή, τοπική περιοχή της εικόνας και μέσω της συνέλιξης, που λειτουργεί ως συρόμενο παράθυρο, γίνεται η σταδιακή ανάλυση της εικόνας, ώστε να αποφευχθεί η υπολογιστική πολυπλοκότητα ολόκληρης της εικόνας.

Β. Πρακτικές

ANN - Καρκίνος Μαστού

Μια πρόσφατη μελέτη προσπάθησε να εντοπίσει το πιο ισχυρό μοντέλο για την πρόβλεψη του καρκίνου του μαστού συγκρίνοντας την ακρίβεια διαφορετικών τεχνικών μηχανικής μάθησης, συμπεριλαμβανομένων Support Vector Machines, τεχνητών νευρωνικών δίκτυων (ANN), και των ταξινομητών Naive Bayes. Για να βελτιωθεί η ανάλυση, χρησιμοποιήθηκε η ανάλυση κύριου συστατικού (PCA) για τη μείωση της διάστασης των δεδομένων, διασφαλίζοντας ότι τα πιο σημαντικά χαρακτηριστικά χρησιμοποιήθηκαν χωρίς απώλεια βασικών πληροφοριών δεδομένων. Τα αποτελέσματα έδειξαν ότι τα ANN ήταν τα πιο αποτελεσματικά, ξεπερνώντας τους παραδοσιακούς αλγόριθμους σε αξιοπιστία για προβλέψεις σε πραγματικό χρόνο. Μαζί με την ικανότητα τους να μαθαίνουν αυτόματα και να εξάγουν χαρακτηριστικά από μεγάλα σύνολα δεδομένων χωρίς εξωτερική χειροκίνητη παρέμβαση τα καθιστά πολύτιμο εργαλείο για τη βελτίωση της διάγνωσης του καρκίνου του μαστού.

CNN - Καρκίνος του Πνεύμονα

Το συνελικτικό νευρωνικό δίκτυο πρόβλεψης καρκίνου του πνεύμονα (LCP-CNN) της Optellum αποτελεί έναν κλινικό σύμμαχο στην έγκαιρη ανίχνευση του καρκίνου του πνεύμονα. Αυτή η πρωτοποριακή τεχνολογία αντιπροσωπεύει το πρώτο στον κόσμο λογισμικό σύστημα υποστήριξης αποφάσεων εικονικών οζίδιων με τεχνητή νοημοσύνη για έγκαιρη ανίχνευση καρκίνου του πνεύμονα, που βοηθά τους γιατρούς στη διάγνωση και την παρακολούθηση ασθενών σε κίνδυνο με την κλινικά επικυρωμένη βαθμολογία Πρόβλεψης Καρκίνου του Πνεύμονα (LCP). Ανεξάρτητες μελέτες έχουν επικυρώσει την αξιοσημείωτη ικανότητα του λογισμικού να διαγνώσει κακοήθεις ακαθόριστων πνευμονικών όζων (IPNs), ξεπερνώντας την ευρέως χρησιμοποιούμενη βαθμολογία κινδύνου Brock με AUC 89,6% για την πρόβλεψη κακοήθειας και μέσω αναδρομικών αξιολογήσεων εντοπίζει κακοήθειες, μειώνοντας σημαντικά τη συχνότητα εμφάνισης ψευδών αρνητικών. Καθώς τα ψηφιακά εργαλεία όπως το LCP-CNN συνεχίζουν να εξελίσσονται, η ενσωμάτωσή τους στην κλινική πρακτική ρουτίνας αναμένεται να γίνει ευρέως διαδεδομένη, αν και η συνεχής επικύρωση σε διαφορετικούς πληθυσμούς ασθενών είναι απαραίτητη.

Γ. State-of-the-art

Η πρόβλεψη της εξέλιξης του καρκίνου είναι δύσκολη λόγω της πολύπλοκης φύσης της στατιστικής ανάλυσης και των πολυάριθμων πιθανολογικών συμβάντων που εμπλέκονται, όπως η υποτροπή του όγκου ή η θνησιμότητα. Τα δίκτυα Bayes προσφέρουν μια ισχυρή λύση με το χειρισμό πιθανολογικών γεγονότων και αβεβαιότητας. Αυτά τα γραφικά μοντέλα αντιπροσωπεύουν τυχαίες μεταβλητές και τις υπό συνθήκη εξαρτήσεις τους, χρησιμοποιώντας κοινή πιθανότητα για τον προσδιορισμό της πιθανότητας ενός χαρακτηριστικού

δεδομένου ενός άλλου, που συχνά εκφράζεται ως P(χαρακτηριστικό/γονικό). Η δομή του οπτικοποιείται χρησιμοποιώντας ένα κατευθυνόμενο ακυκλικό γράφημα (DAG), όπου οι κόμβοι αντιπροσωπεύουν μεταβλητές/γεγονότα με τις κατανομές πιθανοτήτων τους και οι ακμές απεικονίζουν τις υπό όρους εξαρτήσεις μεταξύ των μεταβλητών, με την αριστερή μεταβλητή να θεωρείται ως θυγατρική και τη δεξιά μεταβλητή ως γονέας.

Ένα δίκτυο Bayes είναι ένα οπτικό εργαλείο που αναλύει μια σύνθετη κατανομή πιθανοτήτων σε απλούστερα μέρη, καθιστώντας την ιδιαίτερα χρήσιμη για διάγνωση και λήψη αποφάσεων σε συνθήκες αβεβαιότητας. Οπτικοποιεί τις σχέσεις μεταξύ μεταβλητών, με κατανομές πιθανοτήτων υπό όρους (CPDs) που χαρακτηρίζουν την αρχική κατανομή πιθανότητας, λειτουργούν ως παράμετροι του δίκτυου και σχηματίζουν ένα Bayesian δίκτυο που καλύπτει όλες τις πιθανές περιπτώσεις για τη κάθε μεταβλητή. Καθώς γίνονται διαθέσιμα νέα στοιχεία ή δεδομένα, η εκτίμηση του Bayes ενημερώνει την πιθανότητα μιας υπόθεσης, κινούμενη από προγενέστερες σε μεταγενέστερες πεποιθήσεις. Αυτή η προσέγγιση ενσωματώνει νέες πληροφορίες για να βελτιωθεί η κατανόησή για την πιθανότητα ενός γεγονότος, καθοδηγούμενη από το θεώρημα του Bayes. Στη στατιστική ανάλυση, η τυχαιότητα στα δεδομένα διαχειρίζεται υποθέτοντας ότι ακολουθεί μια προβλέψιμη μορφή που περιγράφεται από μια συνάρτηση κατανομής πιθανότητας (PDF), που περιγράφει τις ιδιότητες των πιθανών τιμών, μέσω μιας παραμετρικής οικογένειας, εστιάζοντας στην εκτίμηση βασικών παραμέτρων (τιμών) για έναν πληθυσμό, μέσω της συλλογής δεδομένων. Οι στατιστικές Bayes επεκτείνουν το θεώρημα του Bayes για εξαγωγή συμπερασμάτων και προβλέψεων από δεδομένα μοντελοποιώντας την τυχαιότητα δεδομένων με μια PDF για την περιγραφή της πιθανότητας διαφόρων αποτελεσμάτων, εκτιμώντας στατιστικές ιδιότητες όπως ο μέσος όρος και το εύρος.

Οι παραδοσιακές στατιστικές μέθοδοι, όπως η εκτίμηση μέγιστης πιθανότητας (MLE), υποθέτουν σταθερές, άγνωστες παραμέτρους πληθυσμού, παρέχοντας εκτιμήσεις θεωρητικά κοντά στις πραγματικές τιμές με επαρκή δεδομένα και ακολουθούμενες από έλεγχο υποθέσεων αξιολογώντας την πιθανότητα τα δεδομένα να προέρχονται από μια συγκεκριμένη υπόθεση ή τιμή παραμέτρου (p-value). Ωστόσο, αυτές οι μέθοδοι αγνοούν την προηγούμενη γνώση σχετικά με τις παραμέτρους, εστιάζοντας μόνο στην κοινή κατανομή πιθανοτήτων των δεδομένων. Αντίθετα, οι στατιστικές Bayes αντιμετωπίζουν τις παραμέτρους ως τυχαίες μεταβλητές, αντικαθιστώντας μια μεμονωμένη πιθανότητα με μια ολόκληρη κατανομή πιθανότητας, γνωστή ως προηγούμενη κατανομή, για να ενσωματώσουν πεποιθήσεις, σχετικά με το ποιες τιμές παραμέτρων είναι πιθανές ή απίθανες πριν από τη συλλογή δεδομένων. Το θεώρημα του Bayes ενημερώνει αυτές τις πεποιθήσεις με νέα δεδομένα για να σχηματίσει μια μεταγενέστερη κατανομή, που αντικατοπτρίζει πιθανότητες διαφορετικών τιμών παραμέτρων μετά την εξέταση των νέων στοιχείων. Αυτή η ενοποίηση της προηγούμενης γνώσης με τις συναρτήσεις πιθανότητας για τον σχηματισμό

μεταγενέστερων κατανομών επιτρέπει μια λεπτή εξέταση των πιθανοτήτων πριν και μετά τη συλλογή δεδομένων, ενισχύοντας τη σαφήνεια και την ακρίβεια των προβλέψεων και των συμπερασμάτων, όπως φαίνεται σε μοντέλα όπως το Beta-binomial που χρησιμοποιήσε η Pfizer για την αποτελεσματικότητα του εμβολίου COVID.

3. Προτεινόμενες τεχνικές

A. Θεωρητική περιγραφή

I. Παρουσίαση

Ψηφιακή Ιατρική Απεικόνιση

Ο τομέας της ιατρικής απεικόνισης έχει μεταβεί από τα φυσικά φίλμ ακτίνων X σε ψηφιακές σαρώσεις που είναι αποθηκευμένες στα Συστήματα Αρχειοθέτησης και Επικοινωνίας ασθενών (PACS). Αυτή η μετάβαση έχει φέρει επανάσταση στην αποθήκευση, ανάκτηση και ανάλυση τεράστιων ποσοτήτων δεδομένων ιατρικής απεικόνισης. Η αξιοποίηση της τεχνητής νοημοσύνης, ιδιαίτερα της όρασης υπολογιστών και της ραδιενέργειας, επιτρέπει την εξαγωγή πολύτιμων πληροφοριών από αυτά τα ψηφιακά οπτικά δεδομένα. Αν και ξεχωριστοί, αυτοί οι δύο κλάδοι λειτουργούν παράλληλα για να μεγιστοποιήσουν τη μάθηση από την ιατρική απεικόνιση.

Computer Vision

Το Computer vision αναπτύσσει αλγόριθμους που επιτρέπουν στους υπολογιστές να παρατηρούν και να κατανοούν οπτικές πληροφορίες, όπως εικόνες και βίντεο. Στον ιατρικό τομέα, η υπολογιστική όραση ξεκινά με τη λήψη μιας εικόνας (π.χ. ακτινογραφίας, αξονικής τομογραφίας) με μια συσκευή ανίχνευσης. Στη συνέχεια, η εικόνα αποστέλλεται σε άλλο σύστημα εξοπλισμένο με μια τεράστια συλλογή σχετικών οπτικών δεδομένων (π.χ. PACS). Αυτό το σύστημα χρησιμοποιεί αλγόριθμους αναγνώρισης μοτίβων, όπως μοντέλα βαθιάς εκμάθησης CNN, για να αναλύσει την εικόνα, να συγκρίνει τα μοτίβα με τη βιβλιοθήκη γνωστών μοτίβων της και να προσδιορίσει εάν οποιοδήποτε περιεχόμενο στην εικόνα ταιριάζει με αυτά τα μοτίβα. Στο τέλος με βάση τα αναγνωρισμένα μοτίβα, το σύστημα παρέχει διαγνωστικές πληροφορίες.

Radiomics

Από την άλλη πλευρά, το Radiomics αξιοποιεί τεχνικές όρασης υπολογιστή, αλλά υπερβαίνει την οπτική ερμηνεία ποσοτικοποιώντας τα χαρακτηριστικά απεικόνισης για υποστήριξη αποφάσεων βάσει δεδομένων. Αυτή η τεχνολογία αναλύει μοτίβα και υφές μέσα στις εικόνες για να εξαγάγει πολυάριθμα ποσοτικά χαρακτηριστικά που χρησιμοποιούνται για εργασίες όπως ο χαρακτηρισμός του όγκου και η πρόβλεψη της ανταπόκρισης στη θεραπεία. Παραδοσιακά, αυτό απαιτούσε χειροκίνητη κατάτμηση όγκων, εξαγωγή χαρακτηριστικών και ανάλυση. Ωστόσο, η βαθιά μάθηση έχει μεταμορφώσει τη διαδικασία μαθαίνοντας αυτόματα χαρακτηριστικά από δεδομένα, εξαλείφοντας την ανάγκη για μη αυτόματο ορισμό χαρακτηριστικών. Τα χαρακτηριστικά υφής όπως το μέγεθος, το σχήμα, η ένταση και η ετερογένεια των αντικειμένων στις

εικόνες εξάγονται αυτόματα από τις επισημασμένες περιοχές ενδιαφέροντος (ROI). Αυτά τα χαρακτηριστικά χρησιμοποιούνται στη συνέχεια για την εκπαίδευση μοντέλων για ταξινόμηση και πρόβλεψη, ενισχύοντας σημαντικά την ακρίβεια και τη χρησιμότητα.

Παρά τις προόδους αυτές, εξακολουθούν να υπάρχουν πολλές προκλήσεις σε αυτόν τον συνδυασμό (Radiomics – DL). Ένα σημαντικό ζήτημα είναι ο τεράστιος όγκος δεδομένων που απαιτείται από τα μοντέλα βαθιάς μάθησης, η συλλογή και η επιμέλεια των οποίων είναι χρονοβόρα και δαπανηρή. Επιπλέον, η διασφάλιση της συνέπειας στις παραμέτρους λήψης και ανακατασκευής εικόνας σε διαφορετικά ιδρύματα και παραλλαγές υλικού είναι μια άλλη σημαντική πρόκληση. Επιπρόσθετα, τα παραδοσιακά χαρακτηριστικά radiomics έχουν σαφείς μαθηματικές διατυπώσεις, ενώ τα χαρακτηριστικά βαθιάς μάθησης είναι συχνά ασαφή και δύσκολο να εξηγηθούν. Αυτή η έλλειψη διαφάνειας εγείρει ανησυχίες σχετικά με την κλινική χρησιμότητα και την αξιοπιστία τους.

Ψηφιακά Αρχεία Υγείας

Η αυξανόμενη αποθήκευση των αρχείων υγείας σε ηλεκτρονική μορφή έχει επεκτείνει τις υποδομές HMY, διευκολύνοντας την ταχεία αποθήκευση και ανάκτηση κλινικών δεδομένων. Αυτή η πρόοδος υποστηρίζει πρωτοβουλίες όπως η Digital Cancer Waiting Times Database, βελτιώνοντας τις διαδικασίες παραπομπής ασθενών για θεραπεία καρκίνου. Τα δεδομένα υγείας προέρχονται τοπικά από νοσοκομεία ή σε εθνικό επίπεδο από μητρώα δημόσιας υγείας, χρησιμοποιώντας ενοποιημένες βάσεις δεδομένων για τη διατήρηση της συνέπειας και τη διασφάλιση της διαλειτουργικότητας.

Natural Language Processing (NLP)

Η τεχνητή νοημοσύνη ενισχύει τον έλεγχο κινδύνου ασυμπτωματικών ασθενών επεξεργάζοντας μεγάλες ποσότητες δομημένων (π.χ. ημερομηνίες εισαγωγής, αποτελέσματα εξετάσεων αίματος) και μη δομημένων δεδομένων (π.χ. σημειώσεις ελεύθερου κειμένου, διαγνωστικές αναφορές) που βρίσκονται στα HMY. Το NLP μετατρέπει τα μη δομημένα κλινικά δεδομένα σε μορφές που μπορούν να αναλυθούν από τα υπολογιστικά συστήματα, αυτοματοποιώντας εργασίες που απαιτούν πολλούς πόρους και βελτιώνοντας την αποτελεσματικότητα της επεξεργασίας δεδομένων. Ανιχνεύει τα μη φυσιολογικά αποτελέσματα προσυμπτωματικού ελέγχου καρκίνου, αναλύει διαγνωστικά πρότυπα και αξιολογεί τις βλάβες ως προς τη διαστρωμάτωση του κινδύνου καρκίνου. Επιπλέον, το NLP προσδιορίζει ασθενείς κατάλληλους για κλινικές δοκιμές και βοηθά στην ανάπτυξη μοντέλων πρόβλεψης καρκίνου σε πραγματικό χρόνο, βελτιώνοντας έτσι τη διάγνωση και τη θεραπεία του καρκίνου.

II. Περιγραφή

Δεδομένης της πολυπλοκότητας της βιολογίας του όγκου, τα μοντέλα που βασίζονται σε μεμονωμένες πηγές δεδομένων ενδέχεται να παραβλέπουν κρίσιμες προγνωστικές πληροφορίες που προκύπτουν από την περίπλοκη αλληλεπίδραση μεταξύ αλληλοεξαρτώμενων βιολογικών συστημάτων, αποτυγχάνοντας έτσι να συλλάβουν την πλήρη εικόνα. Για την αντιμετώπιση αυτής της πρόκλησης, η προσέγγιση μου χρησιμοποιεί τεχνικές computer vision και radiomics σε συνδυασμό με την επεξεργασία φυσικής γλώσσας (NLP) για την ενσωμάτωση πολυτροπικών δεδομένων σε μια ενιαία βάση δεδομένων που στεγάζεται σε κάθε νοσοκομείο, παρέχοντας έτσι μια πιο ολοκληρωμένη περιγραφή του τοπίου του όγκου και ενισχύοντας τη διαγνωστική ακρίβεια. Περιλαμβάνει επίσης διαχείριση ουράς δεδομένων FIFO και λύσεις multi-cloud για την ανάπτυξη του Bayesian δικτύου για την εξαγωγή προβλέψεων που υποδεικνύουν την πιθανότητα ανίχνευσης μελανώματος. Αποτελείται από πέντε διακριτές ροές εργασίας δεδομένων, οι οποίες αναλύονται στην συνέχεια:

1. Αγωγός Ανάπτυξης CNN.
2. Αγωγός Ανάπτυξης Μοντέλου NLP.
3. Ροή Εργασίας Ενοποιημένων Τοπικών Βάσεων Δεδομένων.
4. Ροή Εργασίας Πολλαπλών Σύννεφων.
5. Επικοινωνία και Διαχείριση Δεδομένων.

B. Υλοποίηση-Πειράματα

I. Μελέτη Περίπτωσης

Η καινοτόμος προσέγγιση για την ενσωμάτωση πολυτροπικών πηγών δεδομένων δείχνει τις δυνατότητές της σε μια πρωτοποριακή μελέτη περίπτωσης για την έγκαιρη διάγνωση του καρκίνου του δέρματος:

Ο καρκίνος του δέρματος, ιδιαίτερα το μελάνωμα, είναι από τους πιο επιθετικούς και δύσκολους να ανιχνευθούν στα πρώτα του στάδια. Οι παραδοσιακές μέθοδοι που βασίζονται στην οπτική ανίχνευση και ανάλυση μεμονωμένων σημείων δεδομένων συχνά υστερούν στον ακριβή εντοπισμό των παραμικρών λεπτομερειών που μπορεί να υποδηλώνουν την παρουσία καρκίνου. Αξιοποιώντας τη δύναμη των τεχνικών computer vision για την ανάλυση της εικόνας, των radiomics για την εξαγωγή ποσοτικών δεδομένων από δερμοσκοπικές εικόνες και τον συνδυασμό τους με πληροφορίες από κλινικές σημειώσεις που εξάγονται μέσω του NLP, αυτή η προσέγγιση παρέχει μια πιο ολοκληρωμένη εικόνα των ύποπτων βλαβών.

Radiomics + Computer Vision

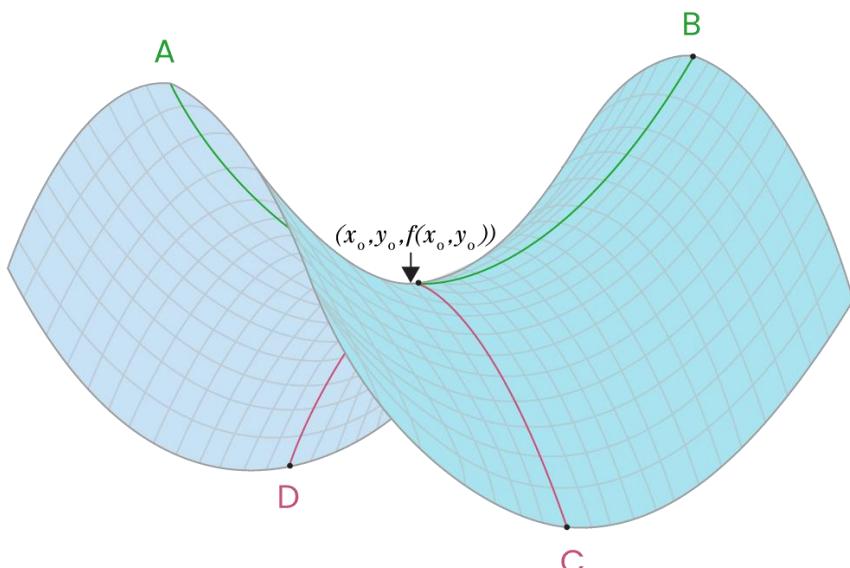
Χρησιμοποιώντας διάφορες τεχνικές υπολογιστικής όρασης, το περιεχόμενο της εικόνας αναλύεται εντοπίζοντας χαρακτηριστικά χαμηλού επιπέδου και καθορίζοντας βασικά σημεία και όρια πιθανών καρκινικών περιοχών:

Εντοπισμός Σαγματικού Σημείου

Το σαγματικό σημείο περιορίζει το χωρικό επίπεδο ενός καρκίνου.

Μαθηματικός Ορισμός

Έστω συνάρτηση f που είναι παραγωγίσιμη σε ένα κρίσιμο σημείο (a,b) . Τότε η f έχει ένα σαγματικό σημείο στο (a,b) , εάν σε κάθε ανοιχτό δίσκο με κέντρο το (a,b) υπάρχουν σημεία (x,y) , για τα οποία $f(x,y) > f(a,b)$ και $f(x,y) < f(a,b)$.



Σχήμα 1.1: Σαγματικό Σημείο

Ένα σαγματικό σημείο δεν είναι ούτε τοπικό μέγιστο ούτε τοπικό ελάχιστο σημείο, καθώς, αν σχεδιαστεί ένα ίχνος από το σημείο A στο σημείο B στην επιφάνεια, το κρίσιμο σημείο μας $(x_0, y_0, f(x_0, y_0))$ είναι ένα ελάχιστο. Ωστόσο, αν σχεδιαστεί ένα ίχνος από το σημείο C στο σημείο D το κρίσιμο σημείο είναι μέγιστο. Όμως ένα σημείο δεν μπορεί να είναι ταυτόχρονα μέγιστο και ελάχιστο σημείο (δηλαδή, minimax point). Για αυτό, περιγράφονται τέτοια σημεία ως σημεία σέλας, καθώς μοιάζουν με σέλα αλόγου.

Αυτό το σημείο καμπής έχει βαθιά σημασία στην ανίχνευση καρκίνου από ιατρικές σαρώσεις. Στο πλαίσιο των καρκινικών περιοχών, προκύπτουν δύο διαφορετικά μοτίβα έντασης: Οι αισιόδοξοι καρκίνοι που παρουσιάζουν μη φυσιολογική αύξηση της φωτεινότητας σε σύγκριση με τον υγιή ιστό, στο

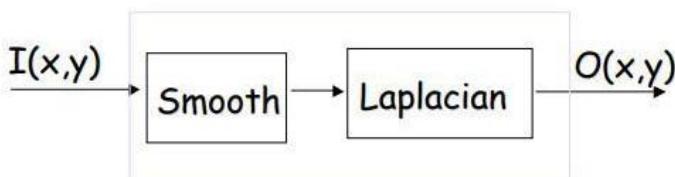
σημείο της σέλας η ένταση αυξάνεται ξαφνικά, ενώ οι αμυδροί καρκίνοι εκδηλώνονται ως πιο σκούρες, μειωμένης έντασης περιοχές, κάτω από το φυσιολογικό επίπεδο των ιστών. Έτσι, στο σημείο της σέλας, η ένταση μείωνεται ξαφνικά. Συνεπώς, τα σαγματικά σημεία οριοθετούν τα όρια αυτών των ανώμαλων περιοχών, σηματοδοτώντας τη συγκυρία όπου η ένταση μεταβαίνει απότομα από την κλιμάκωση στη μείωση ή το αντίστροφο. Εντοπίζοντας αυτά τα σημεία, οι ακτινολόγοι και οι ειδικοί στην απεικόνιση μπορούν να οριοθετήσουν με ακρίβεια την έκταση των πιθανών καρκινικών περιοχών.

Από την Θεωρία στην Πράξη

Για το προσδιορισμό του σαγματικού σημείου γίνεται χρήση δύο βασικών τεχνικών, της Laplacian of the Gaussian (LoG) και της Διαφορά των Gaussian (DoG).

LoG

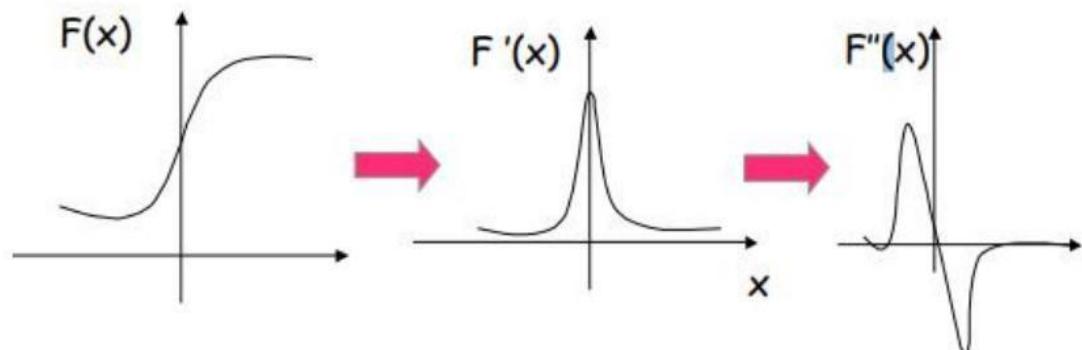
Στην παραδοσιακή ανίχνευση ακμών με χρήση της Laplacian γίνεται απευθείας συνέλιξη της εικόνας με το λαπλασιανό φίλτρο, που είναι το άθροισμα των παραγώγων δεύτερης τάξης ως προς το x και y , υποδεικνύοντας περιοχές όπου η ένταση μεταβάλλεται πιο γρήγορα. Ωστόσο, η λήψη παραγώγων τείνει να ενισχύει το θόρυβο, καθιστώντας την Laplacian ευαίσθητη στο θόρυβο, ως αποτέλεσμα την ανίχνευση ψευδών ακμών. Σε αντίθεση η LoG, πρώτα εξομαλύνει την εικόνα χρησιμοποιώντας ένα φίλτρο Gauss, για να μειώσει το θόρυβο και, στη συνέχεια, εφαρμόζει τον τελεστή Laplacian για να ανιχνεύσει περιοχές ταχείας αλλαγής έντασης φωτεινότητας (μηδενικές διασταυρώσεις), οι οποίες συχνά αντιστοιχούν σε ακμές ή χαρακτηριστικά που μοιάζουν με blobs, που είναι μια ομάδα συνδεδεμένων pixel σε μια εικόνα που μοιράζονται κάποια κοινή ιδιότητα (π.χ. τιμή κλίμακας του γκρι). Αυτό το εργαλείο είναι σαν ένα ειδικό ζευγάρι γυαλιά που κάνει μόνο τις ακμές των αντικειμένων να ξεχωρίζουν, όπως τα περιγράμματα ενός φύλλου ή το περίγραμμα ενός προσώπου.



Σχήμα 1.2: Laplacian of the Gaussian

$$O(x,y) = \nabla^2(I(x,y) * G(x,y))$$

Για την εύρεση των μηδενικών διασταυρώσεων, γίνεται εντοπισμός σημείων στην εικόνα όπου η ένταση αλλάζει γρήγορα (μεγάλη πρώτη παράγωγος), που αντιστοιχούν αντιστοιχούν σε κορυφές ή κοιλάδες και η αλλαγή στην ένταση είναι μέγιστη (μηδενική διασταύρωση στη δεύτερη παράγωγο). Αυτά τα σημεία είναι πιθανό να είναι ακμές στην εικόνα. Η χρήση δεύτερων παραγώγων επιτρέπει την ακριβή ανίχνευση των ακμών μέσα στην εικόνα. Αυτό οφείλεται στο ότι οι δεύτερες παράγωγοι ανταποκρίνονται έντονα στις αλλαγές στην ένταση. Η δεύτερη μερική παράγωγος I_{yy} , υποδεικνύει πώς η συνάρτηση αλλάζει ως y , ενώ το x διατηρείται σταθερό. Κατά τον υπολογισμό της παραγώγου σε σχέση με το y , επισημαίνει τις οριζόντιες ακμές, επειδή καταγράφει τις αλλαγές στις εντάσεις των εικονοστοιχείων κατά μήκος της κατακόρυφης διεύθυνσης. Αντίθετα, η δεύτερη μερική παράγωγος I_{xx} , υποδεικνύει πώς η συνάρτηση αλλάζει ως x , ενώ το y διατηρείται σταθερό. Συνεπώς, κατά τον υπολογισμό της παραγώγου σε σχέση με το x , δίνει έμφαση στις κάθετες ακμές, καθώς μετρά τις αλλαγές στις εντάσεις των εικονοστοιχείων κατά μήκος της οριζόντιας κατεύθυνσης.



Σχήμα 1.3: Εύρεση Μηδενικών Διασταυρώσεων

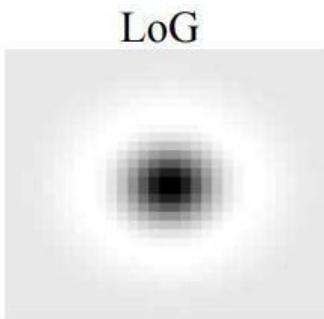
Το φίλτρο Laplacian of Gaussian (LoG) είναι ένα ισχυρό εργαλείο για την ανίχνευση περιοχών ενδιαφέροντος σε εικόνες. Αυτή η διαδικασία περιλαμβάνει δύο βασικά βήματα:

Αρχικά, ο αλγόριθμος εντοπισμού ακμών LoG συνελλίσεται με την εικόνα. Το LoG προσδιορίζει σημεία όπου η φωτεινότητα της εικόνας αλλάζει απότομα, τα οποία συχνά αντιστοιχούν στις ακμές ή στα όρια των αντικειμένων. Αυτό τονίζει τα περιγράμματα πιθανών blob ή περιοχών που ξεχωρίζουν λόγω της έντασής τους σε σύγκριση με τη γύρω περιοχή, μέσα στην εικόνα. Ένα blob μπορεί να οριστεί ως μια περιοχή συνδεδεμένων pixel που μοιράζονται μια κοινή ιδιότητα όπως η ένταση ή το χρώμα.

Στη συνέχεια, ένας αλγόριθμος ανίχνευσης blob επεξεργάζεται τα αποτελέσματα από το φίλτρο LoG. Οι ακμές που ανιχνεύονται συχνά σχηματίζουν κλειστά περιγράμματα ή όρια. Προσδιορίζοντας αυτά τα κλειστά περιγράμματα,

μπορούμε να αναγνωρίσουμε blobs που περικλείονται, που αποτελούν περιοχές με μέγιστη αλλαγή έντασης, που εμφανίζονται ως σκούρες κηλίδες σε ανοιχτό φόντο.

Αυτό μετατρέπει το αρχικό πρόβλημα της εύρεσης blobs στην εικόνα στην απλοποιημένη εργασία ανίχνευσης blobs εντός του χάρτη ακμών που δημιουργείται από το LoG. Ως εκ τούτου, η ανίχνευση ακμών με το LoG χρησιμεύει ως ένα χρήσιμο βήμα προεπεξεργασίας για ισχυρή ανίχνευση κηλίδων. Ουσιαστικά, λοιπόν, εντοπίοντας τις ακμές (όρια) των αντικειμένων σε μια εικόνα, μπορούν να αναγνωριστούν και τα blobs (αντικείμενα) στην εικόνα.



Σχήμα 1.4: Αναγνώριση Blob

Η συνέλιξη αποτελεί μια θεμελιώδης έννοια στην επεξεργασία εικόνας, περιλαμβάνει τη σύγκριση ενός μικρού φίλτρου με όλες τις τοπικές περιοχές της κύριας εικόνας για τον εντοπισμό χαρακτηριστικών. Αυτή η διαδικασία μπορεί να θεωρηθεί ως μετακίνηση του φίλτρου σε ολόκληρη την εικόνα, όπου σε κάθε σημείο, το φίλτρο συγκρίνεται με την τοπική περιοχή. Μια υψηλή τιμή υποδηλώνει ισχυρή αντιστοίχιση, υποδηλώνοντας ότι τα χαρακτηριστικά της τοπικής περιοχής είναι παρόμοια με το φίλτρο, όπως μια σκούρα κηλίδα σε ανοιχτόχρωμο φόντο ενώ μια χαμηλή τιμή υποδηλώνει ανομοιότητα.

DoG

Ο τελεστής Laplacian of Gaussian (LoG), που χρησιμοποιείται για την ανίχνευση ακμών, μπορεί να προσεγγιστεί με τη διαφορά των Gaussian (DoG), η οποία είναι υπολογιστικά λιγότερο ακριβής. Το DoG προκύπτει αφαιρώντας δύο εξομαλυμένες εικόνες Gaussian με διαφορετικές τυπικές αποκλίσεις, εκμεταλλευόμενη τη μαθηματική ιδιότητα ότι ο τελεστής LoG είναι ισοδύναμος με τη διαφορά δύο Gaussian σε διαφορετικές κλίμακες. Η προσέγγιση DoG εκτελεί φίλτραρισμα ζώνης, επιτρέπει τη διέλευση σημάτων εντός μιας συγκεκριμένης περιοχής συχνοτήτων και μπλοκάρει σήματα εκτός αυτής της περιοχής. Δεδομένου ότι ένα Gaussian είναι ένα χαμηλοπερατό φίλτρο (επιτρέπει σήματα χαμηλής συχνότητας και αποκλείει τον θόρυβο υψηλής συχνότητας), η διαφορά δύο Gaussian (DoG) θα έχει ως αποτέλεσμα ένα φίλτρο διέλευσης ζώνης μπλοκάροντας τόσο τον θόρυβο υψηλής συχνότητας όσο και

τις διακυμάνσεις χαμηλής συχνότητας, αφήνοντας μόνο τις συχνότητες που αντιστοιχούν στα επιθυμητά χαρακτηριστικά εικόνας, όπως ακμές ή blobs.

Gaussian Pyramid

Η πυραμίδα του Γκάους είναι μια τεχνική επεξεργασίας εικόνας που δημιουργεί μια σειρά από όλο και πιο θολές εκδόσεις μιας αρχικής εικόνας, γνωστές ως χονδροειδή στρώματα, εφαρμόζοντας επανειλημμένα ένα φίλτρο Gauss για εξομάλυνση και υποδειγματοληψία για μείωση μεγέθους. Κάθε επόμενο στρώμα στην πυραμίδα είναι μια πιο χονδροειδής αναπαράσταση του προηγούμενου στρώματος, που περιέχει λιγότερες λεπτομέρειες και μικρότερο μέγεθος. Αυτή η διαδικασία επιτρέπει την αποτελεσματική κωδικοποίηση εικόνας αξιοποιώντας την ικανότητα πρόβλεψης των χονδροειδών στρωμάτων για να προσεγγίσει την εμφάνιση του επόμενου λεπτότερου στρώματος.

Laplacian Pyramid

Η Πυραμίδα του Λαπλάς είναι ένας από τους αλγόριθμους που μπορούν να χρησιμοποιηθούν για την κωδικοποίηση της εικόνας, που είναι η διαδικασία μετατροπής μιας εικόνας σε μια μορφή που είναι πιο αποτελεσματική για αποθήκευση ή μετάδοση. Αντί να αποθηκεύει ολόκληρη την εικόνα λεπτής κλίμακας, αποθηκεύει μόνο τη διαφορά μεταξύ του προβλεπόμενου στρώματος, που λαμβάνεται από το χοντρό στρώμα, κατά την εξομάλυνση και την υποδειγματοληψία του προηγούμενου επιπέδου, και του πραγματικού λεπτότερου στρώματος, αποτυπώνοντας αποτελεσματικά τη χαμένη λεπτομέρεια κατά την εξομάλυνση και την υποδειγματοληψία. Αυτό σημαίνει ότι η Πυραμίδα Λαπλάς αποθηκεύει τις λεπτομέρειες που χάνονται κατά την εξομάλυνση και την υποδειγματοληψία, αντί για την ίδια την εικόνα. Αυτό είναι χρήσιμο γιατί μας επιτρέπει να ανακατασκευάσουμε την αρχική εικόνα από την Πυραμίδα Λαπλάς, ενώ εξοικονομούμε χώρο αποθήκευσης, καθιστώντας την μια πολύτιμη τεχνική για αποτελεσματική συμπίεση και αποθήκευση, χωρίς να θυσιάζει πολύ την ποιότητα της εικόνας.

Hessian-LoG

Χρησιμοποιείται για τον εντοπισμό ενδιαφερόντων σημείων σε μια εικόνα, τα οποία μπορούν να χρησιμοποιηθούν για την αντιστοίχιση διαφορετικών εικόνων. Υπολογίζεται ως το γινόμενο των διαγωνίων στοιχείων μείον το τετράγωνο του εκτός διαγώνιου στοιχείου.

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - f_{xy}^2$$

Σχήμα 1.5: Ορίζουσα Hessian

Ο ανιχνευτής Hessian-LoG είναι ένα ισχυρό εργαλείο στην ανάλυση εικόνας, που συνδυάζει τον τελεστή Laplacian με τον πίνακα Hessian για την αποτελεσματική αναγνώριση και αντιστοίχιση χαρακτηριστικών. Αυτή η υβριδική προσέγγιση αξιοποιεί την ορίζουσα Hessian για να προσδιορίσει περιοχές υψηλής καμπυλότητας για χωρική επιλογή, ενώ η κανονικοποιημένη Laplacian καθορίζει την κατάλληλη κλίμακα για κάθε σημείο δεδομένων. Αυτή η διπλή ικανότητα διασφαλίζει ότι ο ανιχνευτής είναι αμετάβλητος σε μεταφράσεις, περιστροφές και ομοιόμορφη κλιμάκωση, εξασφαλίζοντας συνεπή ανίχνευση χαρακτηριστικών σε διάφορους μετασχηματισμούς. Άρα, αυτά που τον καθιστούν ανεκτίμητο στην ανίχνευση του σαγματικού σημείου, είναι η αυτόματη προσαρμογή του σε διαφορετικές κλίμακες και η στιβαρότητά του στις χωρικές αλλαγές.

Ανάλυση Οριοθετημένων Περιοχών

Η ανάλυση των οριοθετημένων περιοχών είναι ζωτικής σημασίας για την κατανόηση των υφών και των μοτίβων στην ιατρική απεικόνιση. Αρχικά, τα radiomics μπορούν να ταξινομήσουν όγκους ως καλοήθεις ή κακοήθεις με υψηλή ακρίβεια μετατρέποντας ιατρικές εικόνες σε μεγάλα σύνολα δεδομένων, τα οποία στη συνέχεια χρησιμοποιούνται για την πρόβλεψη της πιθανότητας καρκίνου. Αυτή η ικανότητα προσδιορισμού του εάν μια ασθένεια θα είναι αργά αναπτυσσόμενη ή επιθετική είναι κρίσιμη για την έγκαιρη διάγνωση και τον σχεδιασμό της θεραπείας. Ο στόχος είναι να εντοπιστεί ένα μοτίβο συγκεκριμένων ραδιενεργών ετικετών (radiolabels) που μπορούν να προσκολληθούν στα καρκινικά κύτταρα, βοηθώντας πιθανώς τους γιατρούς να προβλέψουν τον χρόνο επιβίωσης και την ανταπόκριση ενός καρκινοπαθούς στη θεραπεία, σηματοδοτώντας ένα σημαντικό βήμα προς τα μπροστά, οδηγώντας στην αποτελεσματικότερη θεραπεία του καρκίνου.

GLCM

To Gray-Level Co-currence Matrix (GLCM) είναι μια τεχνική που χρησιμοποιείται για την ανάλυση της υφής των ιατρικών εικόνων καθώς και ως βιοδείκτης ετερογένειας, παρέχοντας λεπτομερείς πληροφορίες σχετικά με το μικροπεριβάλλον του όγκου. Η υφή της εικόνας παρέχει πληροφορίες για τη

χωρική διάταξη των χρωμάτων/εντάσεων, βοηθώντας μας να κατανοήσουμε πώς αυτές οι εντάσεις κατανέμονται και οργανώνονται στην εικόνα.

Το GLCM υπολογίζεται με βάση δύο παραμέτρους: τη σχετική απόσταση (D) μεταξύ των ζευγών εικονοστοιχείων και τον σχετικό προσανατολισμό (Θήτα), που είναι η γωνία περιστροφής. Λειτουργεί επιλέγοντας ζεύγη pixel από περιοχές με έντονη υφή και λείες περιοχές και μετρώντας την ομοιότητά τους. Ένας χαμηλός συσχετισμός μεταξύ αυτών των ζευγών υποδηλώνει υψηλό επίπεδο υφής, ενώ ένας υψηλός συσχετισμός υποδηλώνει πιο ομαλή υφή.

Εξάγοντας πολλαπλά χαρακτηριστικά από το GLCM, όπως αντίθεση, συσχέτιση, ομοιογένεια και ενέργεια, σε διαφορετικούς προσανατολισμούς και αποστάσεις, λαμβάνεται ένα πιο ολοκληρωμένο σύνολο χαρακτηριστικών που περιγράφουν την υφή της περιοχής ενδιαφέροντος (ROI) στην εικόνα. Αυτά τα χαρακτηριστικά στη συνέχεια συνδυάζονται για να σχηματίσουν ένα διάνυσμα χαρακτηριστικών. Τέλος, τα πολυδιάστατα διανύσματα χαρακτηριστικών διευκολύνουν τη διαφοροποίηση και την ταξινόμηση των αντικειμένων μέσα στην εικόνα.

NLP

Η επικύρωση των ραδιενεργών χαρακτηριστικών είναι ζωτικής σημασίας για τη διασφάλιση της κλινικής τους σημασίας. Αυτό περιλαμβάνει τη συσχέτιση των εξαγόμενων χαρακτηριστικών με γνωστά κλινικά αποτελέσματα για την επαλήθευση της προγνωστικής τους ισχύος. Οι αλγόριθμοι Επεξεργασίας Φυσικής Γλώσσας (NLP) επεξεργάζονται μη δομημένες κλινικές αναφορές, εξάγοντας σχετικό ιστορικό ασθενών, παράγοντες κινδύνου και περιγραφές συμπτωμάτων. Αυτή η ενοποίηση δεδομένων κειμένου ενισχύει τη συνολική ανάλυση παρέχοντας ένα πληρέστερο προφίλ ασθενούς.

II. Πρακτική Εφαρμογή

Στόχος είναι η δημιουργία μιας βελτιωμένης ροής εργασίας, που χρησιμοποιεί εξειδικευμένο αλγόριθμο για την ανάλυση μεγάλων ποσοτήτων δεδομένων που κοινοποιούνται από τοποθεσίες σε όλο τον κόσμο μέσα από ένα κατανεμημένο σύστημα, διατηρώντας το απόρρητο και την ασφάλεια των ασθενών. Επιτρέπει στα νοσοκομεία να εντοπίζουν ομάδες ατόμων υψηλού κινδύνου για κλινικές δοκιμές, να προβλέπουν την ανταπόκριση των ασθενών στις θεραπείες και να βρίσκουν σχετικούς βιοδείκτες για την απόκτηση νέων γνώσεων σχετικά με την αντιμετώπιση της θανατηφόρας νόσου του καρκίνου του δέρματος.

CNN

Αρχικά, αναπτύσσεται το μοντέλο του συνελικτικού νευρωνικού δικτύου (CNN), που χρησιμοποιεί προεπεξεργασμένες εικόνες ως εισόδους. Αυτές οι εικόνες βελτιώνονται με ανιχνευμένες πληροφορίες υφής, ακμών, φωτεινότητας και

ομαδοποίησης που λαμβάνονται από αναλύσεις GLCM και τεχνικές LoG, DoG και Hessian-LoG. Αυτό το βήμα προεπεξεργασίας επιτρέπει στο CNN να πραγματοποιεί πιο αποτελεσματική αυτόνομη ανίχνευση και ταξινόμηση του μελανώματος. Κατά συνέπεια, αυτή η προσέγγιση οδηγεί σε CNN υψηλής ακρίβειας που αυτοματοποιεί την αξιολόγηση των εικόνων δερματικών αλλοιώσεων, μειώνοντας σημαντικά τη χειρωνακτική προσπάθεια που απαιτείται σήμερα για τη διάγνωση μελανώματος από τους δερματολόγους.

Μοντέλο NLP

Αναπτύσσεται ένα μοντέλο NLP, που στοχεύει στην αυτοματοποίηση της κατανόησης, εξαγωγής και ανάλυσης δεδομένων από κλινικές αναφορές. Αρχικά, συλλέγονται ιατρικές αναφορές και κείμενα από ηλεκτρονικά ιατρικά αρχεία (EHR) και κλινικές σημειώσεις. Στη συνέχεια, τα δεδομένα καθαρίζονται αφαιρώντας σημεία στίξης, ειδικούς χαρακτήρες και κανονικοποιώντας το κείμενο. Το κείμενο χωρίζεται σε διακριτικά, τα οποία μετατρέπονται σε διανύσματα χρησιμοποιώντας μοντέλα ενσωμάτωσης, επιτρέποντας στο μοντέλο να κατανοήσει καλύτερα το περιεχόμενο και το πλαίσιο των λέξεων μέσω αμφίδρομης επεξεργασίας κειμένου. Το μοντέλο είναι εκπαιδευμένο να κατανοεί ιατρικές αναφορές, να εξάγει οντότητες όπως ασθένειες, φάρμακα και αποτελέσματα εργαστηρίου, να κατηγοριοποιεί τις αναφορές σε ενότητες όπως διαγνώσεις, θεραπείες και ιστορικό ασθενών και να δημιουργεί συνοπτικές περιλήψεις αναφορών. Αυτή η ολοκληρωμένη προσέγγιση ενισχύει σημαντικά την αποτελεσματικότητα και την ακρίβεια της διαχείρισης των κλινικών δεδομένων.

Ενοποιημένες Τοπικές Βάσεις Δεδομένων

Κάθε νοσοκομείο ή ιατρικό κέντρο συλλέγει τα δικά του δεδομένα από διάφορες πηγές, συμπεριλαμβανομένων των κλινικών δεδομένων ασθενών και των ιατρικών απεικονίσεων, τη διατήρηση των δεδομένων από διαφορετικά συστήματα υγείας ξεχωριστά και τον χειρισμό αυξημένων απαιτήσεων υποδομής. Η μετάβαση από ξεχωριστά συστήματα σε μια κοινή υποδομή μειώνει το κόστος και απλοποιεί τις λειτουργίες, διατηρώντας παράλληλα τα δεδομένα ασφαλή. Τα δεδομένα αυτά αποθηκεύονται, φυσικά με την συγκατάθεση του ασθενούς, στις τοπικές ενοποιημένες βάσεις δεδομένων, που ενσωματώνουν δεδομένα που εξήχθησαν τόσο από το μοντέλο NLP, όσο και από το CNN, αποθηκεύοντάς όλες τις ιατρικές πληροφορίες ενός ατόμου κρυπτογραφημένες, καθιστώντας τες προσβάσιμες αποκλειστικά σε όσους κλινικούς γιατρούς του συγκεκριμένου νοσοκομείου, έχουν εξουσιοδότησή να εξετάσουν παρόμοιες περιπτώσεις. Το κύριο πλεονέκτημα της ενοποιημένης βάσης δεδομένων είναι ότι συγκεντρώνει όλα τα δεδομένα, εξαλείφοντας την ανάγκη ανάλυσης διαφορετικών συνόλων ξεχωριστά. Αυτά τα ενοποιημένα δεδομένα στη συνέχεια μετατρέπονται σε πολύτιμες πληροφορίες, βοηθώντας

στην καλύτερη λήψη αποφάσεων. Με αυτόν τον τρόπο, σύνθετα ερωτήματα δεδομένων μπορούν να εκτελεστούν χωρίς παρεμβολές στις καθημερινές λειτουργίες των κύριων ροών.

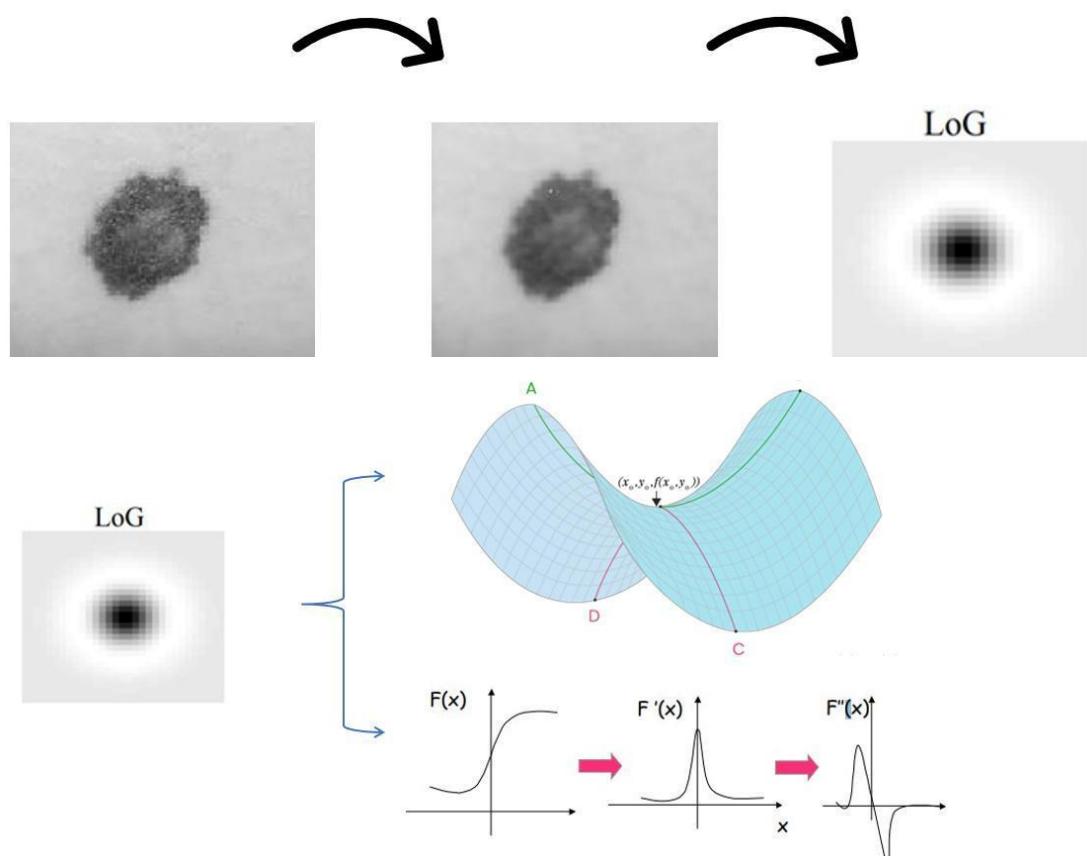
Τα νοσοκομεία χρησιμοποιούν την Ομοσπονδιακή Μάθηση (FL) για να εκπαιδεύσουν τοπικά το Bayesian δίκτυο με τα δεδομένα της ενοποιημένης βάσης τους, εξαλείφοντας την ανάγκη μεταφοράς δεδομένων μεταξύ ιδρυμάτων. Αυτό τους επιτρέπει να προβλέψουν εάν ένας ασθενής θα αναπτύξει καρκίνο, χωρίς συγκεντρωτικά δεδομένα. Αντί να μετακινούνται δεδομένα σε μια κεντρική τοποθεσία, οι αλγόριθμοι αποστέλλονται στη θέση των δεδομένων για εκπαίδευση, διασφαλίζοντας ότι τα δεδομένα παραμένουν τοπικά και ασφαλή, ενώ παράλληλα επιτρέπουν στα μοντέλα μηχανικής εκμάθησης να επωφελούνται από διαφορετικά σύνολα δεδομένων σε διαφορετικά νοσοκομεία. Για κάθε νέα εικόνα δέρματος και αντίστοιχα κλινικά δεδομένα που προστίθενται στη βάση δεδομένων, το Bayesian δίκτυο ενημερώνει τις παραμέτρους του χρησιμοποιώντας τη μέθοδο MAP. Εάν μια νέα εικόνα παρουσιάζει χαρακτηριστικά καρκινικής βλάβης, η a posteriori πιθανότητα να είναι καρκινική ενημερώνεται με βάση αυτά τα νέα δεδομένα. Η εκτίμηση MAP υπολογίζει την πιθανότητα η νέα βλάβη να είναι καρκινική λαμβάνοντας υπόψη τόσο τα νέα δεδομένα (πιθανοφάνεια) όσο και την προηγούμενη κατανομή (prior). Κατά τη διάρκεια της εκπαίδευσης, μόνο οι ενημερωμένες παράμετροι του μοντέλου αποστέλλονται στους διακομιστές των παρόχων cloud, προστατεύοντας τα πραγματικά δεδομένα ασθενών.

Πολλαπλά Σύννεφα

Ο κεντρικός πάροχος επιλέγει Bayesian δίκτυο που θα εκπαιδευτεί. Το μοντέλο αρχικοποιείται και διανέμεται στους παρόχους και στη συνέχεια στις ενοποιημένες βάσεις δεδομένων, διασφαλίζοντας τον συγχρονισμό, ώστε κάθε βάση να ξεκινά με τις ίδιες παραμέτρους του μοντέλου. Κάθε πάροχος cloud είναι υπεύθυνος για συγκεκριμένες ομάδες νοσοκομείων, με στόχο να αξιοποιήσει τις διαφορετικές παραμέτρους από διάφορα ιδρύματα που κανένα νοσοκομείο δεν θα μπορούσε να επιτύχει μόνο του. Αυτοί οι πάροχοι cloud συνδύαζουν τις ενημερώσεις μοντέλων που λαμβάνουν από τα νοσοκομεία που εξυπηρετούν. Οι συνδυασμένες παράμετροι αποστέλλονται στη συνέχεια στον κεντρικό πάροχο cloud για τελική συγχώνευση και ενημέρωση του κεντρικού μοντέλου. Μια συνάρτηση απώλειας ποσοτικοποιεί το σφάλμα μεταξύ των παρατηρούμενων δεδομένων και των προβλεπόμενων πιθανοτήτων. Το ενημερωμένο δίκτυο Bayesian διανέμεται στη συνέχεια στους παρόχους cloud και, από εκεί, στις τοπικές βάσεις δεδομένων των νοσοκομείων. Αυτή η διαδικασία επαναλαμβάνεται μέχρι τη σύγκλιση, όπου οι αλλαγές στις παραμέτρους της κατανομής πιθανοτήτων υπό όρους (CPD) γίνονται αμελητέες. Στη συνέχεια, τα νοσοκομεία χρησιμοποιούν αυτό το ενημερωμένο μοντέλο για να κάνουν προβλέψεις για τους ασθενείς τους.

Η ασφαλής μεταφορά δεδομένων διασφαλίζεται μέσω της επικοινωνίας API. Για τη διαχείριση των εισερχόμενων και εξερχόμενων δεδομένων, δημιουργείται μια ουρά δεδομένων χρησιμοποιώντας την τεχνική FIFO (First In, First Out). Τα πιο πρόσφατα εισερχόμενα δεδομένα προστίθενται στο τέλος της ουράς και επεξεργάζονται με τη σειρά που εισήχθησαν. Κάθε πηγή δεδομένων (τοπική ενοποιημένη βάση – cloud, cloud – main cloud) έχει τη δική της ουρά, διασφαλίζοντας ότι τα δεδομένα υποβάλλονται σε επεξεργασία με τη σωστή σειρά. Το σύστημα μπορεί να κλιμακωθεί οριζόντια ώστε να χειρίζεται περισσότερα δεδομένα, κατανέμοντας τον φόρτο εργασίας, γεγονός που καθιστά τον υπολογισμό πιο γρήγορο και αποτελεσματικό.

Ένα τέτοιο σύστημα, επακριβώς εκπαιδευμένο για την ανίχνευση χαρακτηριστικών του μελανώματος, θα μπορούσε να εντοπίσει γρήγορα περιπτώσεις υψηλού κινδύνου και να ειδοποιήσει αμέσως τους γιατρούς να δώσουν προτεραιότητα στην περαιτέρω αξιολόγηση. Αυτή η προσέγγιση θα μπορούσε να απλοποιήσει τις διαγνωστικές ροές εργασίας, διασφαλίζοντας ότι τα μελανώματα ανιχνεύονται στα πιο πρώιμα και πιο θεραπεύσιμα στάδια.

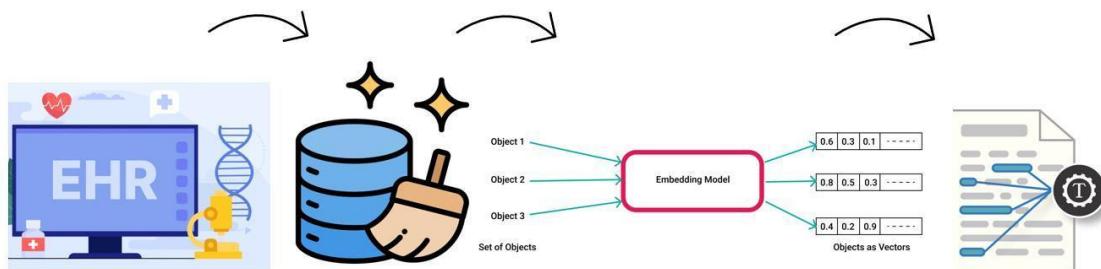


Σχήμα 1.6: Εντοπισμός Σαγματικού Σημείου και Οριοθέτηση Περιοχών Ενδιαφέροντος

Αγωγός Ανάπτυξης CNN:

1. Λήψη ιατρικών εικόνων υψηλής ποιότητας.

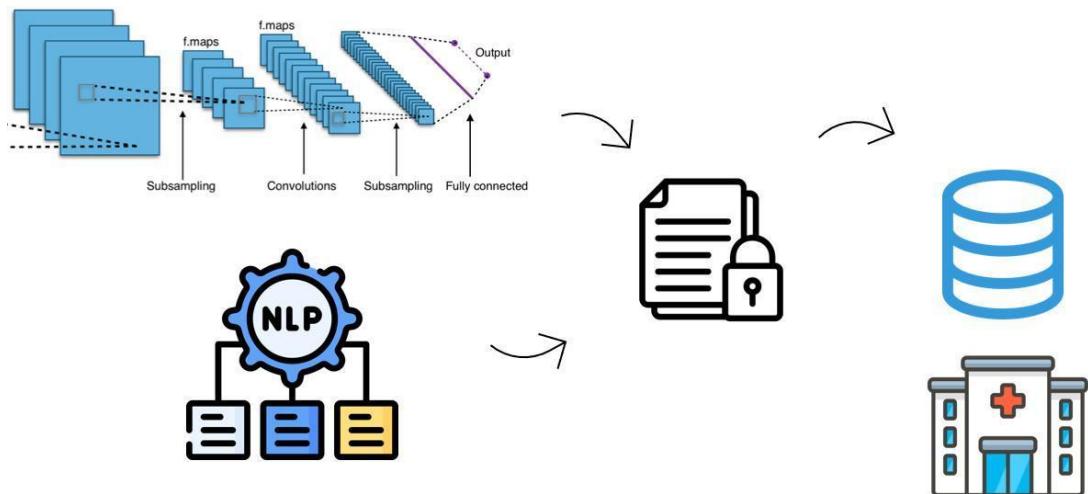
2. Εντοπισμός σαγματικού σημείου, μέσω των τεχνικών Computer Vision (LoG, DoG, Hessian-LoG), για την οριοθέτηση περιοχών ενδιαφέροντος στην εικόνα.
3. Εξαγωγή χαρακτηριστικών υφής των οριοθετούμενων περιοχών, μέσω της τεχνικής GLCM της Radiomics.
4. Εκπαίδευση CNN, χρησιμοποιώντας τα επιλεγμένα χαρακτηριστικά για την αυτόνομη ανίχνευση και ταξινόμηση μελανώματος με υψηλή ακρίβεια, μειώνοντας τη χειρωνακτική προσπάθεια που απαιτείται από τους δερματολόγους.
5. Εφαρμογή του μοντέλου για την αυτοματοποίηση της αξιολόγησης εικόνων δερματικών αλλοιώσεων, ενισχύοντας τη διαγνωστική ταχύτητα και ακρίβεια.



Σχήμα 1.7: Αγωγός Ανάπτυξης Μοντέλου NLP

Αγωγός Ανάπτυξης Μοντέλου NLP:

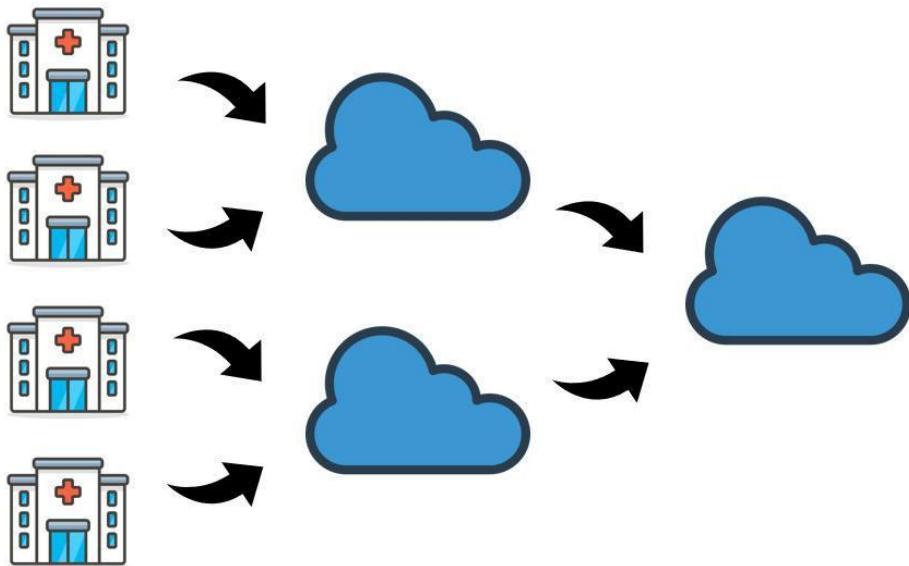
1. Συλλογή ιατρικών αναφορών και κειμένων από ηλεκτρονικά ιατρικά αρχεία (EHR) και κλινικές σημειώσεις.
2. Καθαρισμός των δεδομένων αφαιρώντας σημεία στίξης, ειδικούς χαρακτήρες και κανονικοποιώντας το κείμενο.
3. Προσαρμογή του κειμένου και μετατροπή των διακριτικών σε διανύσματα χρησιμοποιώντας μοντέλα ενσωμάτωσης.
4. Εξαγωγή οντοτήτων όπως ασθένειες, φάρμακα και εργαστηριακά αποτελέσματα.
5. Δημιουργία σύντομης περίληψης των αναφορών.



Σχήμα 1.8: Ροή Εργασίας Ενοποιημένων Τοπικών Βάσεων Δεδομένων

Ροή Εργασίας Ενοποιημένων Τοπικών Βάσεων Δεδομένων:

1. Ενσωμάτωση δεδομένων τόσο από το μοντέλο NLP όσο και από το CNN, του κάθε νοσοκομείου.
2. Κρυπτογράφηση όλων των ιατρικών πληροφοριών, καθιστώντας τις προσβάσιμες μόνο σε εξουσιοδοτημένους κλινικούς ιατρούς με τη συγκατάθεση του ασθενούς.
3. Μετατροπή των δεδομένων σε πολύτιμες πληροφορίες, διευκολύνοντας τη λήψη αποφάσεων και την εκτέλεση σύνθετων ερωτημάτων χωρίς την παρέμβαση στις ανεξάρτητες ροές.
4. Χρήση του FL από κάθε νοσοκομείο για την τοπική εκπαίδευση του Bayesian δικτύου με τα ενοποιημένα δεδομένα τους.
5. Εφαρμογή του θεωρήματος Bayes για ενημέρωση των προηγούμενων πιθανοτήτων με τα παρατηρούμενα δεδομένα.
6. Αποστολή μόνο των ενημερωμένων παραμέτρων των κατανομών πιθανότητας που αναπαριστούν τις εξαρτήσεις μεταξύ των κόμβων του δικτύου στους παρόχους cloud.



Σχήμα 1.9: Ροή Εργασίας Πολλαπλών Σύννεφων

Ροή Εργασίας Πολλαπλών Σύννεφων

Εκπαίδευση Bayesian δικτύου:

1. Ο κεντρικός πάροχος επιλέγει Bayesian δίκτυο που θα εκπαιδευτεί.
2. Το μοντέλο αρχικοποιείται και προετοιμάζεται για διανομή στους παρόχους και έπειτα στις ενοποιημένες βάσεις δεδομένων. Αυτά τα δύο επίπεδα συγχρονισμού διασφαλίζουν ότι κάθε βάση ξεκινά με τις ίδιες παραμέτρους μοντέλου.
3. Διαχείριση συγκεκριμένων ομάδων νοσοκομείων από κάθε πάροχο.
4. Συνδυασμός των ενημερωμένων παραμέτρων, που λαμβάνονται από νοσοκομεία.
5. Αποστολή στον κεντρικό πάροχο cloud για τελική συγχώνευση και ενημέρωση.
6. Διανομή του ενημερωμένου Bayesian δικτύου στους παρόχους cloud και στη συνέχεια στην ενοποιημένη βάση δεδομένων του κάθε νοσοκομείου.
7. Το καθολικό μοντέλο ενημερώνεται χωρίς ο κεντρικός πάροχος να έχει ποτέ πρόσβαση στα πραγματικά δεδομένα από τις βάσεις.
8. Δημιουργήστε μια συνάρτηση απώλειας που ποσοτικοποιεί το σφάλμα μεταξύ των παρατηρούμενων δεδομένων και των προβλεπόμενων πιθανοτήτων.
9. Επαναλάβετε τη διαδικασία μέχρι τη σύγκλιση, δηλαδή μέχρι οι αλλαγές στα CPD να γίνουν αμελητέες.
10. Τα νοσοκομεία χρησιμοποιούν το ενημερωμένο μοντέλο για προβλέψεις ασθενών.

Επικοινωνία:

1. Χρήση API για ασφαλή μεταφορά δεδομένων.
2. Δημιουργία ουράς δεδομένων FIFO (First In, First Out) για την επεξεργασία των πιο πρόσφατων εισερχόμενων δεδομένων με τη σειρά που ελήφθησαν.
3. Κάθε πηγή δεδομένων (τοπική ενοποιημένη βάση – σύννεφο, σύννεφο – κύριο νέφος) έχει τη δική της ουρά για να διασφαλιστεί η σωστή σειρά επεξεργασίας.
4. Το σύστημα είναι οριζόντια επεκτάσιμο για να χειρίζεται αυξανόμενους όγκους δεδομένων, βελτιώνοντας την υπολογιστική απόδοση.

Αυτή η βελτιωμένη ροή εργασίας ενσωματώνει μοντέλα CNN και NLP, ενοποιημένες τοπικές βάσεις δεδομένων, διαχείριση ουράς δεδομένων FIFO και λύσεις multi-cloud για ανάπτυξη του δικτύου Bayesian, ώστε να εξάγει μια πρόβλεψη που υποδεικνύει την πιθανότητα η βλάβη να είναι κακοήθης, η οποία μπορεί να καθοδηγήσει περαιτέρω κλινική δράση για τη βελτίωση της φροντίδας των ασθενών. Το σύστημα υποστηρίζει προηγμένη ανάλυση δεδομένων με προηγμένους αλγορίθμους, διατηρώντας το απόρρητο και την ασφάλεια των ασθενών. Αξιοποιώντας διάφορα σύνολα δεδομένων από ιδρύματα σε όλο τον κόσμο, βοηθά στον εντοπισμό ομάδων υψηλού κινδύνου για κλινικές δοκιμές, στην πρόβλεψη των ανταποκρίσεων των ασθενών στις θεραπείες και στην εύρεση σχετικών βιοδεικτών, προσφέροντας νέες ιδέες για τη θεραπεία του μελανώματος.

4. Μελλοντικές Βελτιώσεις και Επόμενα Βήματα

Το μέλλον της εξατομικευμένης θεραπείας του καρκίνου έγκειται στην ενσωμάτωση της ανάλυσης DNA, για την δημιουργία μοναδικού προφίλ καρκίνου για κάθε ασθενή. Αυτή η προσέγγιση επιτρέπει πιο ακριβή και εξατομικευμένα σχέδια θεραπείας, βελτιώνοντας τα αποτελέσματα και μειώνοντας τις περιπτές παρεμβάσεις. Ωστόσο, καθώς προχωράμε προς αυτές τις εξελίξεις, είναι σημαντικό να αντιμετωπίσουμε τους φόβους και τη δυσπιστία γύρω από την τεχνητή νοημοσύνη και τις εφαρμογές της στην υγειονομική περίθαλψη.

A. Αντιμετώπιση ανησυχιών για AI στην Υγεία

Η ταχεία πρόοδος στα νευρωνικά δίκτυα και την τεχνητή νοημοσύνη έχει οδηγήσει σε ανησυχίες σχετικά με αυτά τα συστήματα, που θέτουν ανεξάρτητα στόχους και λαμβάνουν αποφάσεις χωρίς ανθρώπινη επίβλεψη. Αυτός ο φόβος είναι ιδιαίτερα έντονος στην υγειονομική περίθαλψη, όπου το διακύβευμα είναι απίστευτα υψηλό. Για την οικοδόμηση εμπιστοσύνης, είναι απαραίτητο να διατηρηθεί η διαφάνεια στην ανάπτυξη και στην εφαρμογή της τεχνητής νοημοσύνης. Σαφείς κατευθυντήριες γραμμές και πρότυπα δεοντολογίας θα πρέπει να διέπουν τη χρήση της τεχνητής νοημοσύνης στην υγειονομική περίθαλψη, διασφαλίζοντας ότι αυτά τα συστήματα ενισχύουν την ανθρώπινη λήψη αποφάσεων αντί να την αντικαθιστούν. Πιο συγκεκριμένα, δίνοντας έμφαση στη συνεργασία μεταξύ της τεχνητής νοημοσύνης και των κλινικών ιατρών, μπορούμε να αξιοποιήσουμε τα δυνατά σημεία και των δύο για να βελτιώσουμε την φροντίδα των ασθενών, ενώ αντιμετωπίζουμε ζητήματα ηθικής και ασφάλειας. Αυτή η συλλογική προσέγγιση είναι ιδιαίτερα σημαντική, καθώς στοχεύουμε στην δημιουργία ενός μοναδικού μοριακού προφίλ ασθενή μέσω του DNA.

B. Μετρήσεις σε Πραγματικό Χρόνο

Η ροή εργασιών του συστήματος μπορεί να επεκταθεί με την ανάπτυξη λογισμικού για ζωντανή απομακρυσμένη εξέταση, η οποία καθίσταται δυνατή με την ασφαλή σύνδεση ιατρικών συσκευών στον διακομιστή cloud μέσω HTTPS. Αυτή η ασφαλής σύνδεση διασφαλίζει ότι τα ευαίσθητα ιατρικά δεδομένα μεταδίδονται με ασφάλεια, διατηρώντας το απόρρητο και την ακεραιότητα των δεδομένων του ασθενούς. Επιτρέποντας τη μετάδοση δεδομένων σε πραγματικό χρόνο σε απομακρυσμένες συσκευές, οι πάροχοι υγειονομικής περίθαλψης μπορούν να έχουν άμεση πρόσβαση σε κρίσιμες πληροφορίες ασθενών, ανεξάρτητα από τη φυσική τους τοποθεσία. Τα ιατρικά δεδομένα και οι ακτινογραφίες μεταδίδονται και διαγιγνώσκονται διαδικτυακά μέσα σε λίγα λεπτά, δημιουργώντας την διαγνωστική γεφύρωση των απομακρυσμένων περιοχών. Το cloud απλοποιεί τη διαδικασία για το μη τεχνικό προσωπικό, όπως οι νοσηλευτές των απομακρυσμένων περιοχών, οι οποίοι μπορούν να διεξάγουν την εξέταση και να τη δουν να ανεβάζεται αυτόμata στο cloud για να την ελέγχουν οι ιατροί.

Συνδυάζοντας την ασφαλή μετάδοση δεδομένων με την προηγμένη προγνωστική ανάλυση των δικτύων Bayes, η συνολική αποτελεσματικότητα του συστήματος τηλεϊατρικής ενισχύεται σημαντικά. Αυτή η ενσωμάτωση υποστηρίζει τους επαγγελματίες υγείας στη λήψη καλά ενημερωμένων αποφάσεων γρήγορα, εξασφαλίζοντας έγκαιρη και κατάλληλη φροντίδα. Κατά συνέπεια, η χρήση αυτών των τεχνολογιών όχι μόνο βελτιώνει τη διαγνωστική ακρίβεια και τα αποτελέσματα της θεραπείας, αλλά και βελτιστοποιεί τη ροή εργασιών εντός του συστήματος υγειονομικής περιθαλψης, οδηγώντας τελικά σε καλύτερες εμπειρίες ασθενών και βελτιωμένα αποτελέσματα υγείας.

Γ. Καρκίνος, DNA και Εξατομικευμένη Θεραπεία

Ο καρκίνος συχνά προκύπτει από την αποτυχία ή την κακή αποκατάσταση της βλάβης του DNA, που οδηγεί σε μεταλλάξεις και ανεξέλεγκτη κυτταρική ανάπτυξη, καθιστώντας τον καρκίνο γενετική ασθένεια. Ωστόσο, επηρεάζονται κυρίως τα σωματικά κύτταρα, που αποτελούν τους ιστούς και τα όργανα, και όχι τα αναπαραγωγικά κύτταρα που είναι υπεύθυνα για τη μετάδοση γενετικών πληροφοριών στους απογόνους. Τα περισσότερα από τα κύτταρα του σώματος περιέχουν ένα πλήρες σύνολο DNA, γνωστό ως γονιδίωμα, το οποίο λειτουργεί σαν εγκυκλοπαίδεια. Αυτό το γονιδίωμα χωρίζεται σε χρωμοσώματα, παρόμοια με τα κεφάλαια της εγκυκλοπαίδειας, και μέσα σε αυτά τα χρωμοσώματα βρίσκονται τα γονίδιά μας, παρόμοια με μεμονωμένες σελίδες εγκυκλοπαίδειας. Η σειρά των γονιδίων είναι κρίσιμη και δεν είναι όλα τα γονίδια ενεργά σε κάθε κύτταρο. Τα συγκεκριμένα γονίδια που διαβάζονται ή εκφράζονται σε κάθε κύτταρο καθορίζουν τη λειτουργία και τα χαρακτηριστικά του κυττάρου.

Μέσα στα κύτταρα, τα γονίδια διαβάζονται σε μόρια, κυρίως πρωτεΐνες, που εκτελούν διάφορες λειτουργίες. Ο τύπος και η ποσότητα των πρωτεΐνων που παράγονται, και ο χρόνος παραγωγής τους, υπαγορεύουν τις δραστηριότητες του κυττάρου. Η έκφραση πρωτεΐνων είναι εξαιρετικά περίπλοκη και δυναμική, με συγκεκριμένα μοτίβα καθιερωμένα για διαφορετικούς τύπους κυττάρων. Το DNA, το θεμελιώδες δομικό στοιχείο της ζωής, είναι εγγενώς ασταθές και μπορεί να καταστραφεί έως και 100.000 φορές την ημέρα σε κάθε κύτταρο. Αυτή η αστάθεια, ενώ αποτελεί φυσικό μέρος των κυτταρικών διεργασιών, μπορεί να οδηγήσει σε μεταλλάξεις εάν δεν αντιμετωπιστεί σωστά. Κάθε περίπτωση βλάβης έχει τη δυνατότητα να προκαλέσει μια μετάλλαξη, η οποία μπορεί να έχει ως αποτέλεσμα ένα κύτταρο να χάσει την ικανότητά του να ελέγχει τον εαυτό του, οδηγώντας σε ανεξέλεγκτη κυτταρική ανάπτυξη και διαίρεση, που είναι χαρακτηριστικό του καρκίνου.

Ο κίνδυνος ανάπτυξης καρκίνου επιδεινώνεται περαιτέρω όταν υπάρχουν κληρονομικά ελαττώματα στους μηχανισμούς επιδιόρθωσης του DNA. Αυτά τα ελαττώματα μπορούν να εμποδίσουν την ικανότητα ενός κυττάρου να επιδιορθώσει τη βλάβη του DNA, αυξάνοντας την πιθανότητα μεταλλάξεων και, κατά συνέπεια, την ανάπτυξη καρκίνου. Αν και τα κύτταρα έχουν εξελίξει μηχανισμούς για τη διόρθωση αυτών των σφαλμάτων, όπως η διόρθωση της

πολυμεράσης του DNA και τα διάφορα ένζυμα επιδιόρθωσης του DNA, ορισμένες μεταλλάξεις επιμένουν και μεταβιβάζονται στα θυγατρικά κύτταρα. Με την πάροδο του χρόνου, αυτές οι μεταλλάξεις μπορεί να συσσωρευτούν και ενδεχομένως να οδηγήσουν σε καρκίνο.

Ωστόσο, εάν οι ασθενείς μπορούσαν να λάβουν εξατομικευμένες θεραπείες με βάση τη μοναδική αλληλουχία του γονιδιώματος τους, θα βελτίωνε σημαντικά τα αποτελέσματα και θα άλλαζε την πορεία της νόσου. Το ανθρώπινο γονιδίωμα περιλαμβάνει περίπου 20000 – 25000 γονίδια, οι άνθρωποι μοιάζουν πολύ μεταξύ τους, κατά 99,9% και οι διαφορές μεταξύ τους συγκεντρώνονται στο υπόλοιπο 0,1% των γονιδίων. Για να προσδιοριστεί η πιθανότητα δύο άτομα, στο παρελθόν ή μέλλον, να είναι γενετικά πανομοιότυπα πρέπει κανείς να εξετάσει τον τεράστιο αριθμό συνδυασμών γονιδίων που διαφέρουν μεταξύ τους, μέχρι να καταλήξουμε σε ανθρώπους που είναι πανομοιότυποι. Υπάρχουν 10^{30} παραλλαγές, συνεπώς η πιθανότητα είναι σχεδόν μηδενική, καθιστώντας κάθε άτομο μοναδικό. Μόνο οι ζυγωματικοί δίδυμοι έχουν πανομοιότυπο DNA αλλά μόνο κατά την σύλληψη τους και στην συνέχεια ακόμα και σε αυτούς διαφοροποιείται. Αυτή η γενετική μοναδικότητα αποτελεί τη βάση για την εξατομικευμένη θεραπεία του καρκίνου, δίνοντας την ικανότητα ακριβούς πρόβλεψης των θεραπευτικών σχημάτων, συμπεριλαμβανομένων στοχευμένων φαρμάκων και δόσεων, που ταιριάζουν καλύτερα για κάθε ασθενή με βάση τα μοναδικά μοριακά, γενετικά και βασισμένα στον όγκο χαρακτηριστικά

Η αλληλουχία γονιδιώματος, ιδιαίτερα η αλληλουχία επόμενης γενιάς (NGS), μπορεί να χρησιμοποιήσει μοντέλα μηχανικής μάθησης (ML) και βαθιάς μάθησης (DL) για την πρόβλεψη μεταλλάξεων και επιπέδων γονιδιακής έκφρασης. Αυτά τα μοντέλα μπορούν να αναλύσουν μοτίβα βλάβης και επιδιόρθωσης του DNA, να εντοπίσουν πιθανές μεταλλάξεις και να προβλέψουν την πιθανότητα καρκίνου. Διατηρώντας την ακεραιότητα του DNA και παρεμβαίνοντας νωρίς στη διαδικασία ανάπτυξης του καρκίνου, η τεχνητή νοημοσύνη έχει τη δυνατότητα να μεταμορφώσει την ογκολογία. Αν και οι λύσεις που βασίζονται στην τεχνητή νοημοσύνη βρίσκονται ακόμη στα αρχικά τους στάδια, η ικανότητά τους να ταξινομούν τον καρκίνο με βάση τη γονιδιακή δραστηριότητα και να προβλέπουν την υποτροπή είναι πολλά υποσχόμενη. Για παράδειγμα, η ανάλυση μεθυλίωσης DNA έχει δείξει υψηλή ακρίβεια στην κατηγοριοποίηση των όγκων του εγκεφάλου, υποδεικνύοντας τη δυνατότητα του AI να βελτιώσει τη διάγνωση και την πρόγνωση. Αυτή η πιθανή μεταμόρφωση ανοίγει το δρόμο για πιο εξατομικευμένες θεραπευτικές προσεγγίσεις, προσφέροντας ελπίδα για καλύτερη διαχείριση και αποτελέσματα στη θεραπεία του καρκίνου.

5. Σύνοψη και Συμπέρασμα

Σύνοψη

Αρχικά, όπως προκύπτει από την παρούσα μελέτη, παρά τις τακτικές εξετάσεις στις οποίες υποβάλλονται οι άνθρωποι, πολλές περιπτώσεις καρκίνου δεν εντοπίζονται έγκαιρα λόγω της έλλειψης κατάλληλων τεχνολογικών εργαλείων. Η έγκαιρη διάγνωση της νόσου, που αποτελεί μία από τις κύριες αιτίες θνησιμότητας παγκοσμίως, είναι ιδιαίτερα κρίσιμη. Σε αυτό έρχεται να συμβάλλει τα τελευταία χρόνια η ανάπτυξη της τεχνητής νοημοσύνης, η οποία έχει αυξήσει κατά πολύ την αξία των δεδομένων, καθιστώντας τα απαραίτητα για την βελτίωση της υγείας. Τόσο η τεχνητή νοημοσύνη όσο και η μηχανική μάθηση βελτιώνουν την έγκαιρη ανίχνευση καρκίνου μέσω της ανάλυσης ιατρικών εικόνων και δεδομένων υγείας, κάνοντας στοχευμένες προβλέψεις κινδύνου. Επίσης, ενισχύουν την ακρίβεια της διάγνωσης και διευκολύνουν την ανάπτυξη εξατομικευμένων σχεδίων θεραπείας, βελτιώνοντας την φροντίδα των ασθενών. Η συγκεκριμένη έρευνα προτείνει νέες ιδέες και βελτιώσεις για την εξατομικευμένη διάγνωση, αξιοποιώντας τις τελευταίες εξελίξεις στην τεχνητή νοημοσύνη και τη μηχανική μάθηση καθώς και συγχώνευση διαφορετικών διαγνωστικών στρατηγικών για μια πιο ολοκληρωμένη προσέγγιση.

Στη συνέχεια παρουσιάζονται μερικές πρόσφατες μελέτες και γίνεται αναφορά στο Πλαγκόσμιο Παρατηρητήριο Καρκίνου, το οποίο υπογραμμίζει την κρίσιμη ανάγκη για έγκαιρη διάγνωση και βελτιωμένες θεραπείες, εστιάζοντας σε τρεις κύριους τύπους καρκίνου: του πνεύμονα, του μαστού και του δέρματος και προσδιορίζονται καρκινογόνες ουσίες και προληπτικά μέτρα για κάθε τύπο, τονίζοντας τη σημασία της έγκαιρης διάγνωσης και της συντονισμένης διεθνούς ανταπόκρισης. Έπειτα, αναφέρονται οι έννοιες της Τεχνητής Νοημοσύνης, της Μηχανικής Μάθησης και της Βαθιάς Μάθησης και ο τρόπος που μπορούν να συμβάλλουν στην βελτίωση των αποτελεσμάτων των ασθενών. Επίσης, γίνεται αναφορά στα συνελικτικά νευρωνικά δίκτυα (CNN) και στα δίκτυα Bayes, τα οποία ενισχύουν την ικανότητα διάγνωσης και πρόβλεψης του καρκίνου, παρέχοντας ισχυρά εργαλεία για τη βελτίωση της φροντίδας των ασθενών.

Στην επόμενη ενότητα παρουσιάζονται διάφορες τεχνικές που θα οδηγήσουν σε μια πιο ολοκληρωμένη προσέγγιση για την έγκαιρη διάγνωση και αντιμετώπιση του καρκίνου. Προτείνεται η χρήση τεχνολογιών υπολογιστικής όρασης και radiomics για την ανάλυση ιατρικών εικόνων, σε συνδυασμό με την επεξεργασία φυσικής γλώσσας (NLP) για την ενσωμάτωση πολυτροπικών δεδομένων, δημιουργώντας έτσι μια ενιαία βάση δεδομένων σε κάθε νοσοκομείο. Ωστόσο, μερικές προκλήσεις όπως ο όγκος των δεδομένων και η διαλειτουργικότητα παραμένουν, ενώ νέες προσεγγίσεις ενσωματώνουν πολλαπλές τεχνικές για πληρέστερη διάγνωση. Έπειτα, παρουσιάζεται μία καινοτόμος προσέγγιση για την ενσωμάτωση πολυτροπικών πηγών δεδομένων, η οποία αποδεικνύει τις δυνατότητές της σε μια πρωτοποριακή μελέτη

περίπτωσης για την έγκαιρη διάγνωση του καρκίνου του δέρματος, ειδικά του μελανώματος. Αξιοποιώντας τις τεχνικές computer vision για ανάλυση εικόνας, radiomics για ποσοτικά δεδομένα από δερματοσκοπικές εικόνες και πληροφορίες από κλινικές σημειώσεις μέσω NLP, αυτή η μέθοδος προσφέρει μια ολοκληρωμένη εικόνα για τις βλάβες που προκύπτουν. Επιπλέον, η χρήση τεχνικών όπως LoG, DoG, και GLCM για ανίχνευση και ανάλυση χαρακτηριστικών εικόνας, σε συνδυασμό με τη συσχέτιση των δεδομένων αυτών με κλινικά αποτελέσματα μέσω NLP, οδηγεί σε ακριβέστερη και έγκαιρη διάγνωση και βελτιστοποίηση της θεραπείας.

Τέλος, παρουσιάζονται μερικές μελλοντικές βελτιώσεις και επόμενα βήματα που μπορούν να ακολουθηθούν στην εξατομικευμένη θεραπεία του καρκίνου, εστιάζοντας στην ενσωμάτωση της ανάλυσης DNA για τη δημιουργία εξατομικευμένων προφίλ καρκίνου για κάθε ασθενή. Επιπλέον, είναι πολύ σημαντικό να αντιμετωπιστούν οι ανησυχίες γύρω από την τεχνητή νοημοσύνη στην υγειονομική περίθαλψη, διατηρώντας τη διαφάνεια και τη συνεργασία μεταξύ τεχνητής νοημοσύνης και κλινικών ιατρών, για να διασφαλιστεί ότι αυτά τα συστήματα ενισχύουν αντί να αντικαθιστούν την ανθρώπινη λήψη αποφάσεων. Επίσης, η απομακρυσμένη παρακολούθηση σε πραγματικό χρόνο μέσω ασφαλών συνδέσεων cloud και προηγμένης προγνωστικής ανάλυσης ενισχύει περαιτέρω την αποτελεσματικότητα της υγειονομικής περίθαλψης. Επομένως, η κατανόηση της γενετικής βάσης του καρκίνου υπογραμμίζει τη σημασία των εξατομικευμένων θεραπειών, προσαρμοσμένων στα μεμονωμένα γονιδιωματικά προφίλ, ενισχύοντας σημαντικά τη διαχείριση της νόσου και τα αποτελέσματα των ασθενών.

Συμπέρασμα

Συμπερασματικά, στη σύγχρονη εποχή, η ραγδαία παραγωγή δεδομένων, τα έχει τοποθετήσει ως κρίσιμο πόρο, ιδιαίτερα στην υγειονομική περίθαλψη. Η ικανότητα αξιοποίησης και ανάλυσης τεράστιων ποσοτήτων δεδομένων έχει τη δυνατότητα να φέρει επανάσταση στην έγκαιρη ανίχνευση καρκίνου, βελτιώνοντας έτσι τα αποτελέσματα της θεραπείας και σώζοντας ζωές. Ο καρκίνος παραμένει η κύρια αιτία νοσηρότητας και θνησιμότητας παγκοσμίως, με την έγκαιρη ανίχνευση να ενισχύει σημαντικά τις πιθανότητες αποτελεσματικής θεραπείας. Ωστόσο, οι παραδοσιακές μέθοδοι προσυμπτωματικού ελέγχου συχνά αποτυγχάνουν να ανιχνεύσουν την νόσο στα εκκολαπτόμενα στάδια της, γεγονός που καθιστά αναγκαία την ανάπτυξη προηγμένων διαγνωστικών εργαλείων.

Η τεχνητή νοημοσύνη, η μηχανική μάθηση και η βαθιά μάθηση έχουν αναδειχθεί ως ισχυροί σύμμαχοι σε αυτόν τον αγώνα, ικανοί να αναλύουν πολύπλοκα ιατρικά δεδομένα για να ανιχνεύσουν διακριτικά μοτίβα που θα μπορούσαν να διαφεύγουν της ανθρώπινης παρατήρησης. Ενσωματώνοντας αυτές τις τεχνολογίες σε διαγνωστικές διαδικασίες, βελτιώνεται σημαντικά η

ακρίβεια και η αποτελεσματικότητα, ανοίγοντας το δρόμο στην ιατρική ακρίβειας. Τα τεχνητά νευρωνικά δίκτυα (ANN) και τα συνελικτικά νευρωνικά δίκτυα (CNN) έχουν δείξει πολλά υποσχόμενα στη διάγνωση καρκίνων όπως ο καρκίνος του μαστού και του πνεύμονα με υψηλή ακρίβεια, μέσω της αναγνώρισης προτύπων, ωστόσο συχνά στερούνται ερμηνείας και απαιτούν σημαντικούς υπολογιστικούς πόρους για την εκπαίδευσή τους. Σε αντίθεση τα δίκτυα Bayes προσφέρουν ισχυρές λύσεις στην λήψη αποφάσεων σε συνθήκες αβεβαιότητας, για την πρόβλεψη της εξέλιξης του καρκίνου και τον χειρισμό πιθανολογικών γεγονότων στα κλινικά δεδομένα.

Η προτεινόμενη ροή εργασίας του συστήματος ενσωματώνει πολυτροπικές πηγές δεδομένων, συνδυάζοντας τεχνικές υπολογιστικής όρασης και ραδιενέργειας για την ανάλυση των ιατρικών εικόνων και επικύρωση των δεδομένων μέσω της επεξεργασία φυσικής γλώσσας (NLP) για εξαγωγή κλινικών δεδομένων. Αυτή η ολοκληρωμένη προσέγγιση διασφαλίζει μια πιο ακριβή και λεπτομερή κατανόηση της νόσου, οδηγώντας σε καλύτερα διαγνωστικά και θεραπευτικά αποτελέσματα. Ο σχεδιασμός του συστήματος δίνει έμφαση στο απόρρητο και την ασφάλεια των δεδομένων, χρησιμοποιώντας ομοσπονδιακή μάθηση και λύσεις multi-cloud για να επιτρέψει τη συνεργατική εκπαίδευση μοντέλων μεταξύ των ιδρυμάτων χωρίς να διακυβεύονται τα δεδομένα ασθενών.

Κοιτάζοντας το μέλλον, η ενσωμάτωση της ανάλυσης DNA για τη δημιουργία μοναδικού προφίλ καρκίνου για κάθε ασθενή υπόσχεται περαιτέρω ενίσχυση της εξατομικευμένης θεραπείας. Αυτή η προσέγγιση, σε συνδυασμό με τη μετάδοση και ανάλυση δεδομένων σε πραγματικό χρόνο, έχει τη δυνατότητα να μεταμορφώσει την ογκολογία. Η αντιμετώπιση των ανησυχιών σχετικά με την τεχνητή νοημοσύνη στην υγειονομική περίθαλψη μέσω της διαφάνειας και των ηθικών προτύπων θα είναι ζωτικής σημασίας για την απόκτηση εμπιστοσύνης και τη διασφάλιση της επιτυχούς εφαρμογής αυτών των τεχνολογιών.

Η συνεχής έρευνα και ανάπτυξη είναι ουσιαστικής σημασίας για την αντιμετώπιση των προκλήσεων που θέτει ο καρκίνος και την επίτευξη αποτελεσματικότερης ιατρικής περίθαλψης. Η ενσωμάτωση των παρουσιαζόμενων ροών εργασίας στην ογκολογία συμπληρώνει την ανθρώπινη τεχνογνωσία, παρέχοντας προηγμένα εργαλεία για τη βελτίωση της ακρίβειας και αποτελεσματικότητας της διάγνωσης και της θεραπείας. Αυτή η συνέργεια ανθρώπου και τεχνολογίας μπορεί να οδηγήσει σε μια νέα εποχή στην ιατρική, όπου οι ασθενείς θα λαμβάνουν πιο εξατομικευμένες και αποτελεσματικές θεραπείες, αυξάνοντας έτσι τις πιθανότητες για μεγαλύτερα ποσοστά επιβίωσης. Τελικά, ο πιο κρίσιμος παράγοντας που επιταχύνει την συλλογική αυτή προσπάθεια είναι η ακόρεστη περιέργεια της ανθρώπινης φύσης και η ανάγκη επίλυσης σύνθετων προβλημάτων.

6. Βιβλιογραφία

Almufareh, MF. Tariq, N. Humayun, M. Almas B, A. Federated Learning Approach to Breast Cancer Prediction in a Collaborative Learning Framework. *Healthcare*. 2023; 11(24):3185.

<https://doi.org/10.3390/healthcare11243185>

Briganti, G. and Le Moine, O. (2020) *Artificial Intelligence in medicine: Today and Tomorrow, Frontiers in medicine*.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7012990/>.

Bertolaccini, L. et al. (2017) An overview of the use of artificial neural networks in Lung Cancer Research, *Journal of Thoracic Disease*.
<https://jtd.amegroups.org/article/view/12962/html>.

Capobianco E, Deng J. Radiomics at a Glance: A Few Lessons Learned from Learning Approaches. *Cancers (Basel)*. 2020 Aug 29;12(9):2453. doi: 10.3390/cancers12092453. PMID: 32872466; PMCID: PMC7563283.

Cobo, M., Menéndez Fernández-Miranda, P., Bastarrika, G. et al. Enhancing radiomics and Deep Learning systems through the standardization of medical imaging workflows. *Sci Data* 10, 732 (2023).
<https://doi.org/10.1038/s41597-023-02641-x>

Faruqui N, Yousuf MA, Kateb FA, Abdul Hamid M, Monowar MM. Healthcare As a Service (HAAS): CNN-based cloud computing model for ubiquitous access to lung cancer diagnosis. *Heliyon*. 2023 Oct 27;9(11):e21520. doi: 10.1016/j.heliyon.2023.e21520. PMID: 37942151; PMCID: PMC10628703.

Gholipour, M., Khajouei, R., Amiri, P. et al. Extracting cancer concepts from clinical notes using natural language processing: a systematic review. *BMC Bioinformatics* 24, 405 (2023).
<https://doi.org/10.1186/s12859-023-05480-0>

Heaton, J. Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning. *Genet Program Evolvable Mach* **19**, 305–307 (2018).
<https://doi.org/10.1007/s10710-017-9314-z>

Hunter, B.; Hindocha, S.; Lee, R.W. The Role of Artificial Intelligence in Early Cancer Diagnosis. *Cancers* 2022, 14, 1524. <https://doi.org/10.3390/cancers14061524>

Iannantuono GM, Bracken-Clarke D, Floudas CS, Roselli M, Gulley JL, Karzai F. Applications of large language models in cancer care: current evidence and future perspectives. *Front Oncol*. 2023 Sep 4;13:1268915.

doi: 10.3389/fonc.2023.1268915. PMID: 37731643; PMCID: PMC10507617.

Ijaz MF, Woźniak M. Editorial: Recent Advances in Deep Learning and Medical Imaging for Cancer Treatment. *Cancers*. 2024; 16(4):700. <https://doi.org/10.3390/cancers16040700>

Jiang, X. *et al.* (2023) Deep learning for medical image-based cancer diagnosis, MDPI. <https://doi.org/10.3390%2Fcancers15143608>.

Koh, DM., Papanikolaou, N., Bick, U. *et al.* Artificial intelligence and machine learning in cancer imaging. *Commun Med* 2, 133 (2022). <https://doi.org/10.1038/s43856-022-00199-0>

Krzyszczuk P, Acevedo A, Davidoff EJ, Timmins LM, Marrero-Berrios I, Patel M, White C, Lowe C, Sherba JJ, Hartmansenn C, O'Neill KM, Balter ML, Fritz ZR, Androulakis IP, Schloss RS, Yarmush ML. The growing role of precision and personalized medicine for cancer treatment. *Technology* (Singap World Sci). 2018 Sep-Dec;6(3-4):79-100. doi: 10.1142/S2339547818300020. Epub 2019 Jan 11. PMID: 30713991; PMCID: PMC6352312.

Kumar Y, Gupta S, Singla R, Hu YC. A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis. *Arch Comput Methods Eng*. 2022;29(4):2043-2070. doi: 10.1007/s11831-021-09648-w. Epub 2021 Sep 27. PMID: 34602811; PMCID: PMC8475374.

Lin H, Ni L, Phuong C, Hong JC. Natural Language Processing for Radiation Oncology: Personalizing Treatment Pathways. *Pharmgenomics Pers Med*. 2024;17:65-76 <https://doi.org/10.2147/PGPM.S396971>

Liu, X. *et al.* (2019) 'A comparison of deep learning performance against health-care professionals in detecting diseases from Medical Imaging: A systematic review and meta-analysis', *The Lancet Digital Health*, 1(6). doi:10.1016/s2589-7500(19)30123-2.

Mahmoud NM, Soliman AM. Early automated detection system for skin cancer diagnosis using artificial intelligent techniques. *Sci Rep*. 2024 Apr 28;14(1):9749. doi: 10.1038/s41598-024-59783-0. PMID: 38679633; PMCID: PMC11056372.

Mazhar T, Haq I, Ditta A, Mohsan SAH, Rehman F, Zafar I, Gansau JA, Goh LPW. The Role of Machine Learning and Deep Learning Approaches for the Detection of Skin Cancer. *Healthcare*. 2023; 11(3):415. <https://doi.org/10.3390/healthcare11030415>

McKay, C. (2023, September 9). *Paige and Microsoft are building the world's largest AI model to detect cancer.* Maginative. <https://www.maginative.com/article/paige-and-microsoft-are-building-the-worlds-largest-ai-model-to-detect-cancer/>

Mostavi, M., Chiu, YC., Huang, Y. *et al.* Convolutional neural network models for cancer type prediction based on gene expression. *BMC Med Genomics* 13 (Suppl 5), 44 (2020). <https://doi.org/10.1186/s12920-020-0677-2>

Pati, S., Baid, U., Edwards, B. *et al.* Federated learning enables big data for rare cancer boundary detection. *Nat Commun* 13, 7346 (2022). <https://doi.org/10.1038/s41467-022-33407-5>

Pesapane F, De Marco P, Rapino A, Lombardo E, Nicosia L, Tantrige P, Rotili A, Bozzini AC, Penco S, Dominelli V, Trentin C, Ferrari F, Farina M, Meneghetti L, Latronico A, Abbate F, Origgi D, Carrafiello G, Cassano E. How Radiomics Can Improve Breast Cancer Diagnosis and Treatment. *J Clin Med.* 2023 Feb 9;12(4):1372. doi: 10.3390/jcm12041372. PMID: 36835908; PMCID: PMC9963325.

Post AR, Birmingham Z, Halwani AS. Electronic Health Record Data in Cancer Learning Health Systems: Challenges and Opportunities. *JCO Clin Cancer Inform.* 2022 Mar;6:e2100158. doi: 10.1200/CCI.21.00158. PMID: 35353547; PMCID: PMC9005105.

Savova GK, Danciu I, Alamudun F, Miller T, Lin C, Bitterman DS, Tourassi G, Warner JL. Use of Natural Language Processing to Extract Clinical Cancer Phenotypes from Electronic Medical Records. *Cancer Res.* 2019 Nov 1;79(21):5463-5470. doi: 10.1158/0008-5472.CAN-19-0579. Epub 2019 Aug 8. PMID: 31395609; PMCID: PMC7227798.

Sebastian AM, Peter D. Artificial Intelligence in Cancer Research: Trends, Challenges and Future Directions. *Life (Basel).* 2022 Nov 28;12(12):1991. doi: 10.3390/life12121991. PMID: 36556356; PMCID: PMC9786074.

Sheba Macheka, Peng Yun Ng, Ophira Ginsburg, Andrew Hope, Richard Sullivan, Ajay Aggarwal - Prospective evaluation of artificial intelligence (AI) applications for use in cancer pathways following diagnosis: a systematic review: *BMJ Oncology* 2024;3:e000255. <https://doi.org/10.1136/bmjonc-2023-000255>

Silva HECD, Santos GNM, Leite AF, Mesquita CRM, Figueiredo PTS, Stefani CM, de Melo NS. The use of artificial intelligence tools in cancer detection compared to the traditional diagnostic imaging methods: An overview of the systematic reviews. *PLoS One.* 2023 Oct 5;18(10):e0292063. doi:

10.1371/journal.pone.0292063. PMID: 37796946; PMCID: PMC10553229.

Sorbini M, Carradori T, Togliatto GM, Vaisitti T, Deaglio S. Technical Advances in Circulating Cell-Free DNA Detection and Analysis for Personalized Medicine in Patients' Care. *Biomolecules*. 2024; 14(4):498. <https://doi.org/10.3390/biom14040498>

Subashchandrabose U, John R, Anbazhagu UV, Venkatesan VK, Thyluru Ramakrishna M. Ensemble Federated Learning Approach for Diagnostics of Multi-Order Lung Cancer. *Diagnostics (Basel)*. 2023 Sep 25;13(19):3053. doi: 10.3390/diagnostics13193053. PMID: 37835796; PMCID: PMC10572651.

Sufyan, M., Shokat, Z., & Ashfaq, U. A. (2023). Artificial Intelligence in cancer diagnosis and therapy: Current status and future perspective. *Computers in Biology and Medicine*, 165, 107356. <https://doi.org/10.1016/j.combiomed.2023.107356>

Tandon R, Agrawal S, Rathore NPS, Mishra AK, Jain SK. A systematic review on deep learning-based automated cancer diagnosis models. *J Cell Mol Med*. 2024 Mar;28(6):e18144. doi: 10.1111/jcmm.18144. PMID: 38426930; PMCID: PMC10906380.

Verma M. Personalized medicine and cancer. *J Pers Med*. 2012 Jan 30;2(1):1-14. doi: 10.3390/jpm2010001. PMID: 25562699; PMCID: PMC4251363.

Wang L. Deep Learning Techniques to Diagnose Lung Cancer. *Cancers*. 2022; 14(22):5569. <https://doi.org/10.3390/cancers14225569>

Wang L, Fu S, Wen A, Ruan X, He H, Liu S, Moon S, Mai M, Riaz IB, Wang N, Yang P, Xu H, Warner JL, Liu H. Assessment of Electronic Health Record for Cancer Research and Patient Care Through a Scoping Review of Cancer Natural Language Processing. *JCO Clin Cancer Inform*. 2022 Jul;6:e2200006. doi: 10.1200/CCI.22.00006. PMID: 35917480; PMCID: PMC9470142.

Xu Y, Wu T, Gao F, Charlton JR, Bennett KM. Improved small blob detection in 3D images using jointly constrained deep learning and Hessian analysis. *Sci Rep*. 2020 Jan 15;10(1):326. doi: 10.1038/s41598-019-57223-y. PMID: 31941994; PMCID: PMC6962386.

Zhang B, Shi H, Wang H. Machine Learning and AI in Cancer Prognosis, Prediction, and Treatment Selection: A Critical Approach. *J Multidiscip*

Healthc. 2023 Jun 26;16:1779-1791. doi: 10.2147/JMDH.S410301. PMID: 37398894; PMCID: PMC10312208.

Zhang W, Guo Y, Jin Q. Radiomics and Its Feature Selection: A Review. Symmetry. 2023; 15(10):1834. <https://doi.org/10.3390/sym15101834>

Zhu Z, Wang SH, Zhang YD. A Survey of Convolutional Neural Network in Breast Cancer. Comput Model Eng Sci. 2023 Mar 9;136(3):2127-2172. doi: 10.32604/cmes.2023.025484. PMID: 37152661; PMCID: PMC7614504.