



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΗΣ ΣΧΕΔΙΑΣΗΣ
ΚΑΙ ΠΑΡΑΓΩΓΗΣ



**«Η ταξινόμηση ακτινολογικών δεδομένων με τη χρήση
νευρωνικών δικτύων»**

Αθήνα 2020

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Ως κάτωθι υπογεγραμμένος/η Ντουβλέτης Ιωάννης, του Σταματίου, φοιτητής του Τμήματος Μηχανικών Βιομηχανικής Σχεδίασης και Παραγωγής, του Πανεπιστημίου Δυτικής Αττικής, πριν αναλάβω την εκπόνηση της Πτυχιακής Εργασίας μου δηλώνω ότι ενημερώθηκα για τα παρακάτω:

«Η Πτυχιακή Εργασία (ΠΕ) αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο του συγγραφέα, όσο και του Ιδρύματος και θα πρέπει να έχει μοναδικό χαρακτήρα και πρωτότυπο περιεχόμενο.

Απαγορεύεται αυστηρά οποιοδήποτε κομμάτι κειμένου της να εμφανίζεται αυτούσιο ή μεταφρασμένο από κάποια άλλη δημοσιευμένη πηγή. Κάθε τέτοια πράξη αποτελεί προϊόν λογοκλοπής και εγείρει θέμα Ηθικής Τάξης για τα πνευματικά δικαιώματα του άλλου συγγραφέα. Αποκλειστικός υπεύθυνος είναι ο συγγραφέας της ΠΕ. Οποιοσδήποτε φέρει και την ευθύνη των συνεπειών, ποινικών και άλλων, αυτής της πράξης.

Πέραν των όποιων ποινικών ευθυνών του συγγραφέα, σε περίπτωση που το Ίδρυμα του έχει απονείμει Πτυχίο, αυτό ανακαλείται με απόφαση της Συνέλευσης του Τμήματος. Η Συνέλευση του Τμήματος με νέα απόφασή της, μετά από αίτηση του ενδιαφερόμενου, του αναθέτει εκ νέου την εκπόνηση ΠΕ με άλλο θέμα και διαφορετικό επιβλέποντα καθηγητή. Η εκπόνηση της εν λόγω ΠΕ πρέπει να ολοκληρωθεί εντός τουλάχιστον ενός ημερολογιακού μήνου από την ημερομηνία ανάθεσής της.

Ο Δηλών

Ημερομηνία

Περιεχόμενα

Κατάλογος Εικόνων	5
Περίληψη	6
1 Γενικά για τα νευρωνικά δίκτυα	1
1.1.1 <i>Βιολογικά νευρωνικά δίκτυα</i>	<i>1</i>
1.1.2 <i>Τεχνητά νευρωνικά δίκτυα</i>	<i>3</i>
1.1.2.1 <i>Σύντομη ιστορική ανασκόπηση</i>	<i>4</i>
1.1.2.2 <i>Βασικές αρχές τεχνητών νευρωνικών δικτύων</i>	<i>6</i>
1.1.2.3 <i>Αρχιτεκτονική τεχνητών νευρωνικών δικτύων</i>	<i>10</i>
1.1.3 <i>Σύγκριση βιολογικών και τεχνητών νευρωνικών δικτύων.....</i>	<i>12</i>
1.1.4 <i>Σημερινές προκλήσεις στην ανάπτυξη των τεχνητών νευρωνικών δικτύων.....</i>	<i>14</i>
2 Ψηφιακή απεικόνιση ιατρικών δεδομένων	16
2.1 <i>Τρόπος λειτουργίας.....</i>	<i>18</i>
2.2 <i>Διαχείριση και κατηγοριοποίηση ιατρικών εικόνων με βάση το DICOM για σωστή διάγνωση.....</i>	<i>20</i>
2.2.1 <i>Αποθήκευση ιατρικών εικόνων</i>	<i>21</i>
2.2.2 <i>Συμπίεση ιατρικών εικόνων</i>	<i>23</i>
2.2.3 <i>Παρεμβολή εικόνας (ψηφιακή μεγέθυνση)</i>	<i>24</i>
2.2.4 <i>Ανακατασκευή εικόνας.....</i>	<i>25</i>
3 Μηχανική μάθηση και τεχνητή νοημοσύνη: ο ρόλος των νευρωνικών δικτύων στην ταξινόμηση ιατρικών απεικονίσεων.....	26
3.1 <i>Τα νευρωνικά δίκτυα στην περίπτωση των ιατρικών ακτινολογικών εικόνων</i>	<i>27</i>
3.1.1 <i>Συνελκτικά νευρωνικά δίκτυα (convolutional neural networks).....</i>	<i>30</i>
3.2 <i>Είδη ταξινόμησης εικόνων ιατρικής απεικόνισης</i>	<i>38</i>
3.2.1 <i>Αλγόριθμοι ταξινόμησης.....</i>	<i>40</i>
3.2.2 <i>Μη εποπτευόμενη ταξινόμηση.....</i>	<i>45</i>

3.2.3	<i>Εποπτευόμενη ταξινόμηση</i>	47
3.3	Ταξινόμηση εικόνων ακτινολογικών δεδομένων	51
3.3.1	<i>Η ανάγκη μετατροπής από rgb σε grayscale των ιατρικών απεικονίσεων στην ταξινόμηση τους</i>	53
3.3.2	<i>Επεξεργασία των ιατρικών απεικονίσεων με αλγόριθμο ιστογραμμάτων κατευθυνόμενων κλίσεων (Histogram of oriented gradients , HOG)</i>	56
4	Ανάπτυξη αλγορίθμου για ταξινόμηση ακτινολογικών ιατρικών απεικονίσεων	63
4.1	Μετατροπή ακτινολογικών δεδομένων	63
4.1.1	<i>Διαδικασία testing του DICOM image classification – Συμπεράσματα</i>	75
4.2	Μελλοντικές προκλήσεις	78
	Βιβλιογραφία	80

Κατάλογος Εικόνων

Εικόνα Εξωφύλλου . [116] .

Εικόνα 1.1. Σχηματική απεικόνιση της δομής ενός νευρώνα [2].	1
Εικόνα 1.2. Σχηματική απεικόνιση φόρτισης νευρώνων [4].	2
Εικόνα 1.3. Σχηματική απεικόνιση ενός απλού μοντέλου νευρωνικού δικτύου [14].	7
Εικόνα 1.4. Γράφημα απλού back-propagation αλγορίθμου [15].	8
Εικόνα 1.5. Αρχιτεκτονική δομή ενός επαναλαμβανόμενου τεχνητού νευρωνικού δικτύου [15].	11
Εικόνα 2.1. Πληροφορίες ασθενούς όπως αποθηκεύονται στο DICOM [21].	19
Εικόνα 2.2. Σχηματική απεικόνιση των υπηρεσιών DICOM [21].	20
Εικόνα 2.3. Μεγέθυνση στα pixel εικόνας από ιατρική απεικόνιση αξονικής τομογραφίας [21].	21
Εικόνα 2.4. Αποθήκευση pixel με βάση το DICOM [21].	22
Εικόνα 2.5. Παράδειγμα αποθήκευσης σε bits πιο περίπλοκου στοιχείου [21].	23
Εικόνα 2.6. Τροποποίηση εικόνας αξονικής τομογραφίας ύστερα από συμπύεση [21].	24
Εικόνα 3.1. Εκμάθηση νευρωνικών δικτύων για ταξινόμηση ιατρικών απεικονίσεων [31].	28
Εικόνα 3.2. Ταξινόμηση, ανίχνευση και τμηματοποίηση με μηχανική όραση [31].	29
Εικόνα 3.3. Ανάλυση εικόνας σε τρία βασικά χρώματα [34].	31
Εικόνα 3.4. Κίνηση Kernel [34].	32
Εικόνα 3.5. Λειτουργία συνέλιξης σε μήτρα εικόνας $M \times N \times 3$ με πυρήνα $3 \times 3 \times 3$ [34].	33
Εικόνα 3.6. Διαδικασία ταξινόμησης ιατρικών εικόνων [37].	39
Εικόνα 3.7. Είδη ταξινόμησης εικόνων ιατρικής απεικόνισης [51].	45
Εικόνα 3.8. Μη εποπτευόμενη ταξινόμηση [52].	46
Εικόνα 4.1. HOG στοιχείο.	68
Εικόνα 4.2. Παράθυρο εντολών για εκπαίδευση νευρωνικού δικτύου.	69
Εικόνα 4.3. Εκπαίδευση νευρωνικών δικτύων εννιά επαναλήψεων.	70
Εικόνα 4.4. Πίνακας σύγκρισης.	70
Εικόνα 4.5. Απόδοση νευρωνικού δικτύου.	71
Εικόνα 4.6. Κελιά κατά την εκπαίδευση του νευρωνικού δικτύου.	72
Εικόνα 4.7. Διάγραμμα ROC νευρωνικού δικτύου.	72
Εικόνα 4.8. Παράθυρο εκτέλεσης εκπαίδευσης νευρωνικού δικτύου με πρόγραμμα Matlab και ακρίβεια.	73
Εικόνα 4.9. Θετικά αποδεχόμενη X-Ray dicom.	75

Εικόνα 4.10. Αρνητικά αποδεχόμενο DICOM.....	76
Εικόνα 4.11. Εικόνα υπερήχου που ταξινομήθηκε ως υπερηχογράφημα θετικά αποδεχόμενη.	76
Εικόνα 4.12. Αρνητικά αποδεχόμενη εικόνα υπερηχογραφήματος.....	77
Εικόνα 4.13. Ταξινόμηση εικόνας στην κατηγορία άλλο.	77

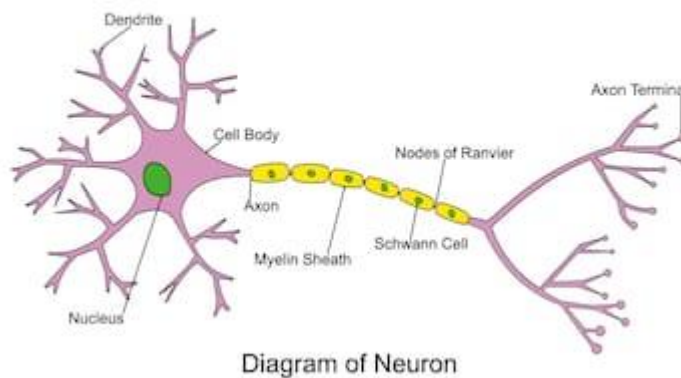
Περίληψη

Στόχος της παρούσας πτυχιακής εργασίας είναι η χρήση του αλγορίθμου HOG σε περιβάλλον Matlab με στόχο την ταξινόμηση ακτινολογικών εικόνων και στη συνέχεια η δημιουργία και εκπαίδευση νευρωνικού δικτύου για την ταξινόμηση αυτών των εικόνων. Για αυτό το σκοπό μελετήθηκε ένα σύνολο δεδομένων 150 εικόνων Dicom, οι οποίες ήταν στη μορφή jpeg. Σε περίπτωση που η εικόνα ήταν σε rgb μετατρέπονταν σε κλίμακα του γκρι αναπροσαρμόζοντας αυτές στις κανονικές διαστάσεις. Η εξαγωγή των χαρακτηριστικών των εικόνων βασισμένη στο HOG η οποία εφαρμόστηκε στις εικόνες χρησιμοποιώντας μια τιμή μεγέθους κελιού 128x128 είχε τη δυνατότητα να δώσει τιμές 324 χαρακτηριστικών από κάθε εικόνα. Με βάση αυτά τα χαρακτηριστικά δημιουργήθηκε ένα νευρωνικό δίκτυο βασισμένο σε patternnet με 20 hidden layers και επιλέγοντας ως τύπο δικτύου την backpropagation Levenberg-Marquardt. Επίσης, ορίστηκε η τυχαία κατανομή δεδομένων με αναλογίες 70% για εκπαίδευση, 15% για επικύρωση και 15% για δοκιμή των δεδομένων. Η απόδοση του νευρωνικού δικτύου μετρήθηκε με τη μέση τιμή σφάλματος.

1 Γενικά για τα νευρωνικά δίκτυα

1.1.1 Βιολογικά νευρωνικά δίκτυα

Τα χαρακτηριστικά του βιολογικού νευρικού δικτύου αποδίδονται στη δομή του και στον τρόπο λειτουργίας του. Η θεμελιώδης μονάδα ενός νευρωνικού δικτύου ονομάζεται νευρώνας ή νευρικό κύτταρο [1]. Στην Εικ. 1.1 απεικονίζεται η στοιχειώδης δομή ενός νευρώνα.



Εικόνα 1.1. Σχηματική απεικόνιση της δομής ενός νευρώνα [2].

Αυτός αποτελείται από ένα κυτταρικό σώμα ή σώματα (cell body) όπου βρίσκεται ο πυρήνας του κυττάρου. Οι νευρικές ίνες που ονομάζονται δενδρίτες (dendrite) συνδέονται με το σώμα των κυττάρων και χρησιμεύουν στο να λαμβάνουν σήματα από άλλους νευρώνες.

Σε επέκταση του κυτταρικού σώματος υπάρχει μια ενιαία μακριά ίνα που ονομάζεται άξονας, η οποία καταλήγει σε σκέλη δενδριτών που συνδέονται με άλλους νευρώνες. Ο άξονας ενός τυπικού νευρώνα οδηγεί σε μερικές άλλες χιλιάδες συνδέσεις που σχετίζονται με άλλους νευρώνες.

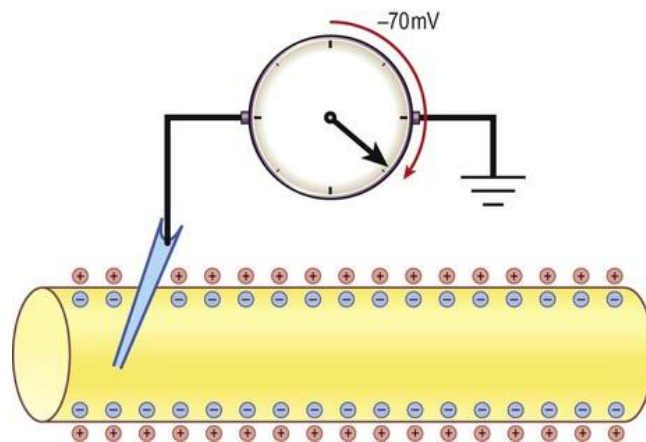
Η μετάδοση ενός σήματος από το ένα κύτταρο αποτελεί μια σύνθετη χημική διαδικασία στην οποία απελευθερώνονται συγκεκριμένες ουσίες 'μεταδότη' από την πλευρά της διασταύρωσης που στέλνει την πληροφορία. Η μετάδοση αυτής της πληροφορίας έχει σαν αποτέλεσμα να δημιουργείται μια διαφορά δυναμικού (σαν πυκνωτής) στο νευρώνα που δέχεται το σήμα. Στην περίπτωση λοιπόν που η διαφορά δυναμικού ξεπεράσει κάποια συγκεκριμένη τιμή αρχίζει η κινητικότητα του κυττάρου με την εμφάνιση κάποιων παλμών

και μπορεί να ειπωθεί ότι η υπάρχει κάποιο είδος ηλεκτρικής δραστηριότητας στο εσωτερικό ενός νευρώνα.

Οι δενδρίτες με τη σειρά τους χρησιμεύουν ως υποδοχείς για σήματα από άλλους νευρώνες, ενώ ο σκοπός του άξονα είναι η μετάδοση της παραγόμενης νευρικής δραστηριότητας σε άλλα νευρωνικά κύτταρα (ενδο-νευρώνες) ή σε μυϊκές ίνες (κινητικός νευρώνας). Ένας τρίτος τύπος νευρώνα, ο οποίος λαμβάνει πληροφορίες από τους μύς ή τα αισθητήρια όργανα, όπως το μάτι ή το αυτί, ονομάζεται υποδοχέας [1].

Αξίζει να αναφερθεί ότι το μέγεθος του κυτταρικού σώματος ενός τυπικού νευρώνα είναι περίπου 10-80 μικρόμετρα (μm) και οι δενδρίτες και οι άξονες έχουν διάμετρο της τάξεως των λίγων μικρομέτρων. Το συνολικό μήκος ενός νευρώνα κυμαίνεται από 0,01 mm για τους εσωτερικούς νευρώνες στον ανθρώπινο εγκέφαλο, μέχρι 1 m για τους νευρώνες στα άκρα [3].

Τέλος, το κυτταρικό σώμα ενός νευρώνα δρα ως ένα είδος αθροιστικής συσκευής που οφείλεται στο καθαρό αποπολωτικό αποτέλεσμα των σημάτων εισόδου του καθώς δέχεται θετικά και αρνητικά σήματα εισόδου (τιμές δυναμικού) των οποίων οι τιμές αθροίζονται (Εικ.1.2.).



Εικόνα 1.2. Σχηματική απεικόνιση φόρτισης νευρώνων [4].

Αν και όλοι οι νευρώνες λειτουργούν με τις ίδιες βασικές αρχές όπως περιγράφηκε παραπάνω, υπάρχουν διάφοροι τύποι νευρώνων οι οποίοι κατηγοριοποιούνται ανάλογα με το μέγεθος και το βαθμό διακλάδωσης του δενδριτικού τους στελέχους, το μήκος των αξόνων τους και άλλες δομικές λεπτομέρειες. Η πολυπλοκότητα του ανθρώπινου κεντρικού νευρικού

συστήματος οφείλεται στον τεράστιο αριθμό των νευρώνων και τις αμοιβαίες τους συνδέσεις.

1.1.2 Τεχνητά νευρωνικά δίκτυα

Τα τεχνητά νευρωνικά δίκτυα (artificial neural networks, ANN) είναι, όπως υποδηλώνει το όνομά τους, υπολογιστικά δίκτυα που προσπαθούν να προσομοιώσουν, κατά μείζονα τρόπο, τη διαδικασία λήψης αποφάσεων που λαμβάνει χώρα στα δίκτυα των νευρικών κυττάρων (νευρώνες) του βιολογικού (ανθρώπινου ή ζωικού) κεντρικού νευρικού συστήματος. Αυτή η προσομοίωση είναι μια προσομοίωση νευρώνα ανά νευρώνα, στοιχείο ανά στοιχείο και 'δανείζεται' στοιχεία από τη νευροφυσιολογική γνώση των βιολογικών νευρώνων και τέτοιων δικτύων.

Πιο αναλυτικά, ένα νευρωνικό δίκτυο είναι ένα διασυνδεδεμένο συγκρότημα απλών στοιχείων επεξεργασίας, μονάδων ή κόμβων, των οποίων η λειτουργικότητα βασίζεται εν μέρει στη λειτουργία των νευρώνων των ζώων.

Ένα νευρωνικό δίκτυο, προσομοιάζοντας ένα βιολογικό νευρικό δίκτυο, αποτελεί ουσιαστικά στην πραγματικότητα μια νέα αρχιτεκτονική υπολογιστών και μια νέα αρχιτεκτονική αλγορίθμων, σε σχέση με τους συμβατικούς υπολογιστές. Αυτό επιτρέπει τη χρήση πολύ απλών υπολογιστικών λειτουργιών (προσθήκες, πολλαπλασιασμός και θεμελιώδη λογικά στοιχεία) για την επίλυση πολύπλοκων μαθηματικών προβλημάτων, μη γραμμικών ή στοχαστικών προβλημάτων.

Ένας συμβατικός αλγόριθμος εφαρμόζει σύνθετα σύνολα εξισώσεων τα οποία θα ισχύουν μόνο για ένα συγκεκριμένο πρόβλημα και ακριβώς και μόνο για αυτό. Τα τεχνητά νευρωνικά δίκτυα διαφέρουν από τις συμβατικές (ψηφιακές ή αναλογικές) υπολογιστικές μηχανές που χρησιμεύουν για την αντικατάσταση, ενίσχυση ή επιτάχυνση των διαδικασιών υπολογισμού του ανθρώπινου εγκεφάλου. Ένα νευρωνικό δίκτυο θα είναι (α) υπολογιστικά και αλγοριθμικά πολύ απλό και (β) θα έχει μια αυτορρυθμιζόμενη δυνατότητα που θα του επιτρέψει να διατηρήσει ένα ευρύ φάσμα προβλημάτων.

Για παράδειγμα, εάν μια μύγα σε ένα σπίτι αποφεύγει ένα εμπόδιο ή αν ένα ποντίκι αποφεύγει μια γάτα, αυτό σίγουρα δεν υπακούει σε διαφορικές εξισώσεις, ούτε χρησιμοποιεί πολύπλοκους αλγόριθμους αναγνώρισης προτύπων. Ο βιολογικός εγκέφαλος είναι πολύ απλός, αλλά απασχολεί λίγα βασικά νευρωνικά κύτταρα τα οποία βασικά υπακούουν σε μια

βασική δομή. Η λύση του τεχνητού νευρικού δικτύου στοχεύει επίσης σε αυτήν την απλότητα.

Οι βασικές αρχές των τεχνητών νευρωνικών δικτύων διατυπώθηκαν για πρώτη φορά από τους McCulloch και Pitts το 1943, με τις εξής παραδοχές:

- i) Η δραστηριότητα ενός δικτύου είναι όλα ή τίποτα.
- ii) Ένας ορισμένος σταθερός αριθμός συνάψεων (δηλαδή μεταφορά σήματος) μεγαλύτερων από 1 πρέπει να μεταφέρεται μέσω ενός 'ερεθίσματος' σε ένα νευρικό δίκτυο.
- iii) Η μόνη σημαντική καθυστέρηση στο τεχνητό νευρικό σύστημα είναι η συναπτική καθυστέρηση.
- iv) Οποιασδήποτε ανασταλτική μετάδοση του σήματος αποτρέπει απολύτως την διέγερση του νευρικού δικτύου εκείνη τη στιγμή.
- v) Η διάρθρωση του δικτύου διασύνδεσης δεν αλλάζει με την πάροδο του χρόνου.

Επομένως με βάση την πρώτη υπόθεση, ο νευρώνας είναι ένα δυαδικό στοιχείο. Αυτές όμως είναι οι πρώτες αρχές πάνω στις οποίες βασίστηκε ο σχεδιασμός τους καθώς σήμερα τα τεχνητά δίκτυα έχουν εξελιχθεί.

1.1.2.1 Σύντομη ιστορική ανασκόπηση

Η έρευνα για τα τεχνητά νευρωνικά δίκτυα μπορεί να διαχωριστεί σε τρεις βασικές περιόδους. Η έναρξής της έγινε τη δεκαετία του 1940 από τους McCulloch και Pitts [5]. Στη συνέχεια το 1960 με το θεώρημα της σύγκλισης οι Rosenblatt [6] και έπειτα οι Minsky και Papert [7] διατυπώνουν τους περιορισμούς ενός απλού perceptron. Αξίζει να σημειωθεί ότι ο νευρώνας perceptron (ή αντίληπτρο) εφευρέθηκε το 1957 και αποτελεί ένα είδος τεχνητού νευρωνικού δικτύου το οποίο μπορεί να συγκαταλεχθεί στην κατηγορία των εμπροσθοτροφοδοτούμενων (feedforward) νευρωνικών δικτύων σαν ένας γραμμικός ταξινομητής [8]. Τα αποτελέσματα των Minsky και Papert [7] ενθουσίασαν τους ερευνητές και ειδικότερα την κοινότητα των υπολογιστών.

Ωστόσο υπήρξε ένα ερευνητικό κενό σχεδόν 20 χρόνων, όπου στη συνέχεια κατά τη δεκαετία του 1980 τα τεχνητά νευρωνικά δίκτυα λαμβάνουν σημαντικό ανανεωμένο ενδιαφέρον. Η αναζωπύρωση του τομέα και οι κύριες εξελίξεις οφείλονται στις ενέργειες του Hopfield [9] το 1982 και στην εκμάθηση του αλγορίθμου της οπίσθιας διάδοσης για perceptrons πολλαπλών στρώσεων (πολλαπλών στρώσεων feedforward δίκτυα) που

προτάθηκαν για πρώτη φορά από τον Werbos [10] και μελετήθηκαν στη συνέχεια το 1986 από τους Rumelhart *et al.* [11]. Ωστόσο καθοριστικό ρόλο στις εξελίξεις των τεχνητών νευρωνικών δικτύων έπαιξαν οι Anderson and Rosenfeldl [12].

Τα τελευταία 15 χρόνια παρατηρείται λοιπόν μια σημαντική αύξηση του ενδιαφέροντος για τα τεχνητά νευρωνικά δίκτυα, καθώς αυτά παρουσιάζουν ικανότητες που λείπουν από άλλες υπολογιστικές μηχανές. Συγκεκριμένα, είναι αποδοτικά για εργασίες που περιλαμβάνουν ελλιπή σύνολα δεδομένων, ασαφείς ή ελλιπείς πληροφορίες και για ιδιαίτερα περίπλοκα και ασαφή προβλήματα, όπου οι άνθρωποι συνήθως αποφασίζουν σε μια συνειδητή βάση. Δηλαδή, για παράδειγμα τα τεχνητά νευρωνικά δίκτυα μπορούν να μάθουν από παραδείγματα και είναι σε θέση να αντιμετωπίσουν μη γραμμικά προβλήματα. Επιπλέον, πλέον παρουσιάζουν ανθεκτικότητα και ανοχή σφάλματος.

Σε περιπτώσεις όμως που απαιτείται υψηλή ακρίβεια στη λογική και στην αριθμητική αυτά δεν μπορούν να χειριστούν αποτελεσματικά τέτοιες καταστάσεις. Μέχρι σήμερα τα τεχνητά νευρωνικά δίκτυα έχουν χρησιμοποιηθεί επιτυχώς για διάφορες λειτουργίες [13]:

1. **Λειτουργία προσέγγισης (function approximation):** σε αυτήν την περίπτωση καθορίζεται η χαρτογράφηση μιας πολλαπλής εισόδου σε μία μόνο έξοδο. Σε αντίθεση με τις περισσότερες στατιστικές τεχνικές, αυτό μπορεί να γίνει με προσαρμοστική εκτίμηση παραμέτρων χωρίς τη χρήση κάποιου μοντέλου.
2. **Σύνδεση μοτίβων και αναγνώριση προτύπων:** αυτό αποτελεί ένα πρόβλημα ταξινόμησης προτύπων. Τα τεχνητά νευρωνικά δίκτυα μπορούν να χρησιμοποιηθούν αποτελεσματικά για την επίλυση δύσκολων προβλημάτων σε αυτόν τον τομέα, για παράδειγμα στην αναγνώριση ήχου, εικόνας ή βίντεο. Επιπλέον αυτή η διαδικασία μπορεί ακόμη να γίνει χωρίς έναν ορισμό του σχεδίου. Σε τέτοιες περιπτώσεις το δίκτυο μαθαίνει να προσδιορίζει εντελώς νέα πρότυπα.
3. **Συνεργατικές μνήμες.** αυτό αποτελεί πρόβλημα της ανάκλησης ενός μοτίβου όταν δίνεται μόνο ένα υποσύνολο. Σε τέτοιες εφαρμογές οι δομές δικτύου που χρησιμοποιούνται είναι συνήθως περίπλοκες, αποτελούμενες από πολλούς δυναμικούς νευρώνες που αλληλεπιδρούν.
4. **Δημιουργία νέων ουσιαστικών προτύπων:** αυτό το γενικό πεδίο εφαρμογής είναι σχετικά νέο. Με τις κατάλληλες νευρωνικές δομές μπορούν να υπάρξουν στοιχεία δημιουργικότητας.

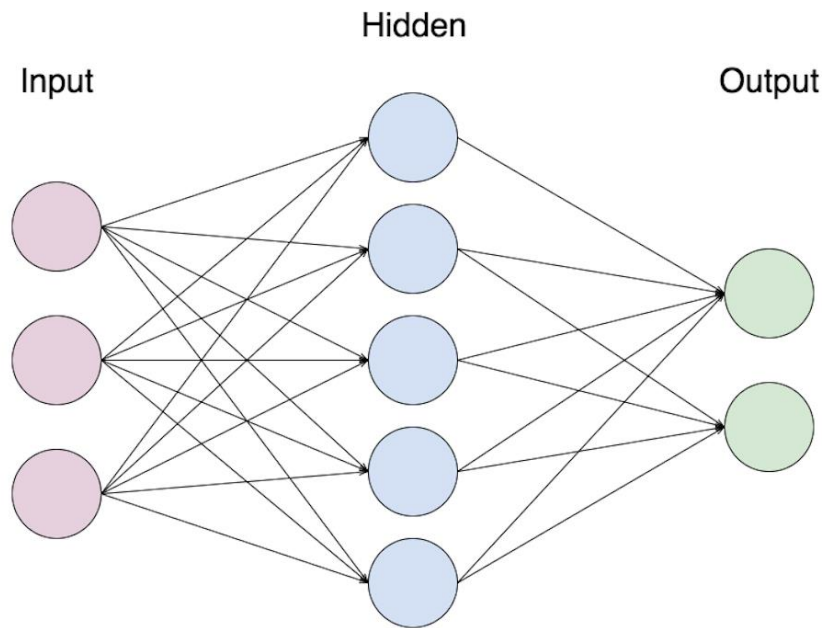
Τα τεχνητά νευρωνικά δίκτυα έχουν εφαρμοστεί επιτυχώς σε διάφορους τομείς των μαθηματικών, της μηχανικής, της ιατρικής, της οικονομίας, της μετεωρολογίας, της ψυχολογίας, της νευρολογίας και σε πολλούς άλλους. Ορισμένες από τις σημαντικότερες είναι στην αναγνώριση του ήχου και του λόγου, στην ανάλυση των ηλεκτρομυογραφημάτων και άλλων ιατρικών υπογραφών, στον εντοπισμό στρατιωτικών στόχων και στην ταυτοποίηση των εκρηκτικών στις βολίτσες των επιβατών.

Χρησιμοποιούνται επίσης για την πρόβλεψη των καιρικών συνθηκών και των τάσεων της αγοράς, για την πρόβλεψη τοποθεσιών εξερεύνησης ορυκτών, για πρόβλεψη ηλεκτρικού και θερμικού φορτίου και για προσαρμοστικό και ρομποτικό έλεγχο. Επιπλέον αυτά χρησιμοποιούνται για τον έλεγχο της διαδικασίας, επειδή μπορούν να δημιουργήσουν πρότυπα μοντέλα της διαδικασίας από πολυδιάστατα δεδομένα συλλεχθέντα συστηματικά από αισθητήρες.

1.1.2.2 Βασικές αρχές τεχνητών νευρωνικών δικτύων

Ένα τεχνητό νευρωνικό δίκτυο, όπως προαναφέρθηκε δανείζεται στοιχεία από τη νευροφυσιολογική γνώση των βιολογικών νευρώνων διαφέροντας κατά αυτόν τον τρόπο από τους συμβατικούς υπολογιστές που χρησιμεύουν για την αντικατάσταση, ενίσχυση ή επιτάχυνση του ανθρώπινου εγκεφαλικού υπολογισμού, ανεξάρτητα από την οργάνωση των υπολογιστικών στοιχείων και τη δικτύωσή τους. Μέχρι σήμερα ωστόσο η προσομοίωση που παρέχεται από τα νευρωνικά δίκτυα είναι πολύ ‘χονδροειδής’.

Η ικανότητα που έχουν τα τεχνητά νευρωνικά δίκτυα να μάθουν τόσο γρήγορα είναι αυτή που τα καθιστά τόσο ισχυρά και χρήσιμα για μια ποικιλία εφαρμογών. Αλλά πώς μαθαίνουν; Η πληροφορία ρέει μέσω ενός νευρωνικού δικτύου με δύο διαφορετικούς τρόπους. Όταν το μοντέλο μαθαίνει (που εκπαιδεύεται) ή λειτουργεί κανονικά (αφού εκπαιδεύεται είτε χρησιμοποιείται είτε δοκιμάζεται), τα πρότυπα πληροφοριών από το σύνολο δεδομένων τροφοδοτούνται στο δίκτυο μέσω των νευρώνων εισόδου, οι οποίοι ενεργοποιούν τα στρώματα κρυμμένων νευρώνων και αυτά με τη σειρά τους φθάνουν στους νευρώνες εξόδου. Αυτό ονομάζεται δίκτυο feedforward (όπως έχει προαναφερθεί) (Εικ.1.3) [14].



Εικόνα 1.3. Σχηματική απεικόνιση ενός απλού μοντέλου νευρωνικού δικτύου [14].

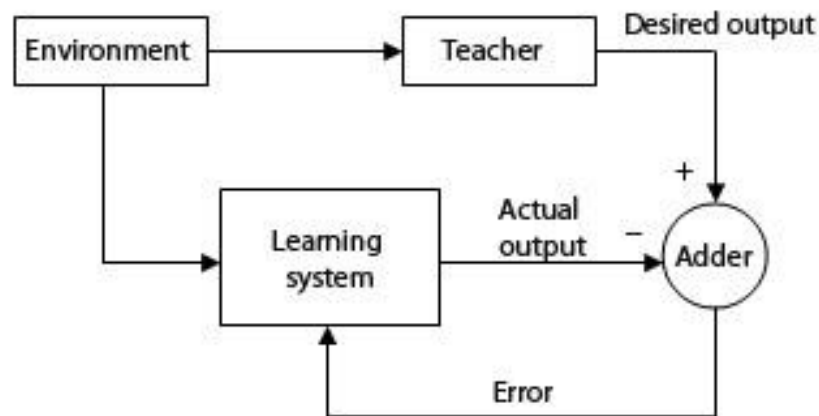
Χαρακτηριστικό είναι ότι δε βρίσκονται σε λειτουργία όλη την ώρα, όλοι οι νευρώνες. Κάθε νευρώνας λαμβάνει εισροές από τους νευρώνες στα αριστερά του και οι είσοδοι πολλαπλασιάζονται με τα βάρη των συνδέσεων που ταξιδεύουν. Κάθε νευρώνας προσθέτει όλες τις εισόδους που λαμβάνει με αυτόν τον τρόπο και (αυτό είναι το απλούστερο νευρωνικό δίκτυο). Αν το άθροισμα είναι πάνω από μια ορισμένη τιμή, ο νευρώνας "ξεκινάει να λειτουργεί" και ενεργοποιεί τους νευρώνες στους οποίους συνδέεται (τους νευρώνες στα δεξιά του).

Για ένα τεχνητό νευρωνικό δίκτυο που πρέπει να μάθει, πρέπει επίσης να μάθει τι έχει κάνει λάθος και τι κάνει σωστά. Αυτή η διαδικασία ονομάζεται ανατροφοδότηση. Σε αυτή λοιπόν τη διαδικασία φαίνονται οι ομοιότητες με τον ανθρώπινο εγκέφαλο. Για παράδειγμα, εάν κάποιος μαθαίνει να παίζει ένα παιχνίδι όπως το τένις, μαθαίνει επίσης ότι αν χτυπήσει πολύ καλά τη μπάλα θα βγει από το γήπεδο και θα χάσει το σημείο ή αν δεν χτυπήσει την μπάλα αρκετά σκληρά δεν θα πάει πέρα από το δίχτυ, αλλά αν τη χτυπήσετε τέλεια θα πάει στην άλλη πλευρά στο γήπεδο και θα μπορούσε να κερδίσει ένα σημείο.

Αυτό είναι ένα κλασικό παράδειγμα ανατροφοδότησης όπου ένα τεχνητό νευρωνικό δίκτυο χάνει το σημείο ή δυνητικά κερδίζει ένα σημείο. Αυτός είναι ο τρόπος με τον οποίο ο άνθρωπος μαθαίνει τι κάνει σωστά ή λάθος και αυτό είναι που πρέπει να μάθει ένα νευρωνικό δίκτυο [14].

Τα νευρωνικά δίκτυα μαθαίνουν τα πράγματα με τον ίδιο ακριβώς τρόπο όπως ο εγκέφαλος, συνήθως με μια διαδικασία ανάδρασης που ονομάζεται back-propagation (μερικές φορές συντομεύεται ως "backprop"). Αυτός είναι ο τρόπος όπου η έξοδος του δικτύου συγκρίνεται με την έξοδο που προοριζόταν να παραχθεί.

Έτσι χρησιμοποιώντας τη διαφορά μεταξύ των εξόδων τροποποιούνται τα βάρη των συνδέσεων μεταξύ των νευρώνων στο δίκτυο σε μια διαδρομή από τις μονάδες εξόδου μέσω των κρυμμένων νευρώνων μέχρι τους νευρώνες εισόδου που πηγαίνουν προς τα πίσω. Με την πάροδο του χρόνου, ο πολλαπλασιασμός της οπίσθιας διάδοσης αναγκάζει το δίκτυο να μάθει κάνοντας μικρότερο το χάσμα ανάμεσα στην έξοδο και την προοριζόμενη έξοδο μέχρι το σημείο όπου τα δύο ταιριάζουν ακριβώς, έτσι ώστε το νευρωνικό δίκτυο να μάθει τη σωστή έξοδο.



Εικόνα 1.4. Γράφημα απλού back-propagation αλγορίθμου [15].

Για παράδειγμα, ένας άνθρωπος περπατάει σε ένα πεζοδρόμιο και βλέπει μια λάμπα, της οποίας τη θέση δεν την είχε δει ποτέ πριν και έτσι περπατάει κατευθείαν επάνω της με αποτέλεσμα να προκαλέσει το σπάσιμό της. Την επόμενη φορά που θα δει μια λάμπα στην ίδια θέση θα παρεκκλίνει μερικά εκατοστά στο πλάι για να συνεχίσει να περπατάει.

Αυτή τη φορά χτυπάει ο ώμος τη θέση του λαμπτήρα. Την τρίτη φορά παρεκκλίνει περισσότερο για να βεβαιωθεί ότι δεν χτυπάει τη θέση του λαμπτήρα. Μόνο που τώρα έχει γίνει κάτι επίσης διαφορετικό, καθώς τώρα ο άνθρωπος περπατάει κατευθείαν στο μονοπάτι ενός γραμματοκιβωτίου που δεν το έχει ξαναδεί. Ο άνθρωπος περπατάει σε αυτό και η όλη διαδικασία συμβαίνει και πάλι.

Αυτή είναι πραγματικά η διαδικασία που κάνει η backpropagation. Ένα τεχνητό νευρωνικό δίκτυο λαμβάνει ένα πλήθος παραδειγμάτων και στη συνέχεια προσπαθεί να πάρει την ίδια απάντηση με το παράδειγμα που δίνεται. Όταν είναι λάθος, υπολογίζεται ένα σφάλμα και οι τιμές σε κάθε νευρώνα και σύναψη πολλαπλασιάζονται προς τα πίσω μέσω του τεχνητού νευρωνικού δικτύου για την επόμενη φορά. Αυτή η διαδικασία παίρνει πολλά παραδείγματα και για εφαρμογές πραγματικού κόσμου, ο αριθμός των παραδειγμάτων μπορεί να είναι εκατομμύρια [15].

Ουσιαστικά λοιπόν ένας τεχνητός νευρώνας είναι μια μαθηματική λειτουργία που θεωρείται ως μοντέλο βιολογικών νευρώνων, ενός νευρικού δικτύου. Οι τεχνητοί νευρώνες είναι στοιχειώδεις μονάδες σε ένα τεχνητό νευρωνικό δίκτυο. Ο τεχνητός νευρώνας λαμβάνει μία ή περισσότερες εισόδους (που αντιπροσωπεύουν τους δενδρίτες) και τους αθροίζει για να παράγει μια έξοδο.

Ωστόσο υπάρχει πάντα το ερώτημα πώς γνωρίζουμε πόσους νευρώνες πρέπει να χρησιμοποιηθούν. Υπάρχουν τα επίπεδα τα οποία είναι απλά σύνολα νευρώνων με ένα στρώμα εισόδου που είναι τα δεδομένα τα οποία παρέχονται στο τεχνητό νευρωνικό δίκτυο. Επίσης υπάρχουν τα 'κρυμμένα' στρώματα, όπου συμβαίνουν διάφορες διεργασίες και τέλος, υπάρχουν τα στρώματα εξόδου, όπου τοποθετούνται οι τελικοί υπολογισμοί του δικτύου για να χρησιμοποιηθούν (Εικ.1.3).

Τα ίδια τα επίπεδα είναι απλά σύνολα νευρώνων. Κατά τα πρώτα χρόνια ανάπτυξης των πολλαπλών επιπέδων perceptrons, θεωρήθηκε ότι ένα στρώμα εισόδου, ένα κρυφό στρώμα και ένα στρώμα εξόδου είναι αρκετά. Από αρκετούς ερευνητές θεωρήθηκε ότι με δεδομένους ορισμένους αριθμούς, χρειάζονται μόνο ένα σύνολο υπολογισμών και, στη συνέχεια, έχουν έξοδο.

Ωστόσο εάν το τεχνητό νευρωνικό δίκτυο δεν υπολόγισε τη σωστή τιμή, προστίθενται στο κρυφό επίπεδο περισσότεροι νευρώνες. Τελικά, με αυτή τη διαδικασία ακριβώς επιτυγχάνεται μια γραμμική χαρτογράφηση από κάθε είσοδο στην έξοδο. Με άλλα λόγια μια συγκεκριμένη είσοδος οδηγεί πάντα σε μια συγκεκριμένη έξοδο.

Ωστόσο στα τεχνητά νευρωνικά δίκτυα που έχουν αναπτυχθεί σήμερα, υπάρχουν περισσότερα από ένα κρυμμένα επίπεδα. Αυτός είναι ένας από τους λόγους που αυτά έχουν βελτιωθεί και επειδή χρειάζονται εκατοντάδες κόμβοι με δεκάδες, αν όχι περισσότερα

επίπεδα. Αυτή η διαδικασία οδηγεί σε μια τεράστια ποσότητα μεταβλητών που πρέπει να παρακολουθούνται κάθε φορά.

Επίσης σήμερα οι πρόοδοι στον παράλληλο προγραμματισμό επιτρέπουν επίσης να τρέχουν σε ‘παρτίδες’ ακόμα μεγαλύτερα τεχνητά νευρωνικά δίκτυα. Τα τεχνητά νευρωνικά δίκτυα είναι τώρα τόσο μεγάλα ώστε δεν μπορούν να τρέχουν σε μια μόνο περίοδο, η οποία είναι μια επανάληψη σε όλο το δίκτυο. Πρέπει η όλη διαδικασία να γίνεται σε ‘παρτίδες που είναι απλώς υποσύνολα ολόκληρου του δικτύου και μόλις ολοκληρώνεται η μια περίοδος, τότε να εφαρμόζεται η backpropagation [15].

Αξίζει να αναφερθεί ότι στα νευρωνικά δίκτυα υπάρχει η εποπτευόμενη μάθηση η οποία είναι ένας τύπος τεχνητού νευρικού δικτύου. Η εποπτευόμενη μάθηση συμβαίνει όταν δεδομένα λαμβάνουν τυχαίες τιμές στις τιμές των βαρών, ώστε να μάθει το δίκτυο. Στην εποπτευόμενη μάθηση, κάθε παράδειγμα είναι ένα ζεύγος που αποτελείται από ένα αντικείμενο εισόδου (αυτό είναι συνήθως ένα διάγραμμα) και μια επιθυμητή τιμή εξόδου, αυτό ονομάζεται εποπτικό σήμα.

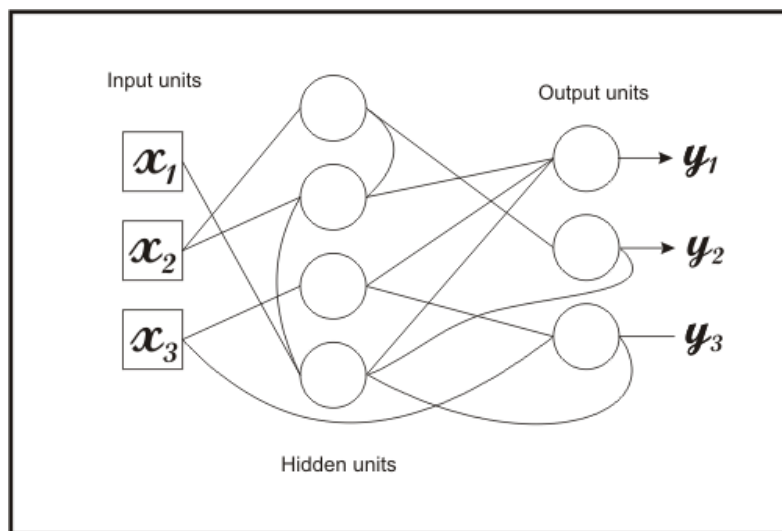
Ένα παράδειγμα ενός προβλήματος εποπτευόμενης μάθησης είναι η κατασκευή αυτόματων αυτοκινήτων, επειδή λαμβάνονται πολλά δεδομένα με τιμές από το LIDAR και τις κάμερες και στη συνέχεια πρέπει να γίνουν υπολογισμοί μηχανικής μάθησης από αυτά.

Κατά τη μη εποπτευόμενη μάθηση για την εξαγωγή συμπερασμάτων χρησιμοποιείται ένας τύπος αλγόριθμου μάθησης μηχανής όπου χρησιμοποιούνται στοιχεία από σύνολα δεδομένων που αποτελούνται από δεδομένα εισόδου χωρίς επισημασμένες απαντήσεις. Η πιο κοινή μέθοδος μάθησης χωρίς επίβλεψη είναι η ανάλυση συμπλέγματος, η οποία χρησιμοποιείται για διερευνητική ανάλυση δεδομένων για να εντοπίσει κρυμμένα μοτίβα ή ομαδοποίηση δεδομένων.

1.1.2.3 Αρχιτεκτονική τεχνητών νευρωνικών δικτύων

Σήμερα υπάρχει μια πληθώρα διαφορετικών αρχιτεκτονικών τεχνητών νευρωνικών δικτύων. Με βάση την κλασική δομή ο κάθε νευρώνας συνδέεται με κάθε άλλο νευρώνα στο επόμενο επίπεδο. Αυτά ονομάζονται ειδικά τροφοδοτούμενα τεχνητά νευρωνικά δίκτυα (αν και γενικά όλα είναι feedforward). Γενικότερα συνδέοντας νευρώνες με άλλους νευρώνες υπό συγκεκριμένα πρότυπα, λαμβάνονται καλύτερα αποτελέσματα σε συγκεκριμένα σενάρια. Μερικά παραδείγματα αρχιτεκτονικής τεχνητών νευρωνικών δικτύων είναι τα ακόλουθα [14]:

Επαναλαμβανόμενα νευρωνικά δίκτυα (recurrent neural networks): τα επαναλαμβανόμενα νευρωνικά δίκτυα δημιουργήθηκαν για να αντιμετωπίσουν το ελάττωμα των τεχνητών νευρωνικών δικτύων που δεν λάμβαναν αποφάσεις με βάση προηγούμενες γνώσεις. Ένα τυπικό τεχνητό νευρωνικό δίκτυο είχε μάθει να λαμβάνει αποφάσεις με βάση το πλαίσιο της κατάρτισης, αλλά μόλις τις χρησιμοποιεί, οι αποφάσεις γίνονται ανεξάρτητες η μια από την άλλη (Εικ.1.5).



Εικόνα 1.5. Αρχιτεκτονική δομή ενός επαναλαμβανόμενου τεχνητού νευρωνικού δικτύου [15].

Για παράδειγμα, μια τέτοια αρχιτεκτονική δομή θα μπορούσε να χρησιμοποιηθεί σε ένα παιχνίδι blackjack. Εάν έδωσε κάποιος 4 και 5 για να ξεκινήσει, ξέρει ότι 2 κάρτες μικρότερου αριθμού δεν πρέπει να πέσουν στο τραπέζι. Αυτές οι πληροφορίες θα μπορούσαν να βοηθήσουν για να προσδιοριστεί αν πρέπει να ‘χτυπήσει’ ή όχι. Η επαναλαμβανόμενη αρχιτεκτονική δομή είναι πολύ χρήσιμη στην επεξεργασία της φυσικής γλώσσας, καθώς οι προηγούμενες λέξεις ή χαρακτήρες είναι χρήσιμα για την κατανόηση του πλαισίου μιας άλλης λέξης.

Υπάρχουν πολλές διαφορετικές υλοποιήσεις, αλλά η πρόθεση είναι πάντα η ίδια, δηλαδή η διατήρηση της πληροφορίας. Αυτή μπορεί να επιτευχθεί μέσω της ύπαρξης αμφίδρομων RNNs, ή μπορεί να εφαρμοστεί ένα επαναλαμβανόμενο κρυφό επίπεδο που τροποποιείται με κάθε feedforward.

Συνεργατικά νευρωνικά δίκτυα (convolutional neural networks): τα συνεργατικά νευρωνικά δίκτυα, που μερικές φορές ονομάζονται LeNets (πήραν το όνομά τους από τον Yann LeCun), είναι τεχνητά νευρωνικά δίκτυα όπου οι συνδέσεις μεταξύ των επιπέδων

φαίνεται να είναι κάπως αυθαίρετες. Ωστόσο, οι συνάψεις πρέπει να ρυθμιστούν με τέτοιο τρόπο ώστε να βοηθήσουν στη μείωση του αριθμού των παραμέτρων που πρέπει να βελτιστοποιηθούν. Αυτό γίνεται κάνοντας μια συγκεκριμένη συμμετρία στον τρόπο με τον οποίο συνδέονται οι νευρώνες και έτσι ουσιαστικά αυτοί μπορεί να "επαναχρησιμοποιηθούν" ώστε να έχουν ταυτόσημα αντίγραφα χωρίς απαραίτητα να χρειάζονται τον ίδιο αριθμό συνάψεων.

Αυτού του είδους τα δίκτυα χρησιμοποιούνται συνήθως για να δουλεύουν με εικόνες χάρη στην ικανότητά τους να αναγνωρίζουν τα σχέδια στα περιβάλλοντα εικονοστοιχεία. Όταν εξετάζετε κάθε μεμονωμένο εικονοστοιχείο υπάρχουν περιττές πληροφορίες που περιέχονται σε σύγκριση με τα εικονοστοιχεία που το περιβάλλουν. Χάρη στις συμμετρικές ιδιότητες τους κάποιες από αυτές τις πληροφορίες μπορούν να συμπιεστούν.

1.1.3 Σύγκριση βιολογικών και τεχνητών νευρωνικών δικτύων

Στην παρούσα υποενότητα γίνεται σύγκριση μεταξύ των βιολογικών και τεχνητών νευρωνικών δικτύων, με βάση κάποια χαρακτηριστικά τους:

Ταχύτητα: τα βιολογικά νευρωνικά δίκτυα είναι αργά στην επεξεργασία πληροφοριών. Στην περίπτωση των πιο προηγμένων υπολογιστών ο χρόνος κύκλου που αντιστοιχεί στην εκτέλεση ενός σταδίου ενός προγράμματος στην κεντρική μονάδα επεξεργασίας χρειάζεται λίγα μόνο νανοδευτερόλεπτα. Ο χρόνος κύκλου που αντιστοιχεί σε ένα νευρικό συμβάν που προκαλείται από ένα εξωτερικό ερέθισμα εμφανίζεται σε χιλιοστά του δευτερολέπτου. Έτσι ένας υπολογιστής μπορεί να επεξεργαστεί πληροφορίες σχεδόν εκατομμύρια φορές πιο γρήγορα.

Επεξεργασία: τα νευρωνικά δίκτυα μπορούν να εκτελέσουν μαζικά παράλληλες λειτουργίες καθώς τα περισσότερα προγράμματα έχουν μεγάλο αριθμό οδηγιών και λειτουργούν με μια διαδοχική σειρά λειτουργιών. Από την άλλη πλευρά, ο βιολογικός εγκέφαλος λειτουργεί με μαζικά παράλληλες λειτουργίες, καθεμία από τις οποίες έχει συγκριτικά λιγότερα βήματα. Αυτό εξηγεί την ανώτερη απόδοση του ανθρώπου σε ορισμένες περιπτώσεις επεξεργασίας πληροφοριών, παρά το γεγονός ότι σε σύγκριση με τους υπολογιστές επεξεργάζεται πιο αργά τις πληροφορίες.

Μέγεθος και πολυπλοκότητα: τα νευρωνικά δίκτυα έχουν μεγάλο αριθμό υπολογιστικών στοιχείων, και ο υπολογισμός δεν περιορίζεται μέσα στους νευρώνες. Ο αριθμός των νευρώνων στον εγκέφαλο εκτιμάται ότι είναι περίπου 10^{11} και ο συνολικός αριθμός των

διασυνδέσεων να είναι περίπου 10^{15} . Λόγω αυτού του μεγέθους και της πολυπλοκότητας των συνδέσεων μπορεί ο βιολογικός εγκέφαλος να έχει τη δύναμη της εκτέλεσης σύνθετων εργασιών που δεν μπορούν να υλοποιηθούν σε έναν υπολογιστή. Η πολυπλοκότητα του εγκεφάλου ενισχύεται περαιτέρω από το γεγονός ότι λαμβάνει χώρα ο υπολογισμός όχι μόνο μέσα στο σώμα του κυττάρου, αλλά και έξω από τους δενδρίτες και τις συνάψεις τους.

Αποθήκευση: η δύναμη της αποθήκευσης των βιολογικών νευρωνικών δικτύων βρίσκεται στις διασυνδέσεις τους. Σε έναν υπολογιστή, οι πληροφορίες αποθηκεύονται στη μνήμη χωρίς ωστόσο να υπάρχει αναπροσαρμογή της μνήμης. Αντίθετα, σε ένα βιολογικό νευρικό δίκτυο όταν προστίθενται νέες πληροφορίες οι δυνάμεις της εκάστοτε διασύνδεσης αναπροσαρμόζονται και οι παλιές πληροφορίες παραμένουν. Έτσι οι πληροφορίες στον εγκέφαλο είναι προσαρμόσιμες, ενώ στον υπολογιστή είναι αυστηρά αντικαταστάσιμες.

Ανοχή σφάλματος: τα βιολογικά νευρωνικά δίκτυα εμφανίζουν ανοχή σφάλματος από τη στιγμή που οι πληροφορίες διανέμονται στις συνδέσεις σε όλο το δίκτυο. Ακόμη και αν συνδέονται μερικές συνδέσεις ή κάποιοι νευρώνες δεν είναι σε λειτουργία, οι πληροφορίες εξακολουθούν να διατηρούνται λόγω των κατανεμημένων εξαιτίας της φύσης κωδικοποιημένων πληροφοριών. Σε αντίθεση, οι υπολογιστές εγγενώς δεν είναι ανεκτικοί σε σφάλματα, με την έννοια ότι οι πληροφορίες που έχουν αλλοιωθεί στη μνήμη δεν μπορούν να ανακτηθούν.

Μηχανισμός ελέγχου: Στην περίπτωση του εγκεφάλου δεν υπάρχει κεντρικός έλεγχος για την επεξεργασία των πληροφοριών. Αντίθετα σε έναν υπολογιστή υπάρχει μια μονάδα ελέγχου η οποία παρακολουθεί όλες τις δραστηριότητες του υπολογιστή. Επίσης σε ένα νευρωνικό δίκτυο κάθε νευρώνας δρα με βάση τις πληροφορίες που είναι διαθέσιμες τοπικά και μεταδίδει τις πληροφορίες στους νευρώνες που συνδέονται με αυτόν, με αποτέλεσμα να μην υπάρχει συγκεκριμένος μηχανισμός ελέγχου.

Ενώ στην περίπτωση αναγνώρισης προτύπων, το ανθρώπινο σύστημα επεξεργασίας πληροφοριών είναι ανώτερο σε σύγκριση με το συμβατικό υπολογιστή, λόγω της βασικής δομής και λειτουργίας του βιολογικού νευρικού συστήματος, υπάρχει η δυνατότητα μερικά από τα χαρακτηριστικά του βιολογικού νευρωνικού δικτύου να ακολουθηθούν χρησιμοποιώντας ένα τεχνητό δίκτυο που αποτελείται από βασικές μονάδες υπολογισμού.

Ένα τέτοιο δίκτυο είναι δυνατόν να παρουσιάζει παράλληλη και κατανεμημένη ικανότητα επεξεργασίας, όπως ο βιολογικός εγκέφαλος. Επιπλέον, οι πληροφορίες μπορεί να

αποθηκευτούν κατανεμημένα σε κατάλληλες βάσεις σύνδεσης έτσι ώστε να επιτυγχάνεται κάποια ανοχή σφάλματος. Στα τεχνητά δίκτυα αυτά τα χαρακτηριστικά απεικονίζονται μέσω πολλών παράλληλων και κατανεμημένων μοντέλων επεξεργασίας για τις διεργασίες των γνωστικών διαδικασιών στους [16].

1.1.4 Σημερινές προκλήσεις στην ανάπτυξη των τεχνητών νευρωνικών δικτύων

Από τη δεκαετία του 1980, μέχρι σήμερα η έρευνα σχετικά με τα τεχνητά νευρωνικά δίκτυα έχει επιφέρει αξιοσημείωτες εξελίξεις. Ωστόσο παρά την εξέλιξη της έρευνας, η εφαρμογή τους αντιμετωπίζει πολλά προβλήματα, όπως η επιλογή της δομής και των παραμέτρων των δικτύων, η επιλογή των μαθησιακών δειγμάτων, η επιλογή των αρχικών τιμών, η σύγκλιση των αλγορίθμων μάθησης κλπ.

Είναι γνωστό ότι η απόδοση των νευρωνικών δικτύων είναι ευαίσθητη στον αριθμό των νευρώνων. Μικρός αριθμός νευρώνων μπορεί να οδηγήσει σε κακή προσέγγιση, ενώ μεγάλος αριθμός νευρώνων μπορεί να συμβάλλει σε προβλήματα υπερφόρτωσης. Η παραπάνω πρόκληση είναι αντιφατική καθώς για να επιτευχθεί καλύτερη απόδοση του δικτύου πρέπει να γίνει απλοποίηση της τοπολογίας του δικτύου.

Έτσι από τη δεκαετία του 1990 για τη βελτιστοποίηση του σχεδιασμού και των παραμέτρων των τεχνητών νευρωνικών δικτύων έχουν χρησιμοποιηθεί με επιτυχία οι αλγόριθμοι εξέλιξης [17]. Αξίζει να αναφερθεί ότι ένας αλγόριθμος τυχαίας αναζήτησης προσομοιώνει την επιλογή της φύσης και τη διαδικασία της εξέλιξης [18] παρουσιάζοντας το πλεονέκτημα της καλής παγκόσμιας ικανότητας αναζήτησης και μάθησης της κατά προσέγγιση βέλτιστης λύσης χωρίς να χρησιμοποιεί να παρουσιάζει αποκλίσεις λόγω λειτουργικού σφάλματος. Επίσης μπορεί να παρέχει νέες ιδέες και μεθόδους για την επίλυση προβλημάτων όσον αφορά στους ειδικούς κανόνες της φυσικής εξέλιξης και την ανωτερότητα του πληθυσμού που βελτιστοποιεί την αναζήτηση.

Σήμερα, η υπολογιστική νοημοσύνη βρίσκεται στο στάδιο της ταχείας ανάπτυξης, των κύριων τεχνικών του, συμπεριλαμβανομένης της ασαφούς τεχνολογίας, του νευρικού δικτύου, του εξελικτικού αλγορίθμου κλπ. Με την ανάπτυξη της τεχνολογίας υπολογιστών, αυτές οι μέθοδοι έχουν ήδη επιτύχει τεράστια πρόοδο και παρουσίασαν μια συγχωνευτική τάση. Μπορούν να αλληλοενισχυθούν μέσω της αλληλοσυμπλήρωσής τους.

Τα τελευταία χρόνια, όλο και περισσότεροι ερευνητές [19] προσπαθούν να συνδυάσουν τους έξυπνους αλγορίθμους με τα τεχνητά νευρωνικά δίκτυα και ελπίζουν ότι συνδυάζοντας τα

πλεονεκτήματα τους, μπορούν να δημιουργήσουν μια πιο αποτελεσματική μέθοδο. Έτσι με τη χρήση των έξυπνων αλγορίθμων γίνεται προσπάθεια για τη βελτιστοποίηση του σχεδιασμού του δικτύου, την προεπεξεργασία των δεδομένων εισόδου του δικτύου και τη συναρμολόγηση των τεχνητών νευρωνικών δικτύων [20].

Σήμερα στον τομέα της ιατρικής με χρήση έξυπνων αλγορίθμων πραγματοποιείται ψηφιοποίηση των εικόνων. Η μεγάλη πλειοψηφία των σημερινών μορφών απεικόνισης ‘κατασκευάζει’ ψηφιακές πληροφορίες από τις εξετάσεις. Τα ψηφιακά συστήματα σε σύγκριση με τα αναλογικά, έχουν πολλά πλεονεκτήματα. Η ανάγκη για ψηφιακή απεικόνιση των ιατρικών δεδομένων αναλύεται στο επόμενο Κεφάλαιο.

2 Ψηφιακή απεικόνιση ιατρικών δεδομένων

Στο παρόν Κεφάλαιο αναλύεται και συζητείται πώς πραγματοποιείται η ψηφιακή απεικόνιση των ιατρικών εικόνων. Τα ψηφιακά δεδομένα αποθηκεύονται σε υπολογιστές, οι οποίοι είναι προσπελάσιμοι από οποιαδήποτε τοποθεσία μέσω ασφαλών δικτύων, επιτρέποντας έτσι την αποθήκευση των δεδομένων εκτός χώρου και μακροπρόθεσμα. Η φορητότητα των ψηφιακών δεδομένων επιτρέπει την απευθείας προώθηση των εικόνων του ασθενούς ή τη μεταφορά τους σε άλλους γιατρούς σε δίσκους δεδομένων. Ένα σημαντικό πλεονέκτημα των ψηφιακών δεδομένων είναι ότι επιτρέπει την μεταγενέστερη επεξεργασία, όπως την αλλαγή της κλίμακας του γκρι, την ενίσχυση των άκρων και την ανακατασκευή εικόνας [21].

Το DICOM (Digital Imaging and Communications in Medicine) είναι ένα παγκόσμιο πρότυπο τεχνολογίας πληροφορικής που έχει σχεδιαστεί για να διασφαλίζει τη διαλειτουργικότητα των συστημάτων που χρησιμοποιούνται για την αποθήκευση, την εκτύπωση και τη μετάδοση πληροφοριών στην ιατρική απεικόνιση. Αυτό περιλαμβάνει έναν ορισμό μορφής αρχείου και ένα πρωτόκολλο επικοινωνίας δικτύου [21].

Τα αρχεία DICOM έχουν ομοιομορφία και μεταφέρουν πληροφορίες και εικόνες σε συστήματα ικανά να λαμβάνουν δεδομένα εικόνας και του ασθενούς σε μορφή DICOM. Ένα αρχείο DICOM περιέχει ταυτοποίηση ασθενούς, τοποθεσία προέλευσης, χαρακτηριστικά της εικόνας που περιλαμβάνουν μέγεθος εικονοστοιχείου και την ίδια την εικόνα. Είναι σημαντικό ότι, εντός των ορίων ενός αρχείου DICOM, τα δεδομένα ασθενών και τα δεδομένα εικόνας δεν μπορούν να διαχωριστούν. Κατά συνέπεια, η εικόνα DICOM είναι πάντα αναγνωρίσιμη για τον εξεταζόμενο ασθενή [22].

Αξίζει να σημειωθεί ότι, αν και το DICOM περιλαμβάνει πληροφορίες για τα μέρη του σώματος, σήμερα η ταξινόμηση των εικόνων αποτελεί πρόκληση στον τομέα της ιατρικής. Κατ' αρχάς, μια ανάγνωση μιας εικόνας μπορεί να περιέχει περίπου 15% σφάλματα και έτσι να περιοριστεί η ακρίβεια της αναγνώρισης μέρος του σώματος. Δεύτερον, οι πληροφορίες κειμένου στο DICOM είναι εξαιρετικά αφηρημένες και μπορεί να μην περιγράφουν με ακρίβεια τις ανατομίες που περιέχονται στη σάρωση. Για παράδειγμα, είναι δύσκολο να διαπιστωθεί εάν μια σάρωση με ετικέτα DICOM (0018,0015) = "TSPINE" περιλαμβάνει το μεσαίο μέρος του ήπατος [22].

Επιπλέον, ο πολυγλωσσικός χαρακτήρας των πληροφοριών του bodypart DICOM γίνεται ένα άλλο εμπόδιο για την ανάκτηση δεδομένων με βάση το κείμενο. Αντίθετα, ένας αξιόπιστος αλγόριθμος αναγνώρισης bodypart με βάση την εικόνα μπορεί να αντιμετωπίσει όλες αυτές τις τρεις προκλήσεις αξιοποιώντας τις εγγενείς ανατομικές πληροφορίες εμφάνισης [22].

Το ακρωνύμιο DICOM αντιπροσωπεύει την ψηφιακή απεικόνιση και την επικοινωνία στην ιατρική και πιο συγκεκριμένα στην ιατρική απεικόνιση. Αυτό είναι αποτέλεσμα χρόνιας προσπάθειας με στόχο τη δημιουργία του πιο καθολικού και θεμελιώδους προτύπου στην ψηφιακή ιατρική απεικόνιση. Ως εκ τούτου, παρέχει όλα τα απαραίτητα εργαλεία για να υπάρξει μια ακριβής διάγνωση και αναπαράσταση και επεξεργασία των δεδομένων της ιατρικής απεικόνισης.

Αξίζει να αναφερθεί, ότι μια απεικόνιση τύπου DICOM δεν είναι απλώς μια μορφή εικόνας ή αρχείου. Αντίθετα, πρόκειται για ένα ολοκληρωμένο πρωτόκολλο μεταφοράς, αποθήκευσης και προβολής δεδομένων το οποίο έχει σχεδιαστεί για να καλύπτει όλες τις λειτουργικές πτυχές της ψηφιακής ιατρικής απεικόνισης (και γι 'αυτό το DICOM θεωρείται από πολλούς ως ένα σύνολο προτύπων, αντί για ένα μόνο πρότυπο) [23].

Με βάση λοιπόν το DICOM έχει διαμορφωθεί πραγματικά το τοπίο της σύγχρονης ιατρικής, καθώς αυτό παρέχει [23]:

1. Ένα καθολικό πρότυπο ψηφιακής ιατρικής. Όλες οι τρέχουσες ψηφιακές συσκευές λήψης εικόνων παράγουν εικόνες DICOM και επικοινωνούν μέσω δικτύων DICOM. Η τρέχουσα ιατρική ροή εργασίας ελέγχεται εμμέσως από ένα πλήθος κανόνων DICOM.
2. Εξαιρετική ποιότητα εικόνας. Για παράδειγμα, το DICOM υποστηρίζει έως 65.536 (16 bit) αποχρώσεις του γκρι για απεικόνιση μονόχρωμης εικόνας, καταγράφοντας έτσι τις παραμικρές αποχρώσεις στην ιατρική απεικόνιση. Ωστόσο η μετατροπή των DICOM εικόνων σε αρχεία JPEG ή bitmap (BMP), είναι πάντα περιορισμένη σε 256 αποχρώσεις του γκρι, με αποτέλεσμα συχνά η διαγνωστική ανάγνωση να καθίσταται ανέφικτη. Το DICOM εκμεταλλεύεται τις πιο σύγχρονες και προηγμένες τεχνικές ψηφιακής αναπαράστασης εικόνας για να παρέχει άριστη ποιότητα διαγνωστικής εικόνας.
3. Πλήρης υποστήριξη για πολλές παραμέτρους απόκτησης εικόνας και διαφορετικών τύπων δεδομένων. Το DICOM αποθηκεύει όχι μόνο τις εικόνες, αλλά καταγράφει επίσης μια πληθώρα άλλων παραμέτρων που σχετίζονται με την εικόνα, όπως η θέση

του ασθενούς σε 3D, τα φυσικά μεγέθη αντικειμένων στην εικόνα, το πάχος, τις παραμέτρους έκθεσης της εικόνας κ.ο.κ. Αυτά τα δεδομένα εμπλουτίζουν πάρα πολύ το ενημερωτικό περιεχόμενο ενός ασθενούς.

4. Πλήρης κωδικοποίηση ιατρικών δεδομένων. Χρησιμοποιούνται περισσότερα αρχεία και μηνύματα DICOM με περισσότερες από 2000 τυποποιημένες ιδιότητες (λεξικό δεδομένων DICOM) για τη μετάδοση διαφόρων ιατρικών δεδομένων, όπως το όνομα του ασθενούς, το βάθος χρώματος μιας εικόνας, κ.ά. Αυτά τα δεδομένα συχνά είναι απαραίτητα για την ακριβή διάγνωση.
5. Σαφήνεια στην περιγραφή συσκευών ψηφιακής απεικόνισης και της λειτουργικότητάς τους. Η ραχοκοκαλιά οποιασδήποτε ιατρικής απεικόνισης είναι η ορθή ψηφιακή απεικόνιση και η λειτουργικότητα των συσκευών. Το DICOM ορίζει τη λειτουργικότητα μιας ιατρικής συσκευής με πολύ ακριβείς και ανεξάρτητους από τη συσκευή όρους. Οι ιατρικές συσκευές λειτουργούν μέσω των διεπαφών τους με το DICOM με μια πολύ απλή διαδικασία, αφήνοντας ελάχιστα περιθώρια για σφάλματα.

2.1 Τρόπος λειτουργίας

Το DICOM χρησιμοποιεί τη δική του γλώσσα λειτουργίας, με βάση το μοντέλο του πραγματικού κόσμου (μοντέλο πληροφοριών DICOM). Πιο αναλυτικά, δηλαδή, όλα τα πραγματικά δεδομένα - ασθενείς, μελέτες, ιατρικές συσκευές κ.λπ. - από το DICOM θεωρούνται ως αντικείμενα με αντίστοιχες ιδιότητες ή χαρακτηριστικά. Οι ορισμοί αυτών των αντικειμένων και ιδιοτήτων είναι τυποποιημένοι σύμφωνα με αυτή [21].

Αυτή χρησιμοποιεί ορισμούς αντικειμένων πληροφοριών (Information Object Definitions, IODs). Τα IODs επομένως είναι συλλογές χαρακτηριστικών, περιγράφοντας κάθε συγκεκριμένο αντικείμενο δεδομένων. Για παράδειγμα ένα IOD ασθενούς, μπορεί να εμπεριέχει το όνομα, τον αριθμό ιατρικού αρχείου (ID), το φύλο, την ηλικία, το βάρος, εάν είναι καπνιστής ή όχι, και ούτω καθεξής, καθώς και πολλά χαρακτηριστικά που χρειάζονται για να συλλεχθούν όλες οι κλινικά σχετικές πληροφορίες ενός ασθενούς. Με μια ευρύτερη έννοια, ένας ασθενής (όπως και κάθε άλλο DICOM αντικείμενο) είναι αποτελείται από ένα σύνολο χαρακτηριστικών όπως απεικονίζεται στην Εικ.2.1 [21].

Patient in
Real World



Patient object in
DICOM World

=

Name	John Smith
ID	123456
DOB	19681108
Weight	75.5
Sex	M
.....	

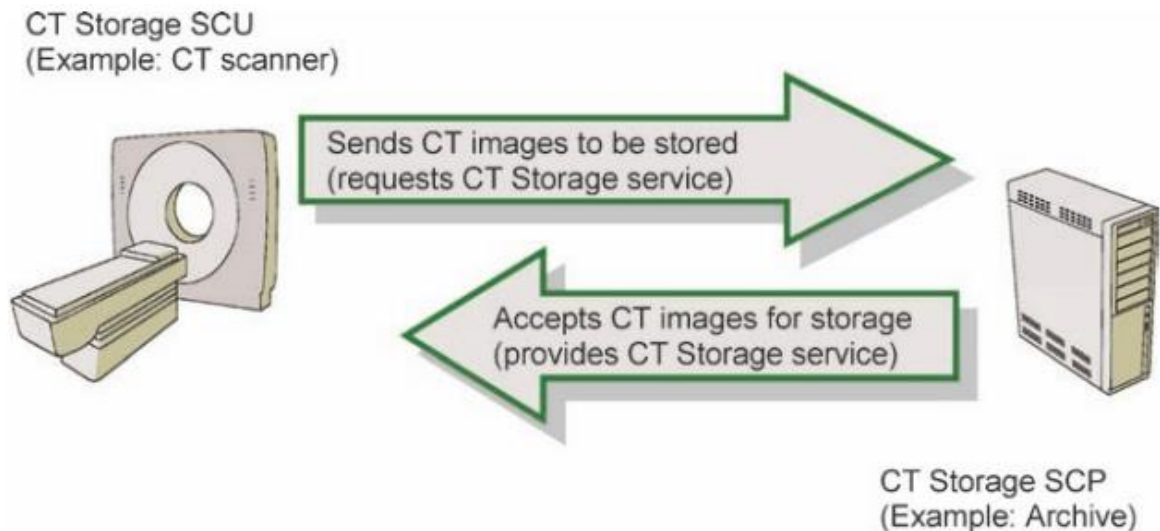
} Patient IOD
(Information
Object
Definition)

Εικόνα 2.1. Πληροφορίες ασθενούς όπως αποθηκεύονται στο DICOM [21].

Το DICOM διατηρεί μια λίστα με όλα τα χαρακτηριστικά (περισσότερα από 2000 από αυτά), γνωστό ως λεξικό δεδομένων DICOM, για να διασφαλιστεί η συνοχή στην ονομασία και στην επεξεργασία χαρακτηριστικών. Για παράδειγμα, οι ιδιότητες των ασθενών μας - όνομα, ημερομηνία γέννησης, το φύλο και ούτω καθεξής - περιλαμβάνονται επίσης στο DICOM Data Dictionary.

Μόλις ληφθούν τα δεδομένα ως χαρακτηριστικά δεδομένων DICOM, μπορούν να μεταδοθούν και να υποβληθούν σε επεξεργασία μεταξύ διαφόρων συσκευών και λογισμικού DICOM (Entities Application (AE), όπως είναι γνωστό στο DICOM). Το DICOM, όπως προαναφέρθηκε, αντιπροσωπεύει ένα μοντέλο παροχής υπηρεσιών.

Οι εφαρμογές DICOM παρέχουν υπηρεσίες μεταξύ τους, επειδή κάθε υπηρεσία συνήθως περιλαμβάνει κάποια ανταλλαγή δεδομένων (που συνήθως εκτελούνται μέσω ενός δικτύου υπολογιστών), συγκεκριμένοι τύποι υπηρεσιών με τα δεδομένα (IOD) που επεξεργάζονται, συνδέονται μεταξύ τους. Για παράδειγμα, η αποθήκευση μιας εικόνας αξονικής τομογραφίας από έναν ψηφιακό σαρωτή τομογράφου, σε ένα ψηφιακό αρχείο PACS αντιστοιχεί στο CT Storage SOP, όπως φαίνεται στην Εικ. 2.2 [21].



Εικόνα 2.2. Σχηματική απεικόνιση των υπηρεσιών DICOM [21].

Σε αυτό το συγκεκριμένο παράδειγμα, η εικόνα CT αντιπροσωπεύει το DICOM IOD (δεδομένα DICOM αντικειμένου). Στην επόμενη υποενότητα συζητείται αναλυτικά η περίπτωση κατηγοριοποίησης ιατρικών εικόνων με βάση το DICOM.

2.2 Διαχείριση και κατηγοριοποίηση ιατρικών εικόνων με βάση το DICOM για σωστή διάγνωση

Το DICOM υποστηρίζει ένα ευρύ φάσμα μορφών εικόνας για την αποθήκευση αυτών (7FE0, 0010) εικονοστοιχεία. Οι μορφές αυτές μπορούν να κατηγοριοποιηθούν σε δύο κύριες ομάδες [21]:

1. Εικόνες του DICOM: οι μορφές που χρησιμοποιούνται μόνο από το DICOM. Είναι συνήθως παλαιότερης μορφής, που παρουσιάστηκαν στην αρχή της εποχής του υπολογιστή πριν αναπτυχθούν καλύτερες μορφές εικόνας. Μοιάζουν με τα ακατέργαστα BMP με διαφορετικούς τρόπους ‘συσκευασίας’ των ψηφιολέξεων pixel.
2. Ανεξάρτητοι τυποποιημένοι μορφότυποι που έγιναν αποδεκτοί από το DICOM: Αυτοί περιλαμβάνουν γνωστές μορφές όπως JPEG, RLE (κωδικοποίηση χρόνου εκτέλεσης), ZIP, και τα γνωστά (αλλά όλο και πιο δημοφιλή) JPEG2000 και JPEG-LS.

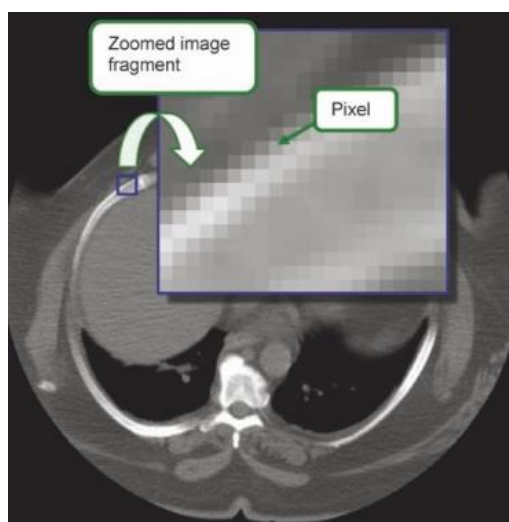
Όλα αυτά τα πρότυπα συνδέονται επίσης με διάφορες τεχνικές συμπίεσης εικόνων, τόσο αναστρέψιμες όσο και μη αναστρέψιμες, γεγονός που τα καθιστά ιδιαίτερα χρήσιμα στην ιατρική απεικόνιση (η μείωση του μεγέθους των δεδομένων μιας εικόνας είναι σημαντική). Η

βασική τυποποιημένη προσέγγιση, όταν το κύριο πρότυπο (DICOM) περιλαμβάνει τα άλλα πρότυπα (όπως JPEG) που χρησιμοποιούνται για συγκεκριμένες εργασίες (όπως η κωδικοποίηση μιας εικόνας), είναι πολύ βολικό, συνεπές και έχει μεγάλη πρακτική έννοια.

2.2.1 Αποθήκευση ιατρικών εικόνων

Μια ψηφιακή εικόνα, όπως είναι γνωστό, είναι ένας ορθογώνιος πίνακας pixel (στοιχεία εικόνας), από μικροσκοπικές κουκκίδες διαφορετικών χρωμάτων που αποτελούν την πραγματική εικόνα. Για παράδειγμα, μια τυπική εικόνα αξονικής τομογραφίας έχει πλάτος 512 εικονοστοιχεία και ύψος 512 pixels, δηλαδή εμπεριέχει $512 \times 512 = 262.144$ pixels. Επομένως αυτά τα εικονοστοιχεία της κάθε γραμμής ξεκινώντας από την πάνω αριστερή γωνία δίνουν μια ακολουθία 262.144 pixel, τα οποία μπορούν να αποθηκευτούν σε ένα αρχείο. Ουσιαστικά, αυτό το αρχείο είναι η ακατέργαστη εικόνα BMP (Εικ.2.3).

Κάθε εικονοστοιχείο μιας εικόνας μπορεί να είναι ένα μίγμα αρκετών τιμών δείγματος. Η πιο χαρακτηριστική περίπτωση είναι ένα έγχρωμο εικονοστοιχείο, το οποίο περιλαμβάνει τρία ανεξάρτητα δείγματα χρώματος: κόκκινο, πράσινο και μπλε (γνωστά ως έγχρωμος χώρος RGB). Ενώ η ισχύς κάθε δείγματος συμβάλλει στη φωτεινότητα του εικονοστοιχείου, το μείγμα τους δημιουργεί το χρώμα.

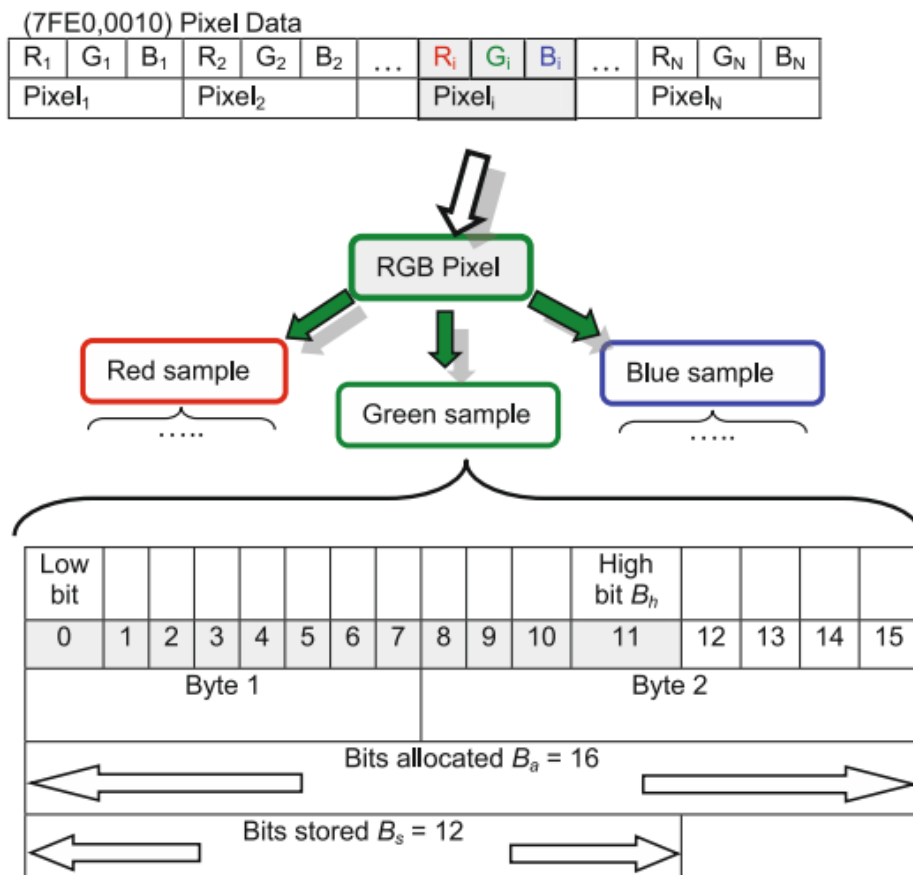


Εικόνα 2.3. Μεγέθυνση στα pixel εικόνας από ιατρική απεικόνιση αξονικής τομογραφίας [21].

Ωστόσο, οι εικόνες σε κλίμακα του γκρι συνήθως αποθηκεύονται με ένα μονόχρωμο δείγμα ανά εικονοστοιχείο, που αντιστοιχεί στο εικονοστοιχείο φωτεινότητα της κλίμακας του γκρι. Όταν χρησιμοποιούνται δείγματα 2-byte (OW VR), το DICOM παρέχει για $22 \times 8 = 65.536$

πιθανές κλίμακες του γκρι. Σε σε κάθε περίπτωση, η επιλογή της δειγματοληψίας εικονοστοιχείων παραμένει σταθερή για όλα τα εικονοστοιχεία σε μια εικόνα, ανάλογα μόνο με την τεχνική απεικόνισης.

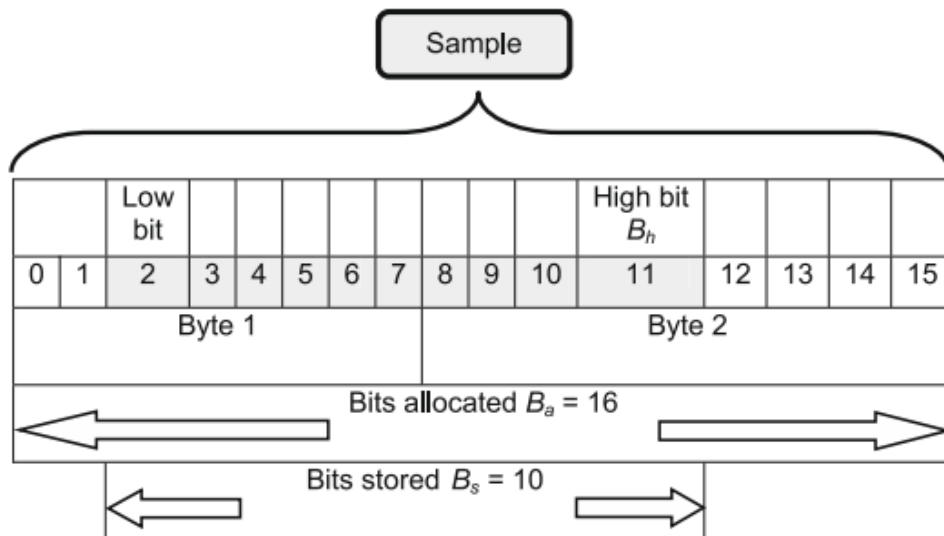
Η Εικ.2.4 συνοψίζει τη δομή ενός εικονοστοιχείου DICOM πολλαπλών δειγμάτων: το εικονοστοιχείο αποτελείται από τρία δείγματα (κόκκινο, πράσινο και μπλε) και κάθε δείγμα σε αυτό το συγκεκριμένο παράδειγμα έχει $B_s=12$ bits. Επειδή όλα τα δεδομένα υπολογιστών αποθηκεύονται σε bytes (όπου 1 byte=8 bits), τα 12 bits στρογγυλεύονται μέχρι $B_a=16$, και $B_h=11$. Συνήθως $B_h=B_s-1$ (ο αριθμός των bits ξεκινά από 0). Η ίδια αποθήκευση 16-bit διατίθεται για τα άλλα δείγματα (κόκκινο και μπλε) και ολόκληρη η εικόνα γράφεται ως μια ακολουθία των δειγμάτων των εικονοστοιχείων της.



Εικόνα 2.4. Αποθήκευση pixel με βάση το DICOM [21].

Στην περίπτωση της Εικ. 2.5 το δείγμα αποθηκεύτηκε σε $B_s=10$ bits, αλλά μετρώντας από το bit δυο. Επειδή το DICOM αποθηκεύει τις τιμές του μεγάλου δυαδικού ψηφίου B_h και των

αποθηκευμένων bits B_s , μπορείτε να ανακτηθεί το πρώτο "χαμηλό bit" ως $B_{h+1}-B_s$ και έπειτα θα ακολουθήσουν τα υπόλοιπα.



Εικόνα 2.5. Παράδειγμα αποθήκευσης σε bits πιο περίπλοκου στοιχείου [21].

Το πιο δύσκολο κομμάτι σε όλες τις κωδικοποιήσεις εικονοστοιχείων της πρώτης κατηγορίας είναι το πώς μπορεί το DICOM να επωφεληθεί από τα αχρησιμοποίητα bits.

2.2.2 Συμπίεση ιατρικών εικόνων

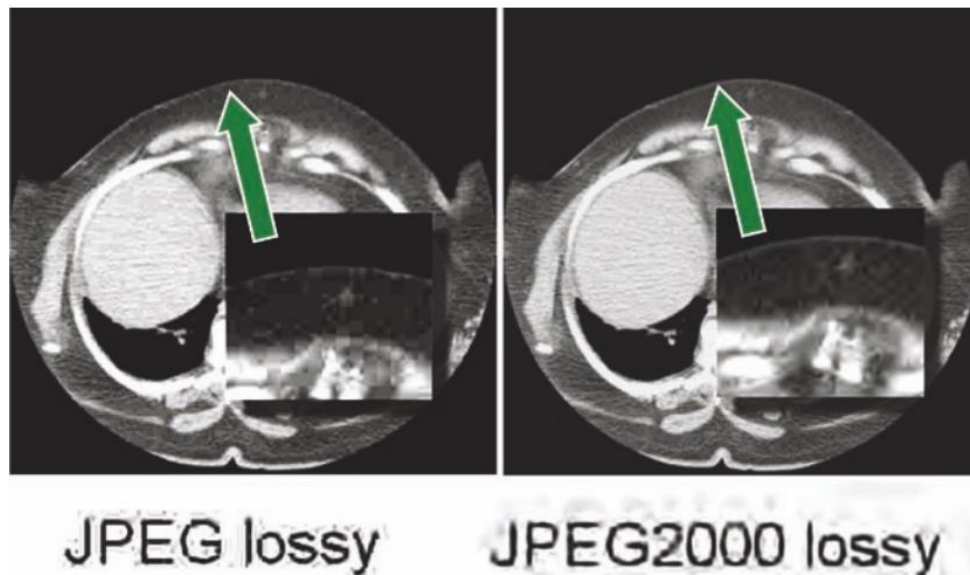
Εξετάζοντας μια τυπική εικόνα αξονικής τομογραφίας με πλάτος εικόνας=ύψος=512, κλίμακα του γκρι (ένα δείγμα ανά εικονοστοιχείο), και αποθηκευμένα bits=12, για την αποθήκευση θα χρειαζόνταν 12 bits, 2 full bytes (16 bits). Ο συνολικός αριθμός εικονοστοιχείων είναι 512×512 , οπότε ο συνολικός αριθμός των byte για την αποθήκευσή τους θα είναι: $512 \times 512 \times 2 = 524.288$. Έτσι, χρειαζόμαστε μισό megabyte για τα δεδομένα pixel. Επομένως κρίνεται απαραίτητη η συμπίεση των εικόνων, λειτουργικότητα που έχει προστεθεί στο DICOM [22].

Με δεδομένο αυτό, η συμπίεση εικόνας έχει συμπεριληφθεί στο πρότυπο DICOM από την ίδρυσή της. Η συμπίεση εικόνας στοχεύει το εικονοστοιχείο εικόνας bytes και τα αναδιατάσσει έξυπνα σε πολύ μικρότερη μορφή. Αυτό μειώνει σημαντικά το μέγεθος της αρχικής εικόνας, το οποίο με τη σειρά του εξοικονομεί χώρο αποθήκευσης και συχνά μειώνει το χρόνο λήψης της εικόνας. Αναφέρεται ότι το DICOM περιλαμβάνει σχεδόν όλους τους αλγόριθμους συμπίεσης όπως RLE, JPEG, JPEG2000, JPEG-LS και ZIP. Όλοι αυτοί οι

αλγόριθμοι έχουν αναπτυχθεί ξεχωριστά, με αποτέλεσμα να διαχωρίζονται από πρότυπα και εφαρμογές του Διεθνούς Οργανισμού Τυποποίησης (ISO) [22, 23].

Ωστόσο η χρήση της συμπίεσης εικόνων θα μπορούσε να έχει σημαντική επίδραση στην εμφάνιση μιας εικόνας. Οποιοσδήποτε αλγόριθμος συμπίεσης θα προσπαθήσει να βρει και να εξαλείψει τις πιο περιττές και επαναλαμβανόμενες πληροφορίες, κάνοντας τα δεδομένα πιο σύντομα.

Η υπερβολική συμπίεση μπορεί να τροποποιήσει σημαντικά μια εικόνα (Εικ.2.6). Η εικόνα στα αριστερά του σχήματος έχει υπερσυμπεστεί με το JPEG, γεγονός που οδηγεί στα περίφημα τετράγωνα JPEG που είναι εύκολα αντιληπτά ως ορθογώνιες περιοχές με σταθερό χρώμα. Η ίδια εικόνα στα δεξιά έχει εκτεθεί σε απώλεια συμπίεσης JPEG2000. Επομένως πρέπει να επιλεγθεί η ορθή μέθοδος συμπίεσης των εικόνων [24].



Εικόνα 2.6. Τροποποίηση εικόνας αξονικής τομογραφίας ύστερα από συμπίεση [21].

2.2.3 Παρεμβολή εικόνας (ψηφιακή μεγέθυνση)

Η παρεμβολή εικόνας (ψηφιακό ζουμ) είναι η διαδικασία αλλαγής της αρχικής ανάλυσης μιας ψηφιακής εικόνας, αυξάνοντας τεχνητά τον αριθμό pixel της. Για παράδειγμα, εάν η αρχική εικόνα είναι $256 \times 256 = 65.536$ εικονοστοιχεία. Αν αυτό παρουσιασθεί σε εικόνα σε πλήρη οθόνη 1024×1024 , πρέπει να το έχετε ως εικόνα $1024 \times 1024 = 1.048.576$ εικονοστοιχεία [25].

Έτσι, όταν γίνεται μεγέθυνση αυτής της εικόνας σε πλήρη οθόνη, τα επιπλέον $1.048.576 - 65.536 = 983.040$ pixels, παράγονται από έναν αλγόριθμο παρεμβολής (συνήθως bicubic

παρεμβολή) και εισάγονται μεταξύ των αρχικών εικονοστοιχείων εικόνας, κάνοντας έτσι την εικόνα μεγαλύτερη. Ως αποτέλεσμα, όταν γίνεται μεγέθυνση σε μια ψηφιακή εικόνα, φαίνονται προοδευτικά περισσότερα pixels που δεν υπήρχαν στα αρχικά δεδομένα αλλά προστέθηκαν από την παρεμβολή [26].

2.2.4 Ανακατασκευή εικόνας

Η ανακατασκευή εικόνας για τη διάγνωση είναι πολύ σημαντική, Ιδιαίτερα οι ανακατασκευές εικόνας - επίπεδες, καμπύλες, επιφάνειες, 3D - έχουν γίνει ζωτικής σημασίας στη σύγχρονη ακτινολογία. Είναι επίσης ένα από τα πιο χαρακτηριστικά πλεονεκτήματα της ψηφιακής απεικόνισης. Αυτή η μετεπεξεργασία ανακατασκευής χωρίς τις πληροφορίες που συλλέγονται στα αντικείμενα δεδομένων DICOM, δε θα ήταν δυνατή [27].

Σήμερα το DICOM χρησιμοποιείται ευρέως για την αποστολή, τη διανομή και την αποθήκευση εικόνων, ανεξάρτητα από τη μηχανή, τον κατασκευαστή ή τον τρόπο λειτουργίας. Το DICOM ελέγχει την σωστή απεικόνιση της εικόνας και χωρίς DICOM δε θα ήταν δυνατή η διάγνωση μέσω υπολογιστή [24]. Ωστόσο αυτό παρουσιάζει το μειονέκτημα ότι βασίζεται στο χρήστη. Σε αυτό το σημείο έγκειται η χρησιμότητα των νευρωνικών δικτύων και της τεχνητής όρασης, όπως αναλύεται στο Κεφάλαιο 3.

3 Μηχανική μάθηση και τεχνητή νοημοσύνη: ο ρόλος των νευρωνικών δικτύων στην ταξινόμηση ιατρικών απεικονίσεων

Λόγω του μεγάλου όγκου ιατρικών δεδομένων η ταχεία ανάπτυξη των ιατρικών εικόνων και των μεθόδων απαιτεί εκτεταμένες προσπάθειες από ιατρικό εμπειρογνώμονα για τη διάγνωση. Ωστόσο η διάγνωση αυτή είναι υποκειμενική, επιρρεπής σε ανθρώπινο λάθος και μπορεί να υπάρχουν μεγάλες διαφορές μεταξύ διαφορετικών εμπειρογνομόνων.

Σε αυτό λοιπόν το σημείο, δηλαδή στο χρονικό σημείο της διάγνωσης, χρησιμοποιούνται ως εναλλακτική λύση οι τεχνικές μηχανικής μάθησης για την αυτοματοποίηση της διαδικασίας διάγνωσης. Υπολογιστικές μηχανές μπορεί να αποκτήσουν την ικανότητα να χειρίζονται μεγάλα ιατρικά δεδομένα εικόνας για ακριβή και αποτελεσματική διάγνωση. Η βαθιά εκμάθηση αυτών των μηχανών όχι μόνο θα βοηθήσει στην επιλογή και εξαγωγή συγκεκριμένων χαρακτηριστικών, αλλά θα μπορούν να κατασκευάσουν και νέα δεδομένα διάγνωσης μιας ασθένειας, από την παροχή μοντέλων πρόβλεψης, για να βοηθήσουν αποτελεσματικά τον ιατρό [28].

Τα τελευταία χρόνια η μηχανική μάθηση (Machine Learning) και η τεχνητή νοημοσύνη (Artificial Intelligence) έχουν εξελιχθεί σημαντικά. Οι τεχνικές ML και AI έχουν διαδραματίσει σημαντικό ρόλο στο ιατρικό πεδίο όπως στην ιατρική επεξεργασία εικόνας, στη διάγνωση μέσω υπολογιστή, στην ερμηνεία των εικόνων, κ.ά. [28].

Το ML και το AI διευκολύνουν και βοηθούν τους ιατρούς ώστε να μπορούν να διαγνώσουν και να προβλέπουν ακριβείς και ταχύτερους κινδύνους από ασθένειες και να τους αποτρέπουν εγκαίρως. Αυτές οι τεχνικές ενισχύουν τις ικανότητες των ιατρών και των ερευνητών να κατανοήσουν πώς να αναλύουν τις εικόνες που θα οδηγήσουν σε ασθένειες. Μέχρι σήμερα, οι αλγόριθμοι βαθιάς μάθησης οι οποίοι βασίζονται σε ερεθίσματα έδειξαν ελπιδοφόρες επιδόσεις στη διάγνωση.

Οι περισσότεροι ερευνητές πιστεύουν ότι μέσα στα επόμενα 15 χρόνια, η τεχνολογία βαθιάς μάθησης θα διεισδύσει στον τομέα της ιατρικής απεικόνισης και ιδιαίτερα στην περίπτωση των ακτινολογικών δεδομένων, σε τέτοιο βαθμό όπου οι εφαρμογές θα αναλάβουν τον άνθρωπο και το μεγαλύτερο μέρος της διάγνωσης θα πραγματοποιείται από έξυπνες μηχανές.

Πιο αναλυτικά, η οφθαλμολογία θα είναι το πρώτο πεδίο που οι αλγόριθμοι βαθιάς μάθησης θα φέρουν επανάσταση, ωστόσο η διάγνωση του καρκίνου έχει λάβει ιδιαίτερη προσοχή.

Χαρακτηριστικό είναι ότι το Google DeepMind Health συνεργάζεται με την Εθνική Υγεία Service, του Ηνωμένου Βασιλείου για την επεξεργασία ιατρικών δεδομένων έως και 1 εκατομμυρίων ασθενών σε όλα τα νοσοκομεία [29].

Πολλές εργασίες διάγνωσης εικόνας απαιτούν αρχική αναζήτηση για τον εντοπισμό ανωμαλιών, ποσοτικοποίηση της μέτρησης και αλλαγή στην πάροδο του χρόνου. Τα αυτόματα εργαλεία ανάλυσης εικόνας με βάση αλγόριθμους μηχανικής μάθησης είναι οι βασικοί παράγοντες που επιτρέπουν τη βελτίωση της ποιότητας της διάγνωσης και της ερμηνείας της εικόνας διευκολύνοντας τη διάγνωση μέσω αποτελεσματικών ταυτοποιήσεων του ευρήματος [30].

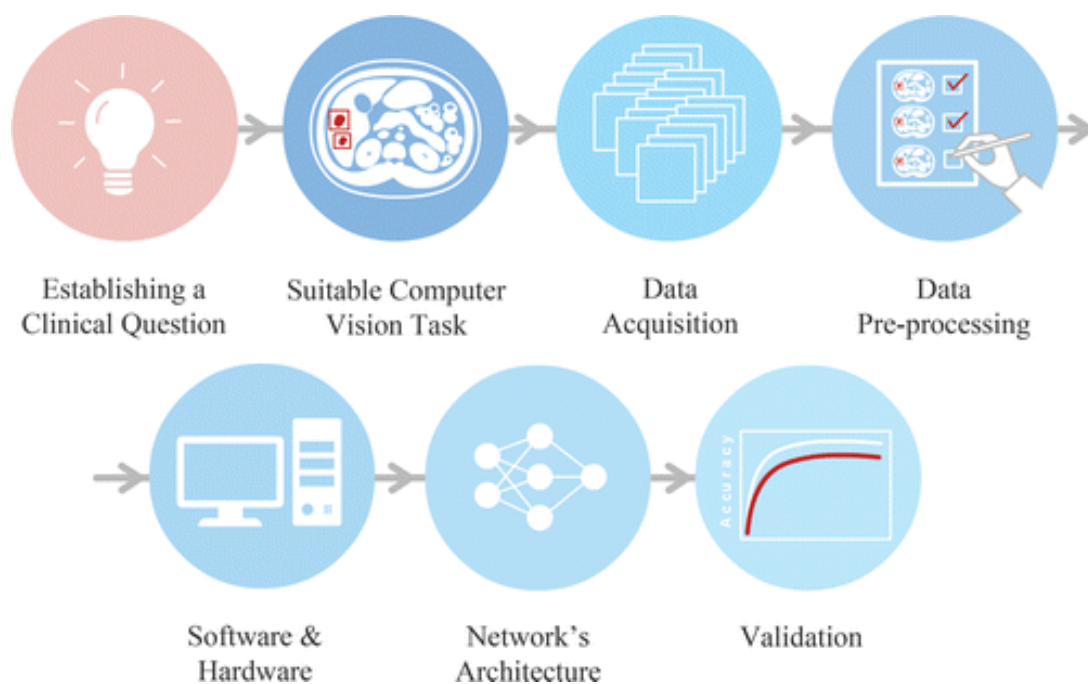
Η βαθιά επομένως εκμάθηση είναι μια εκτεταμένη τεχνική που παρέχει πίσω διαγνώσεις ακρίβειας. Χαρακτηριστικό είναι ότι εφαρμογές βαθιάς μάθησης στην υγειονομική περίθαλψη καλύπτουν ένα ευρύ φάσμα προβλημάτων που κυμαίνονται από τον προσυμπτωματικό έλεγχο και την παρακολούθηση ασθένειας οδηγώντας σε εξατομικευμένες προτάσεις θεραπείας [30].

Σήμερα διάφορες πηγές δεδομένων, από ακτινολογική απεικόνιση (ανίχνευση ακτίνων X, CT και MRI), απεικόνιση παθολογίας και πρόσφατα, γονιδιωματικές αλληλουχίες έφεραν ένα τεράστιο όγκο δεδομένων στη διάθεση των ιατρών. Ωστόσο, ακόμα και σήμερα δεν υπάρχουν τα κατάλληλα εργαλεία για μετατροπή αυτών των δεδομένων σε χρήσιμες πληροφορίες.

3.1 Τα νευρωνικά δίκτυα στην περίπτωση των ιατρικών ακτινολογικών εικόνων

Ο σχεδιασμός μιας μελέτης βαθιάς μάθησης συνεπάγεται ένα κοινό πρότυπο που περιλαμβάνει ποικίλα στάδια. Το πρώτο στάδιο είναι η διατύπωση ενός κλινικού ερωτήματος. Μετά τη διατύπωση του ερωτήματος, επιλέγεται με τις κατάλληλες μετρήσεις μια κατάλληλη ‘διεργασία’ ηλεκτρονικής όρασης.

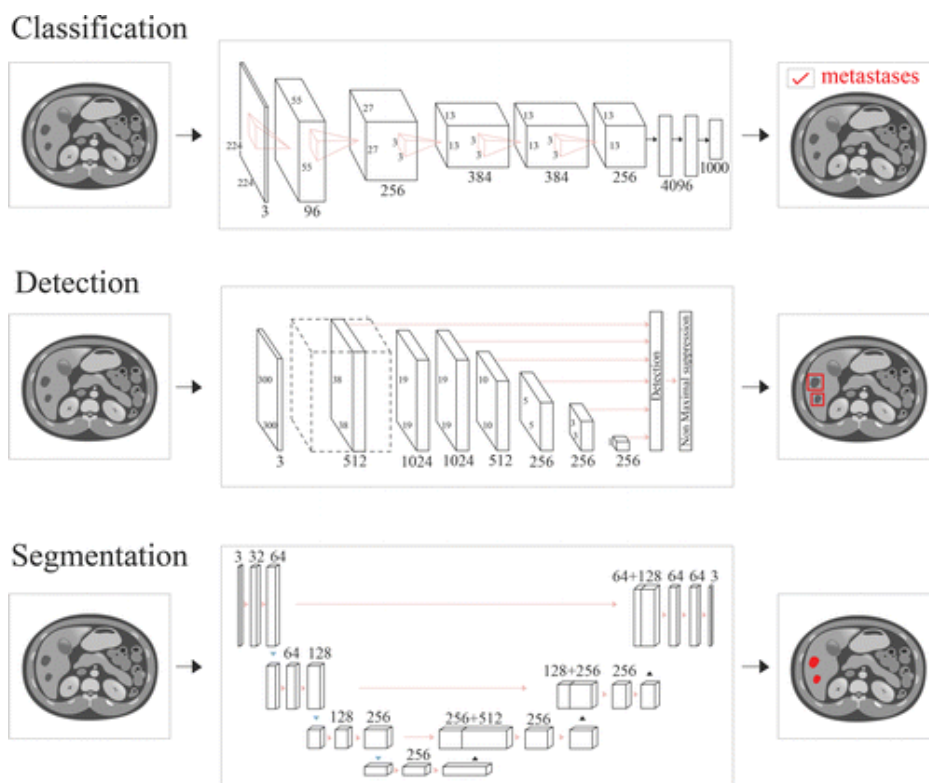
Στη συνέχεια, εξετάζεται το στάδιο της απόκτησης δεδομένων και της προεπεξεργασίας των δεδομένων. Αυτά περιλαμβάνουν τον προγραμματισμό των δεδομένων τόσο για την εκπαίδευση όσο και για τις δοκιμές, καθώς και την επισήμανση των ιατρικών δεδομένων. Στη συνέχεια, επιλέγεται το πλαίσιο λογισμικού, η πλατφόρμα του υλικού και σχεδιάζεται η αρχιτεκτονική του δικτύου. Τέλος, χρησιμοποιώντας τις επιλεγμένες μετρήσεις τα αποτελέσματα επικυρώνονται με βάση τα δεδομένα των δοκιμών που διεξήχθησαν (Εικ.3.1).



Εικόνα 3.1. Εκμάθηση νευρωνικών δικτύων για ταξινόμηση ιατρικών απεικονίσεων [31].

Το κλινικό ερώτημα που επιλέγεται εξελίσσεται από τα διάφορα ακτινολογικά πεδία και τις σχετικές λεπτομέρειες της απεικόνισης. Συνήθως μια κλινική ερμηνεία των ακτινολογικών δεδομένων βασίζεται στην εμπειρία των ακτινολόγων. Ωστόσο, η βαθιά εκμάθηση έχει το ξεχωριστό πλεονέκτημα ότι μπορεί να επεξεργαστεί μη δομημένα δεδομένα, ενώ η κλασική μηχανική μάθηση προτιμάται για δεδομένα που χαρακτηρίζονται ως καλά δομημένα και έχουν καλά καθορισμένα χαρακτηριστικά.

Η μηχανική όραση (machine vision) του υπολογιστή που είναι ιδιαίτερα εφαρμόσιμη στον τομέα της ακτινολογίας περιλαμβάνει την ταξινόμηση, ανίχνευση και τμηματοποίηση των εικόνων (Εικ.3.2) [31].



Εικόνα 3.2. Ταξινόμηση, ανίχνευση και τμηματοποίηση με μηχανική όραση [31].

Η ταξινόμηση είναι το καθήκον κατηγοριοποίησης ή επισήμανσης μιας εικόνας σε μια συγκεκριμένη κατηγορία - για παράδειγμα, ταξινόμηση των ακτινογραφιών στο στήθος είτε ως φυσιολογικές είτε με χαρακτηριστικά της φυματίωσης. Αυτή η εργασία στον τομέα της μηχανικής όρασης είναι θεμελιώδης για την πραγματοποίηση περαιτέρω εργασιών ενός νευρωνικού δικτύου [31].

Η ανίχνευση επιτρέπει την ταυτοποίηση της θέσης βλαβών, οργάνων ή άλλων αντικειμένων ενδιαφέροντος - για παράδειγμα, εντοπίζοντας τις συντεταγμένες x, y των ηπατικών μαζών. Η τμηματοποίηση εφαρμόζεται για να οριστούν τα ακριβή όρια pixel-wise ενός οργάνου ή ενός παθολογικού χαρακτηριστικού.

Η βελτιστοποίηση μιας εικόνας περιλαμβάνει εργασίες όπως η βελτίωση της ανάλυσης της εικόνας, καθώς και τη σύνταξη της συνθετικής εικόνας εισόδου. Κάθε συγκεκριμένη εργασία χρησιμοποιεί μια συγκεκριμένη στατιστική μέθοδο (μετρήσεις) για να παρουσιάσει τα αποτελέσματα [31].

Για παράδειγμα, τα αποτελέσματα ταξινόμησης μπορούν να αξιολογηθούν χρησιμοποιώντας λειτουργικές χαρακτηριστικές καμπύλες του δέκτη, τα αποτελέσματα της ανίχνευσης μπορούν να μετρηθούν με τη χρήση του πραγματικού θετικού ρυθμού και του ψευδώς

θετικού ρυθμού και τα αποτελέσματα κατάτμησης μπορούν να εκτιμηθούν χρησιμοποιώντας τον συντελεστή Ζαριών ή τον δείκτη Jaccard, που χρησιμοποιείται για τη σύγκριση της ομοιότητας δύο δειγμάτων.

Με βάση τη βιβλιογραφία, για την ταξινόμηση των εικόνων, τα συνελκτικά νευρωνικά δίκτυα έχουν αποδειχθεί ιδιαίτερα αποτελεσματικά στην επεξεργασία ακατέργαστων δεδομένων εικονοστοιχείων. Αξίζει να αναφερθεί ότι οι μέχρι σήμερα εξελίξεις στον τομέα της μηχανικής εκμάθησης και ιδιαίτερα σε αυτόν της μηχανικής όρασης έχουν βασιστεί κυρίως πάνω σε ένα συγκεκριμένο αλγόριθμο, αυτόν των νευρωνικών συνελκτικών δικτύων (convolutional neural network) [32].

3.1.1 Συνελκτικά νευρωνικά δίκτυα (convolutional neural networks)

Ένα συνελκτικό δίκτυο είναι ένας αλγόριθμος βαθιάς μάθησης που μπορεί να πάρει μια εικόνα εισόδου, να αναδείξει τα σημαντικά σημεία σε διάφορες πτυχές/αντικείμενα μιας εικόνας και έχει τη δυνατότητα να διαφοροποιήσει το ένα από το άλλο. Η προεπεξεργασία που απαιτείται σε ένα ConvNet είναι πολύ χαμηλότερη σε σύγκριση με άλλους αλγόριθμους ταξινόμησης. Ενώ στις πρωτότυπες μεθόδους τα φίλτρα είναι χειροκίνητα, με αρκετή εκπαίδευση, ένα ConvNet έχει τη δυνατότητα να μάθει αυτά τα φίλτρα/χαρακτηριστικά.

Η αρχιτεκτονική ενός ConvNet είναι ανάλογη με εκείνη του μοτίβου συνδεσιμότητας των νευρώνων στον ανθρώπινο εγκέφαλο και εμπνεύστηκε από την οργάνωση του οπτικού φλοιού. Οι μεμονωμένοι νευρώνες ανταποκρίνονται σε ερεθίσματα μόνο σε μια περιορισμένη περιοχή του οπτικού πεδίου που είναι γνωστή ως πεδίο υποδοχής. Μια συλλογή τέτοιων πεδίων επικαλύπτεται για να καλύψει ολόκληρη την οπτική περιοχή [33].

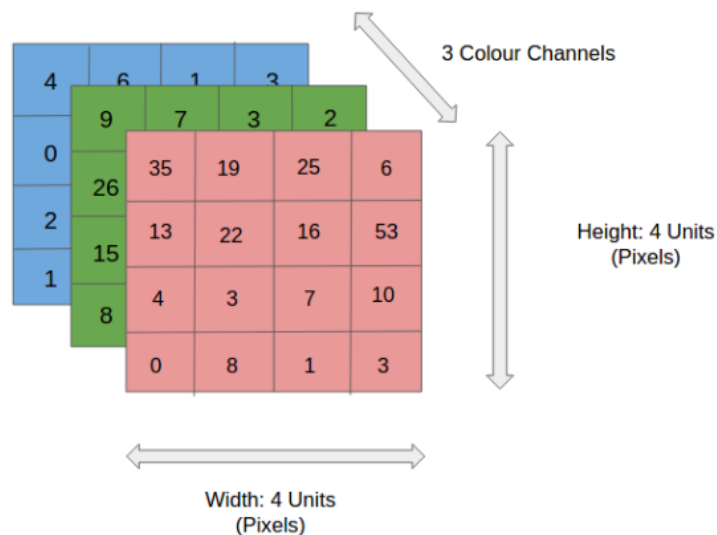
Δηλαδή, αυτά τα δίκτυα χρησιμοποιούν ιεραρχική τοπολογία συνδέσεων εμπνευσμένων από βιολογικά οπτικά συστήματα, όπου οι κόμβοι χαμηλού επιπέδου στο δίκτυο επεξεργάζονται ένα χωρικά περιορισμένο πλέγμα εικονοστοιχείων και οι κόμβοι υψηλότερου επιπέδου κωδικοποιούν όλο και πιο περίπλοκα χαρακτηριστικά συνδυάζοντας απλούστερα χαρακτηριστικά από χαμηλότερα επίπεδα.

Τα βάρη μοιράζονται μεταξύ των κόμβων στο ίδιο στρώμα με τρόπο που επιτρέπει την αναγνώριση του ίδιου μοτίβου εικόνας σε οποιοδήποτε αριθμό χωρικών θέσεων. Τα συνεργατικά νευρωνικά δίκτυα αυξανόμενου βάθους και πολυπλοκότητας έχουν χρησιμοποιηθεί για να προωθήσουν την κατάσταση της τεχνολογίας στην ταξινόμηση

εικόνων. Αυτό έχει καταδειχθεί με τις πιο πρόσφατες επιτυχίες στην ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [25, 26].

Ένα συνελκτικό νευρωνικό δίκτυο, μέσω της εφαρμογής των σχετικών φίλτρων, είναι σε θέση να καταγράψει με επιτυχία τις χωρικές και προσωρινές εξαρτήσεις σε μια εικόνα. Η αρχιτεκτονική αυτών των δικτύων, λόγω της μείωσης του αριθμού των παραμέτρων που εμπλέκονται και της επαναχρησιμοποίησης των βαρών, έχει καλύτερη προσαρμογή στο σύνολο των δεδομένων. Με άλλα λόγια, το δίκτυο μπορεί να εκπαιδευτεί ώστε να κατανοήσει καλύτερα και λεπτομερώς μια εικόνα.

Για παράδειγμα, στην Εικ. 3.3 απεικονίζεται μια RGB εικόνα που χωρίστηκε στα τρία επίπεδα χρωμάτων της, κόκκινο, πράσινο και μπλε. Γενικότερα, υπάρχουν αρκετοί τέτοιοι χρωματικοί χώροι στους οποίους υπάρχουν εικόνες - Grayscale, RGB, HSV, CMYK, κλπ.



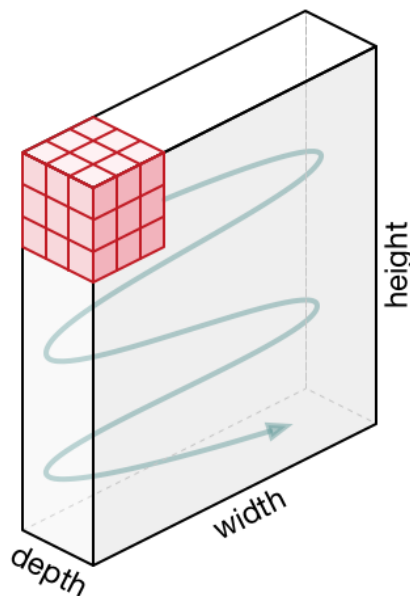
Εικόνα 3.3. Ανάλυση εικόνας σε τρία βασικά χρώματα [34].

Στην περίπτωση λοιπόν που οι εικόνες χρησιμοποιούνται σε διαστάσεις, για παράδειγμα 8K (7680 × 4320), θα υπήρχε μεγάλη υπολογιστική δυσκολία. Ο ρόλος του ConvNet είναι να μειώσει το μέγεθος των εικόνων σε μια μορφή που είναι ευκολότερη στη διαδικασία, χωρίς να χάνουν τα κύρια χαρακτηριστικά τους που είναι κρίσιμα για την επίτευξη μιας καλής πρόβλεψης. Σε αυτό το σημείο αναφέρεται ότι όταν σχεδιάζεται η αρχιτεκτονική ενός τέτοιου δικτύου, η μη αλλοίωση των χαρακτηριστικών μιας εικόνα είναι σημαντικό. Επιπλέον, η αρχιτεκτονική του δικτύου θα πρέπει επίσης να είναι επεκτάσιμη σε μαζικά σύνολα δεδομένων.

Για παράδειγμα, στην περίπτωση εικόνας με διαστάσεις, 5 (Υψος) x 5 (πλάτος) x 1 (Αριθμός καναλιών, π.χ. RGB). Το στοιχείο που εμπλέκεται στην εκτέλεση της συνέλιξης το πρώτο μέρος ενός συνελκτικού στρώματος ονομάζεται Kernel/Filter, K , επιλέγοντας για παράδειγμα το K ως μήτρα $3 \times 3 \times 1$.

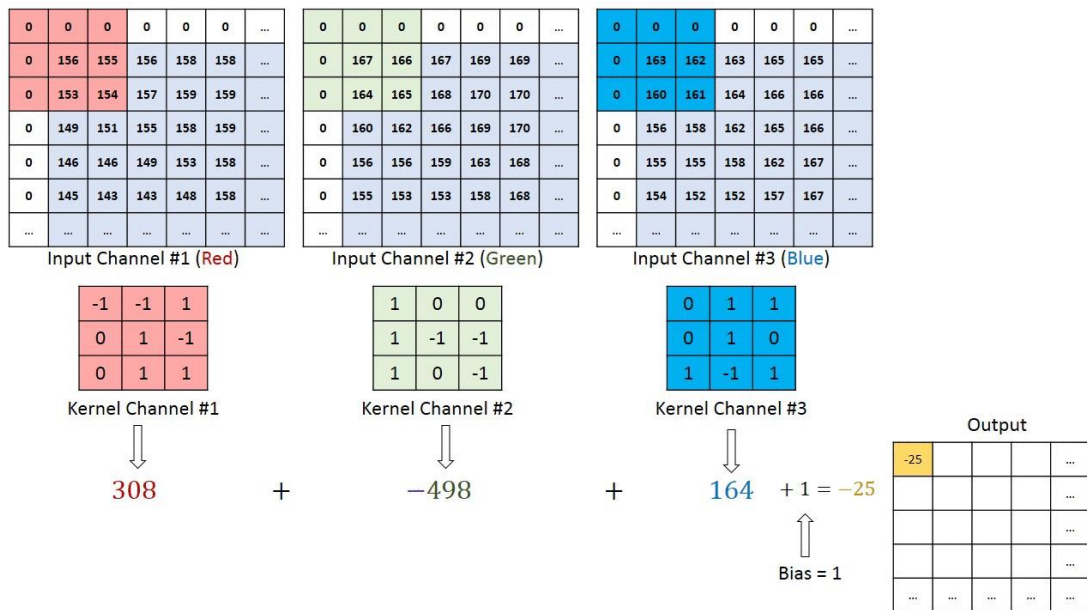
$$\text{Kernel/Filter, } K = \begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{matrix}$$

Κάθε φορά που εκτελεί μια λειτουργία πολλαπλασιασμού μήτρας μεταξύ του K και του τμήματος P της εικόνας πάνω από το οποίο ο πυρήνας κινείται, ο πυρήνας μετατοπίζεται 9 φορές εξαιτίας του $\text{Stride Length} = 1$ (Μη Στρεβλωμένο) (Εικ.3.4).



Εικόνα 3.4. Κίνηση Kernel [34].

Το φίλτρο μετακινείται προς τα δεξιά με μια συγκεκριμένη τιμή Stride μέχρι να αναλύσει το πλήρες πλάτος. Προχωρώντας, πέφτει προς τα κάτω στην αρχή (αριστερά) της εικόνας με την ίδια τιμή Stride και επαναλαμβάνει τη διαδικασία έως ότου διασχίσει ολόκληρη την εικόνα.



Εικόνα 3.5. Λειτουργία συνέλιξης σε μήτρα εικόνας $M \times N \times 3$ με πυρήνα $3 \times 3 \times 3$ [34].

Στην περίπτωση εικόνων με πολλαπλά κανάλια (π.χ. RGB), ο πυρήνας έχει το ίδιο βάθος με εκείνο της εικόνας εισόδου. Ο πολλαπλασιασμός της μήτρας εκτελείται μεταξύ του K_n και του I_n stack ($[K_1, I_1], [K_2, I_2], [K_3, I_3]$) και όλα τα αποτελέσματα αθροίζονται για να δώσουν ένα συνελικτικό κανάλι βάθους.

Ο στόχος της λειτουργίας Convolution είναι η εξαγωγή των χαρακτηριστικών υψηλού επιπέδου, όπως οι άκρες, από την εικόνα εισόδου. Τα ConvNets δεν χρειάζεται να περιορίζονται σε μόνο ένα επίπεδο μετατροπής. Συμβατικά, το πρώτο ConvLayer είναι υπεύθυνο για τη λήψη των χαρακτηριστικών χαμηλού επιπέδου, όπως οι άκρες, το χρώμα, ο προσανατολισμός κλίσης κ.λπ. Με πρόσθετα επίπεδα, η αρχιτεκτονική προσαρμόζεται επίσης στα χαρακτηριστικά υψηλού επιπέδου, δίνοντάς ένα δίκτυο που έχει την ορθή κατανόηση των εικόνων στο σύνολο δεδομένων.

Υπάρχουν δύο τύποι αποτελεσμάτων – ένα κατά το οποίο το περιστρεφόμενο χαρακτηριστικό μειώνεται σε διαστάσεις σε σύγκριση με την είσοδο και το άλλο στο οποίο η διάσταση είτε αυξάνεται είτε παραμένει η ίδια. Αυτό γίνεται με την εφαρμογή Valid Padding στην περίπτωση του πρώτου, ή του ίδιου Padding στην περίπτωση του δεύτερου.

Για παράδειγμα, όταν μια εικόνα $5 \times 5 \times 1$ αυξάνεται σε μια εικόνα $6 \times 6 \times 1$ και στη συνέχεια εφαρμόζεται ο πυρήνας $3 \times 3 \times 1$ πάνω σε αυτό, διαπιστώνεται ότι η συνελικτική μήτρα έχει διαστάσεις $5 \times 5 \times 1$ (Same padding). Από την άλλη πλευρά, αν γίνει η ίδια λειτουργία χωρίς

κάποια αύξηση των διαστάσεων,, παρουσιάζεται ένας πίνακας που έχει διαστάσεις του ίδιου του Kernel (3x3x1) (Valid Padding) [34].

Ένα σημαντικό σημείο στην ανάλυση ιατρικών εικόνων είναι η προετοιμασία των δεδομένων. Η χειροκίνητη επισήμανση και σχολιασμός είναι χρονοβόρες. Ωστόσο μπορούν να χρησιμοποιηθούν διάφορες επιλογές για τη διαδικασία της προετοιμασίας των δεδομένων και περιλαμβάνουν τα εξής, με αυξανόμενη σειρά πολυπλοκότητας: α) σήμανση εικόνας (π.χ. "ακτινογραφία με φυματίωση"). (β) σημάνσεις περιοχής ενδιαφέροντος (ROI), όπως τετραγωνικά ή κυκλικά ROI, και (γ) τμηματοποίηση με εικονοστοιχεία.

Οι διαφορετικές εργασίες μηχανικής όρασης στον υπολογιστή απαιτούν διαφορετικούς σχολιασμούς. Για παράδειγμα, ο διαχωρισμός των οργάνων ή των παθολογικών χαρακτηριστικών απαιτεί κατάτμηση με εικονοστοιχεία, ενώ ταξινομεί τις ακτινογραφίες στο στήθος είτε παρουσιάζοντας αυτές με φυματίωση ή τις φυσιολογικές. Αυτή η διαδικασία, απαιτεί μόνο σήμανση εικόνας και έτσι μπορεί δυνητικά να επιτρέψει τη χρήση μιας μεγαλύτερης κοόρτης [31].

Για παράδειγμα, στον τομέα της κατάτμησης εικόνων του εγκεφάλου, οι μέθοδοι ανάλυσης των εικόνων με δενδριτική δομή είναι θέμα που διερευνείται ευρέως. Οι Yoo et al. [42] έκαναν κατάτμηση σε εικόνες πολλαπλών εικόνων MR που έδειχναν αλλοιώσεις σκλήρυνσης κατά πλάκας. Αυτοί χρησιμοποιώντας ένα δίκτυο δημιούργησαν μια περιοχή και εξήγαγαν χαρακτηριστικά από εικόνες T2-σταθμισμένης πυκνότητας και πυκνότητας πρωτονίων για την τελική εποπτευόμενη ταξινόμηση.

Προκειμένου να βελτιωθεί η απόδοση του μοντέλου από το θόρυβο και την ευρωστία έναντι της υπερκατασκευής, οι Maier et al. [43] πρότειναν μια επιπλέον δενδριτική δομή για να εντοπίσει, να ταξινομήσει και να ποσοτικοποιήσει την υποξεία ισχαιμική βλάβη του εγκεφαλικού. Αυτοί χρησιμοποίησαν τοπικά χαρακτηριστικά voxel-wise, όπως ένταση, σταθμισμένο τοπικό μέσο, τοπικό ιστόγραμμα και κεντρική απόσταση 2D.

Ωστόσο, η μέθοδος τους μπορεί να 'ασχοληθεί' μόνο με τις σταθμισμένες με T1 και σταθμισμένες με διάχυση σειρές δεδομένων και εικόνες υψηλής ποιότητας. Τα πολυτροπικά δεδομένα από τον ίδιο ασθενή μπορούν να παρέχουν επιπλέον χρήσιμες πληροφορίες για τη διάγνωση [43].

Οι τεχνικές προεπεξεργασίας, όπως η τυποποίηση, δεν πρέπει να αποτελούν πρόβλημα, εφόσον όλα τα δείγματα στα δεδομένα εκπαίδευσης και δοκιμών υπόκεινται στις ίδιες

σταθερές κανονικοποίησης. Ωστόσο, τα στρώματα κανονικοποίησης παρτίδων, όπου οι συντελεστές κλιμάκωσης εξαρτώνται από τις παρτίδες και όχι από ολόκληρο το σετ εκπαίδευσης, μπορεί να αφαιρέσουν κάποιες από τις πληροφορίες και είναι γνωστό ότι μειώνουν την απόδοση σε περιοχές ανάλυσης ιατρικών εικόνων όπως η CT.

Η κλίμακα αποτελεί έναν άλλον τύπο ποσοτικών πληροφοριών που υπάρχει συχνά στην ιατρική ανάλυση εικόνας. Για παράδειγμα, το μέγεθος ενός εικονοστοιχείου σε εικόνες ακτίνων X παρέχεται συνήθως στην επικεφαλίδα DICOM. Γνωρίζοντας την ποσότητα των εικονοστοιχείων που καταλαμβάνει μία ανωμαλία και το μέγεθος των εικονοστοιχείων από την κεφαλίδα, μπορούμε να υπολογιστεί το μέγεθος της βλάβης στην εικόνα.

Αυτό το ακριβές μέγεθος μπορεί να είναι ένα σημαντικό χαρακτηριστικό. Στη μαστογραφία (εικόνες ακτίνων X του μαστού), για παράδειγμα, τα καρκινώματα δεν έχουν αμετάβλητη κλίμακα: η αύξηση του μεγέθους ενός μικρού καρκίνου, δεν σημαίνει ότι αντιπροσωπεύει στην πραγματικότητα ένα μεγάλο όγκο (Εικ.3.1) [40].

Επίσης σε αντίθεση με τις φυσικές εικόνες, στην περίπτωση των ιατρικών, ο προσανατολισμός γενικά δεν είναι ένα σημαντικό χαρακτηριστικό των ιατρικών δεδομένων και τα αντικείμενα στις ιατρικές εικόνες δεν έχουν κανονικό προσανατολισμό. Για παράδειγμα, οι υμένες ιστών σε ψηφιακές παθολογικές εικόνες τοποθετούνται στο γυαλί χωρίς προκαθορισμένη ευθυγράμμιση. Κάτι παρόμοιο ισχύει για τους όγκους στο στήθος, το στήθος ή τον εγκέφαλο και πολλές άλλες ασθένειες. Παρόλο που ο ιστός υποβάθρου μπορεί να έχει κάποια δομή που εξαρτάται από τον προσανατολισμό, οι ίδιες οι ανωμαλίες δεν έχουν.

Στην περίπτωση των ακτινολογικών δεδομένων επίσης η κατάτμηση των εικόνων είναι ένα απαραίτητο βήμα στην αποτελεσματική διάγνωση και θεραπεία των ασθενειών. Πιο συγκεκριμένα βοηθάει τους κλινικούς να κατανοήσουν τις δομικές πληροφορίες και τις χωρικές ανατομικές σχέσεις. Ωστόσο η ανάλυση των εικόνων εξαρτάται από την εμπειρία των κλινικών ιατρών και είναι χρονοβόρα [41]. Επομένως, για τη βελτίωση της ανάλυσης διάγνωσης και για την αναπαραγωγικότητα των κλινικών μελετών μεγάλης κλίμακας είναι απαραίτητες οι αυτόματες μέθοδοι ταξινόμησης.

Με τη σειρά τους οι Mitra et al. [44] πρότειναν τη χρήση χαρακτηριστικών από πολυτροπικά δεδομένα για την ταξινόμηση των ισχαιμικών αλλοιώσεων, λευκής ύλης και άλλων δευτερογενών βλαβών. Στη μελέτη τους, ο συνδυασμός των αλγορίθμων συνδυάζει την

εκτίμηση πιθανότητας μεγιστοποίησης προσδοκίας και το τυχαίο πεδίο Bayesian-Markov για τον διαχωρισμό των πιθανών περιοχών αλλοιώσεων από τα δεδομένα FLAIR και στη συνέχεια στα πολυτροπικά δεδομένα εφαρμόζεται δενδριτική δομή.

Επομένως, τα νευρωνικά δίκτυα και οι τεχνικές βαθιάς μάθησης είναι ισχυρά εργαλεία όσον αφορά στις εργασίες κατάτμησης των εικόνων μαγνητικής του εγκεφάλου. Οι Si et al. [45] με χρήση τεχνητού νευρικού δικτύου με επίπεδα γκρίζου και στατιστικά χαρακτηριστικά ως εισροές, πρότειναν μια ημιαυτόματη μέθοδο ταξινόμησης των εικονοστοιχείων της μαγνητικής τομογραφίας του εγκεφάλου σε τραυματισμένους και υγιείς ιστούς. Αυτοί παρατήρησαν ότι η κατάτμηση των ιστών πρώιμου εγκεφάλου είναι πιο δύσκολη από αυτή των εγκεφάλων ενηλίκων λόγω της χαμηλότερης αντίθεσης ιστών [46], ενώ οι πολλαπλές μορφές εικόνας περιέχουν συμπληρωματικές πληροφορίες για ανεπαρκή αντίθεση ιστών [47].

Οι Zhang et al. [48] έδειξαν ότι οι κλασματικές εικόνες ανισοτροπίας είναι πιο ισχυρές για τη διάκριση της φαιάς ύλης και της λευκής ύλης και ότι οι T2-ζυγισμένες εικόνες έχουν υψηλότερες επιδόσεις στη σύλληψη του εγκεφαλονωτιαίου υγρού. Αυτοί πρότειναν για τη βελτίωση της απόδοσης τμηματοποίησης μια CNN μέθοδο που συνδυάζει αυτά τα δεδομένα εικόνας πολλαπλών μεταβολών.

Παρομοίως, ο Kleesiek [49] κατάτμησε τους εγκεφαλικούς και τους μη εγκεφαλικούς ιστούς με την τροφοδοσία των δεδομένων σε ένα νευρωνικό δίκτυο με επτά κρυμμένα συνελκτικά στρώματα. Το πρότυπό τους μπορεί να εφαρμοστεί σε κάθε μονήρη εικόνα ή σε συνδυασμό με διάφορες μορφές με διαφορετικό μέγεθος.

Οι μέθοδοι βαθιάς εκμάθησης μπορούν επίσης να κατατμήσουν αυτόματα τις εικόνες MRI του ανθρώπινου εγκεφάλου σε πολλές ανατομικές περιοχές. Όπως φαίνεται στην Εικ.3.1, οι Chen et al. [50] επέκτεινε το ResNet σε ογκομετρική ανατομική κατάτμηση του εγκεφάλου. Αυτοί ενσωμάτωσαν τις δυνατότητες εμφάνισης της εικόνας χαμηλής στάθμης, τις 'κρυφές' πληροφορίες για το σχήμα και το πλαίσιο υψηλού επιπέδου για περαιτέρω βελτίωση της απόδοσης ογκομετρικής κατάτμησης.

Το επόμενο βήμα στη διαδικασία σχεδιασμού μιας έρευνας είναι η επιλογή μιας πλατφόρμας υλικού και ενός πλαισίου λογισμικού. Η βαθιά εκμάθηση έγινε δυνατή χάρη στην ανάπτυξη νέας τεχνολογίας υλικού, η οποία επιτρέπει την εξαιρετικά γρήγορη επεξεργασία των λειτουργιών του πίνακα. Η εισαγωγή της μονάδας επεξεργασίας γραφικών, ή της GPU, άνοιξε τη δυνατότητα εκπαίδευσης μεγάλων δικτύων με αποτελεσματικό τρόπο και

εξοικονόμησης χρόνου. Παράλληλα με την ανάπτυξη υλικού, δημιουργήθηκαν επίσης εξελιγμένα εργαλεία λογισμικού που επέτρεψαν το σχηματισμό βαθιών νευρωνικών δικτύων.

Τέλος, η αρχιτεκτονική ενός νευρωνικού δικτύου είναι από τα πιο σημαντικά. Η αρχιτεκτονική ορίζει τη δομή των στρωμάτων του νευρικού δικτύου. Αν και υπάρχουν διάφορες αρχιτεκτονικές μορφές, αυτές βασίζονται γενικά σε ένα επαναλαμβανόμενο πρότυπο ακολουθιών.

Όπως αναφέρθηκε σε προηγούμενο Κεφάλαιο, η αρχιτεκτονική περιλαμβάνει ένα στρώμα εισόδου, κρυφά επίπεδα μάθησης (τα οποία στις περισσότερες περιπτώσεις αποτελούνται από υποστρώματα συνέλιξης και συγκέντρωσης) και ένα στρώμα εξόδου. Οι διαφορετικές εργασίες απαιτούν διαφορετικές αρχιτεκτονικές δικτύων και η επιλογή της κατάλληλης αρχιτεκτονικής μπορεί να βελτιώσει τη συνολική απόδοση. Οι ερευνητές μπορούν να επιλέξουν να χρησιμοποιήσουν προηγουμένως αναπτυγμένα μοντέλα CNN ή μπορούν να κατασκευάσουν τη δική τους «εσωτερική» αρχιτεκτονική.

Τα δεδομένα εισόδου για ένα νευρωνικό δίκτυο μπορεί να είναι είτε μια δισδιάστατη μήτρα είτε ένας τρισδιάστατος τανυστής. Σε ένα μεγάλο μέρος ιατρικών εικόνων, τα δεδομένα έχουν τη μορφή τρισδιάστατων όγκων. Κατά την ανάλυση των ογκομετρικών δεδομένων, μπορεί να χρησιμοποιηθεί μια δισδιάστατη αρχιτεκτονική, ωστόσο, η χρήση μιας ανάλυσης μίας ενότητας μπορεί να οδηγήσει στην απώλεια σημαντικών ογκομετρικών πληροφοριών. Το μειονέκτημα αυτής της προσέγγισης μπορεί να ξεπεραστεί χρησιμοποιώντας μια τρισδιάστατη αρχιτεκτονική. Αυτό, ωστόσο, είναι εις βάρος ενός πολύ υψηλότερου υπολογιστικού κόστους.

Οι προκαθορισμένες αρχιτεκτονικές του CNN, όπως το AlexNet ή το VGG, μπορούν να προωθηθούν σε σύνολα δεδομένων μεγάλης κλίμακας, όπως το ImageNet, πριν από την εκπαίδευση στο μικρό συγκεκριμένο σύνολο ιατρικών δεδομένων. Αυτή η μέθοδος, που αναφέρεται ως εκμάθηση προπαιδευσεως ή μεταφοράς, εφαρμόζεται σήμερα ευρέως στον ιατρικό κόσμο και μπορεί να μετριάσει τον περιορισμό των μικρών συνόλων δεδομένων.

Η εκμάθηση της μεταφοράς μπορεί να γίνει κατανοητή με την εξέταση της ανθρώπινης συμπεριφοράς: Όταν ένα άτομο αντιμετωπίζει ένα νέο έργο, μεταφέρει πληροφορίες που συγκεντρώνονται από άλλα πεδία γνώσης. Ενώ στη μη μεταφορική μάθηση, στα μη ιατρικά δεδομένα μεγάλης κλίμακας, όλα τα βάρη κατανέμονται τυχαία στη φάση προ-κατάρτισης,

στη μάθηση μεταφοράς, ενώ τα βάρη των στρωμάτων του δικτύου προέρχονται από την κατάρτιση.

Επομένως στη μεταβίβαση της γνώσης σχετικά με τον αριθμό των στρώσεων που θα πρέπει να επανεκπαιδευθούν με τα ιατρικά δεδομένα και αν τα βάρη στα στρώματα που θα επανεκπαιδευθούν θα πρέπει να είναι καλά ρυθμισμένα ή εκπαιδευμένα από το μηδέν, υπάρχουν μεταβολές.

3.2 Είδη ταξινόμησης εικόνων ιατρικής απεικόνισης

Η ταξινόμηση είναι ένας σημαντικός τομέας μηχανικής μάθησης η οποία έχει προσελκύσει το ενδιαφέρον καθώς μπορεί να εφαρμοστεί σε πολλές υποσχόμενες εφαρμογές, όπως η εξαγωγή δεδομένων, η χρηματοοικονομική πρόβλεψη, η οργάνωση και η ανάκτηση πολυμέσων και η χρήση βιοπληροφορικής [35].

Με βάση τη βιβλιογραφία μπορούν να επιτευχθούν μέθοδοι ταξινόμησης, που είναι εξειδικευμένες σε συγκεκριμένες εφαρμογές, με πολύ καλές επιδόσεις όσον αφορά στην ακρίβεια και στο χρόνο πολυπλοκότητας, λαμβάνοντας υπόψη διάφορους παράγοντες και προηγούμενες γνώσεις.

Οι Ali & Smith [35] έκαναν μια λεπτομερή ανάλυση και συνέκριναν οκτώ διαφορετικούς αλγόριθμους ταξινόμησης με εκατό διαφορετικά προβλήματα ταξινόμησης. Οι σχετικές επιδόσεις των αλγόριθμων έδειξαν ότι δεν υπήρχε κάποιος ενιαίος αλγόριθμος ταξινόμησης για την επίλυση όλων των εκατοντάδων προβλημάτων ταξινόμησης. Επιπλέον, η διάδοση μεγάλων συνόλων δεδομένων σε πολλές περιοχές δημιούργησε αρκετές προκλήσεις για την εξαγωγή των δεδομένων. Έτσι μέχρι σήμερα η δημιουργία του κατάλληλου αλγορίθμου παραμένει υπό διερεύνηση.

Η ταξινόμηση των ιατρικών εικόνων μπορεί να πραγματοποιείται με εποπτευόμενη ή με μη εποπτευόμενη μέθοδο, που βασίζεται σε μοντέλα κατανομής πιθανοτήτων που θα μπορούσαν να είναι παραμετρικά ή μη παραμετρικά όπως ο ευκλείδειος ταξινομητής, το k-nearest, η ελάχιστη απόσταση και δέντρο απόφασης κ.λπ.

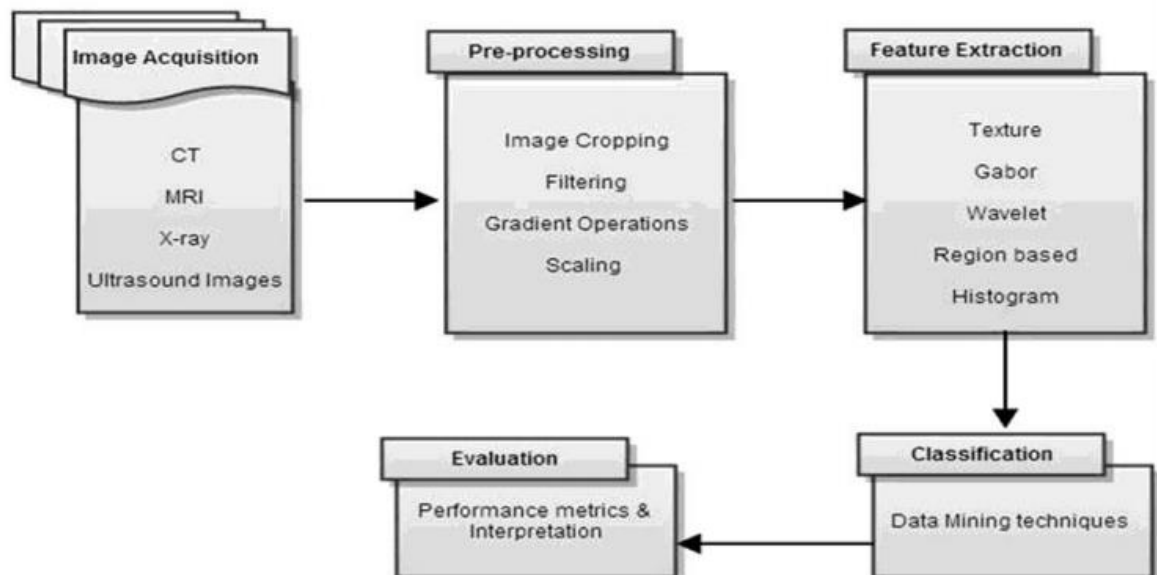
Παρόλο που στην εποπτευόμενη ταξινόμηση παρέχεται μια συλλογή ετικετών (προ-ταξινομημένων) εικόνων, το πρόβλημα είναι να επισημανθούν οι νεοεμφανιζόμενες, μη επισημασμένες εικόνες. Σε γενικές γραμμές, στην προκαθορισμένη σήμανση (συνήθως ονομάζεται εκπαιδευτικό σετ) οι εικόνες χρησιμοποιούνται για να κάνουν την εκμάθηση της

μηχανής της κατηγορίας (ομάδα), η οποία με τη σειρά της χρησιμοποιείται για την επόμενη άγνωστη εικόνα [36].

Η Εικ.3.6 δείχνει τη διαδικασία ταξινόμησης των ιατρικών εικόνων. Η διαδικασία ταξινόμησης περιλαμβάνει πέντε κύρια βήματα τα οποία είναι η απόκτηση εικόνας, η προεπεξεργασία, η εξαγωγή χαρακτηριστικών, η ταξινόμηση και η αξιολόγηση. Το βήμα της απόκτησης περιλαμβάνει την επιλογή εικόνων που κυμαίνονται από υπολογιστικές τομογραφίες (CT), εικόνες μαγνητικού συντονισμού (Μαγνητική τομογραφία) σε ακτίνες X κ.λπ.

Η προεπεξεργασία είναι μια πορεία δράσεων που εκτελείται σε ανεπεξέργαστα δεδομένα με στόχο τη βέλτιστη αιτιολογική σκέψη για αυτά τα σύνολα δεδομένων. Η διαδικασία της προεπεξεργασίας έχει σημαντικό αντίκτυπο στην απόδοση του αλγορίθμου ταξινόμησης. Η φάση της προεπεξεργασίας των δεδομένων περιλαμβάνει την περικοπή εικόνων, το φιλτράρισμα, τη διόρθωση της κλίσης και της κλίμακας.

Η εξαγωγή χαρακτηριστικών εμπλέκει την εκτίμηση χαρακτηριστικών και τις μεθόδους επιλογής χαρακτηριστικών. Ένας σημαντικός αριθμός χαρακτηριστικών είναι διαθέσιμος στη βιβλιογραφία όπως υφή, gabor, ιστόγραμμα wavelet, τα οποία κάθε ένα από αυτά περιγράφει ορισμένες πτυχές του περιεχομένου της εικόνας.



Εικόνα 3.6. Διαδικασία ταξινόμησης ιατρικών εικόνων [37].

Ως εκ τούτου, η εξαγωγή χαρακτηριστικών αντικειμένων και εικόνων είναι μια διαδικασία ανάλυσης . Αξίζει να αναφερθεί ότι η βελτίωση της διαδικασίας εξαγωγής χαρακτηριστικών μπορεί να βελτιωθεί με την αύξηση της απόδοσης ενός αλγορίθμου ταξινόμησης.

Γενικότερα κατά τη διαδικασία της ταξινόμησης χρησιμοποιούνται διάφορες τεχνικές εξαγωγής δεδομένων [38]. Ενώ, για όλα τα προβλήματα ταξινόμησης, η κύρια πηγή αξιολόγησης της ταξινόμησης είναι ένας πίνακας συμπτώσεων ή πίνακας έκτακτης ανάγκης ο οποίος επικυρώνει την απόδοση της μεθόδου ταξινόμησης.

Με την πρόοδο στην εξαγωγή δεδομένων που αφορούν στη διάγνωση και στην πρόγνωση διαφόρων ασθενειών, έχει προταθεί ένας σημαντικός αριθμός αλγορίθμων για ευρεία ποικιλία ταξινόμησης ιατρικών εικόνων .

Για παράδειγμα, η ταξινόμηση με βάση την υφή μπορεί να εφαρμοστεί σε οποιαδήποτε μορφή ψηφιακής εικόνας και βοηθάει στην απόκτηση φασματικών ιδιοτήτων μιας εικόνας [39]. Επιπλέον, η ταξινόμηση των χαρακτηριστικών υφής μπορεί να είναι χρήσιμη ως βοήθεια στην κλινική ερμηνεία του ακτινολόγου, ακθώς μπορεί να περιλαμβάνει το διαχωρισμό του επιλεγμένου χώρου λειτουργιών σύμφωνα με την τάξη/κατηγορία των ιστών.

Σήμερα, για την περιγραφή της υφής μιας εικόνας χρησιμοποιείται μια μεγάλη ποικιλία τεχνικών . Οι μέθοδοι ανάλυσης της υφής μπορούν να χωριστούν σε τέσσερις όπως, στατιστικές, γεωμετρικές, βασισμένη σε μοντέλα και στην επεξεργασία σήματος [39]. Στην επόμενη υποενότητα αναφέρονται συνοπτικά ορισμένοι αλγόριθμοι ταξινόμησης.

3.2.1 Αλγόριθμοι ταξινόμησης

Ο k-Nearest Neighbor (k-NN) είναι ένας μη παραμετρικός αλγόριθμος. Γενικότερα ο αλγόριθμος αποθηκεύει πρώτα τα χαρακτηριστικά διανύσματα για το σετ εκπαίδευσης και στη συνέχεια για την ταξινόμηση ενός νέου στιγμιότυπου, βρίσκει ένα σύνολο k πλησιέστερων παραδειγμάτων εκπαίδευσης και τότε ορίζει το παράδειγμα στην κλάση που έχει περισσότερα παραδείγματα στο δεδομένο σετ [40].

Επιπλέον αυτός μπορεί να χρησιμοποιηθεί τόσο για προβλήματα ταξινόμησης όσο και για προβλήματα παλινδρόμησης. Ωστόσο, χρησιμοποιείται ευρύτερα σε προβλήματα ταξινόμησης γιατί είναι ένας απλός αλγόριθμος που αποθηκεύει όλες τις διαθέσιμες περιπτώσεις και ταξινομεί νέες περιπτώσεις με βάση την πλειοψηφία των γειτονικών του εικόνων.

Αυτές οι λειτουργίες απόστασης μπορούν να είναι η απόσταση Euclidean, Manhattan, Minkowski και Hamming. Οι πρώτες τρεις λειτουργίες χρησιμοποιούνται για τις συνεχείς μεταβλητές και η τέταρτη (Hamming) για τις κατηγορικές μεταβλητές. Αν $K = 1$, τότε η περίπτωση απλά ανατίθεται στην κατηγορία του πλησιέστερου 'γείτονα'. Κατά καιρούς, η επιλογή του K αποδεικνύεται ως πρόκληση κατά την εκτέλεση του μοντέλου kNN.

Οι Suguna & Thanushkodi [41], χρησιμοποίησαν ένα βελτιωμένο k-NN χρησιμοποιώντας γενετικό αλγόριθμο για τη μείωση της υψηλής πολυπλοκότητας, με χαμηλή εξάρτηση από το σετ εκπαίδευσης και καμία διαφορά βάρους μεταξύ κάθε κατηγορίας. Οι Latifoglu et al. [42] πρότειναν ιατρικό σύστημα βασισμένο στην ανάλυση κύριων συστατικών (PCA), προεπεξεργασία βάρους k-NN και τεχνητό σύστημα ανοσολογικής αναγνώρισης (AIRS) για τη διάγνωση της νόσου αθηροσκλήρωσης.

Ωστόσο, υπάρχουν ορισμένοι περιορισμοί αυτού του αλγορίθμου: μπορεί να αποθηκεύσει μόνο τις τοπικές πληροφορίες με υψηλή πολυπλοκότητα υπολογισμού και χρειάζεται περισσότερος χρόνος για τον υπολογισμό του νέου ερωτήματος. Μπορεί να χειρίζεται δυαδικά και συνεχή χαρακτηριστικά αλλά όχι άμεσα διακριτά [43]. Ωστόσο, λίγες πρόσφατες μελέτες προσπαθούν να ξεπεράσουν τον περιορισμό των παραδοσιακών kNN, για παράδειγμα οι εργασίες [42, 44, 45] είναι σε θέση να παράγουν καλύτερα αποτελέσματα.

Το SVM, με τη σειρά του είναι μια θεωρία στατιστικής μάθησης για την ανάλυση δεδομένων και την αναγνώριση προτύπων και αποτελεί μια επιβλεπόμενη μέθοδο μάθησης. Τα πλεονεκτήματα του SVM είναι ότι μπορεί να χειριστεί συνεχή και δυαδικά χαρακτηριστικά με καλή ταχύτητα ταξινόμησης και καλή ακρίβεια. Αλλά υπάρχουν λίγα μειονεκτήματα όπως: Το SVM διαρκεί περισσότερο χρόνο για την κατάρτιση του συνόλου δεδομένων και δεν χειρίζεται χωριστά χαρακτηριστικά [46].

Επίσης αναφέρεται ότι σε αυτόν τον αλγόριθμο, κάθε στοιχείο δεδομένων σχεδιάζεται ως σημείο σε n -διάστατο χώρο (όπου n είναι ο αριθμός των χαρακτηριστικών που υπάρχουν) με την τιμή κάθε χαρακτηριστικού να είναι η τιμή μιας συγκεκριμένης συντεταγμένης. Για παράδειγμα, αν υπήρχαν μόνο δύο χαρακτηριστικά θα έπρεπε πρώτα να σχεδιαστούν αυτές οι δύο μεταβλητές σε δισδιάστατο χώρο όπου κάθε σημείο έχει δύο συντεταγμένες (αυτές οι συντεταγμένες είναι γνωστές ως φορείς υποστήριξης).

Οι Kharrat et al.[46] πρότειναν μια προσέγγιση για την ταξινόμηση των MRI εγκεφάλου (απεικόνιση μαγνητικού συντονισμού) χρησιμοποιώντας γενετικό αλγόριθμο με SVM ικανό

για την ταξινόμηση του εγκεφαλικού ιστού σε φυσιολογικό, καλοήγη ή κακοήγη όγκο. Ωστόσο, το SVM τείνει να αποδίδει πολύ καλύτερα όταν ασχολείται με πολυδιάστατες και συνεχείς λειτουργίες. Επιπλέον, απαιτείται μεγάλο μέγεθος δείγματος προκειμένου να επιτύχει η μέγιστη ακρίβεια πρόβλεψης.

Ο αλγόριθμος Naive-Bayes Πρόκειται για μια τεχνική ταξινόμησης βασισμένη στο θεώρημα του Bayes με την παραδοχή της ανεξαρτησίας μεταξύ των προγνωστικών. Πιο απλά, ένας ταξινομητής Naive Bayes υποθέτει ότι η παρουσία ενός συγκεκριμένου χαρακτηριστικού σε μια τάξη δεν σχετίζεται με την παρουσία οποιουδήποτε άλλου χαρακτηριστικού [40].

Για παράδειγμα, ένας καρπός μπορεί να θεωρηθεί κάτι άλλο αν έχει τα χαρακτηριστικά άλλου αντικειμένου. Ακόμη και αν αυτά τα χαρακτηριστικά εξαρτώνται το ένα από το άλλο ή από την ύπαρξη των άλλων χαρακτηριστικών, ένας ταξινομητής Bayes θα θεωρούσε όλες αυτές τις ιδιότητες να συμβάλλουν ανεξάρτητα στην πιθανότητα ότι αυτός ο καρπός είναι κάποιο διαφορετικό αντικείμενο.

Ένα Bayesian μοντέλο είναι εύκολο να οικοδομηθεί και ιδιαίτερα χρήσιμο για πολύ μεγάλα σύνολα δεδομένων. Μαζί με την απλότητα που το διακατέχει, το Naive Bayes είναι γνωστό ότι ξεπερνάει ακόμη και πολύ εξελιγμένες μεθόδους ταξινόμησης [40].

Ένα άλλο είδος είναι η γραμμική παλινδρόμηση, η οποία χρησιμοποιείται για την εκτίμηση των πραγματικών τιμών με βάση τις συνεχείς μεταβλητές. Με βάση τη θεωρία δημιουργείται μια σχέση μεταξύ ανεξάρτητων και εξαρτημένων μεταβλητών για την εξαγωγή γραμμής προσαρμογής. Αυτή η γραμμή καλύτερης προσαρμογής είναι γνωστή ως γραμμή παλινδρόμησης και αντιπροσωπεύεται από μια γραμμική εξίσωση $Y = a * X + b$ [40].

Το δέντρο αποφάσεων αποτελεί έναν άλλον αλγόριθμο κατηγοριοποίησης, ο οποίος χρησιμοποιείται αρκετά συχνά. Είναι ένας τύπος αλγόριθμου επιβλεπόμενης μάθησης που χρησιμοποιείται κυρίως για προβλήματα ταξινόμησης και λειτουργεί τόσο για κατηγορηματικές όσο και για συνεχείς εξαρτώμενες μεταβλητές. Σε αυτόν τον αλγόριθμο, ο πληθυσμός χωρίζεται σε δύο ή περισσότερα ομοιογενή σύνολα. Αυτό γίνεται με βάση τις πιο σημαντικές ιδιότητες/ανεξάρτητες μεταβλητές, ώστε να γίνονται όσο το δυνατόν διακριτές ομάδες [40].

Ο K-Means είναι ένας τύπος αλγορίθμου χωρίς επίβλεψη ο οποίος επιλύει το πρόβλημα της ομαδοποίησης. Η διαδικασία ακολουθεί έναν απλό και εύκολο τρόπο ταξινόμησης ενός δεδομένου συνόλου δεδομένων μέσω ενός ορισμένου αριθμού συμπλεγμάτων (υποθέτουμε k

clusters). Τα σημεία δεδομένων μέσα σε ένα σύμπλεγμα είναι ομοιογενή και ετερογενή σε ομάδες ομοτίμων [40].

Με βάση αυτόν τον αλγόριθμο υπάρχουν clusters και κάθε cluster έχει το δικό του centroid. Το άθροισμα της τετραγωνικής διαφοράς μεταξύ του κεντροειδούς (centroid) και των σημείων δεδομένων εντός ενός συμπλέγματος αποτελεί το άθροισμα της τετραγωνικής τιμής για το σύμπλεγμα.

Επίσης, όταν προστεθεί το άθροισμα των τετραγωνικών τιμών για όλα τα συμπλέγματα, γίνεται συνολικό στο άθροισμα της τετραγωνικής τιμής για τη λύση συμπλέγματος. Καθώς λοιπόν ο αριθμός των συστοιχιών αυξάνεται, αυτή η τιμή συνεχίζει να μειώνεται, αλλά αν σχεδιαστεί το αποτέλεσμα, το άθροισμα της τετραγωνικής απόστασης μειώνεται απότομα μέχρι κάποια τιμή του k βρίσκοντας το βέλτιστο αριθμό συμπλεγμάτων [40].

Το τυχαίο δάσος είναι ένας αλγόριθμος που χρησιμοποιείται για ένα σύνολο δέντρων αποφάσεων. Στο τυχαίο δάσος, έχουμε συλλογή δέντρων αποφάσεων (γνωστών ως "Δάσος") και για να ταξινομηθεί ένα νέο αντικείμενο με βάση τα χαρακτηριστικά, κάθε δέντρο δίνει μια ταξινόμηση. Το δάσος επιλέγει την ταξινόμηση με τις περισσότερες ψήφους (πάνω από όλα τα δέντρα στο δάσος).

Εάν ο αριθμός των περιπτώσεων στο σετ εκπαίδευσης είναι N , τότε το δείγμα των περιπτώσεων N λαμβάνεται τυχαία αλλά με αντικατάσταση. Αυτό το δείγμα θα είναι το σετ κατάρτισης για την 'ανάπτυξη' του δέντρου. Αν υπάρχουν μεταβλητές εισόδου M , ένας αριθμός $m \ll M$ καθορίζεται έτσι ώστε σε κάθε κόμβο, οι μεταβλητές m να επιλέγονται τυχαία εκτός από το M . Ο καλύτερος διαχωρισμός σε αυτά τα m χρησιμοποιείται για τον διαχωρισμό του κόμβου. Η τιμή του m διατηρείται σταθερή κατά τη διάρκεια της δασικής ανάπτυξης και κάθε δέντρο αναπτύσσεται στο μεγαλύτερο δυνατό βαθμό [40].

Τα νευρωνικά δίκτυα όπως έχει ήδη αναφερθεί έχουν αναδειχθεί ως ένα σημαντικό εργαλείο για την ταξινόμηση. Πρόκειται για υπολογιστικό μοντέλο των οποίων η αρχιτεκτονική είναι ικανή για οποιαδήποτε κατά προσέγγιση λειτουργία και ως εκ τούτου, θεωρούνται καλή επιλογή όταν η λειτουργία που πρέπει να μάθει δεν είναι γνωστή εκ των προτέρων.

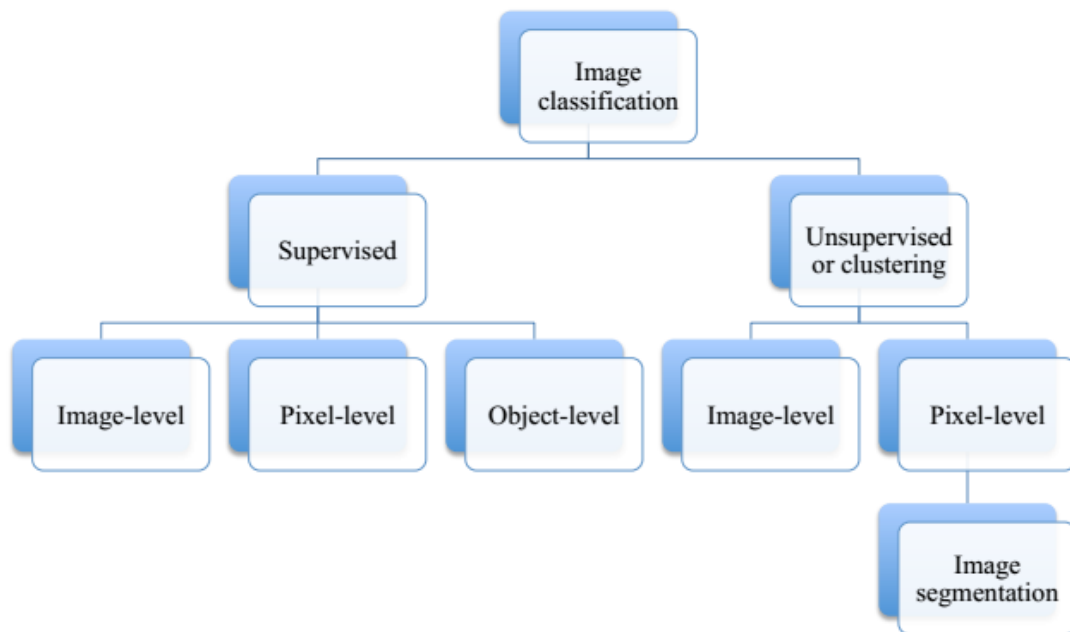
Με την ταχεία λοιπόν ανάπτυξη της πληροφορικής και της ιατρικής απεικόνισης, οι τεχνικές ιατρικής απεικόνισης, μπορούν να απεικονίσουν χωρίς να υπάρχει κάποια επέμβαση, την φυσιολογική κατάσταση ιστών και οργάνων στο ανθρώπινο σώμα. Αυτές οι τεχνικές έχουν γίνει σταδιακά απαραίτητα εργαλεία στην ιατρική έρευνα, στην κλινική διάγνωση και στον

χειρουργικό σχεδιασμό [47]. Ενώ αυτές οι νέες τεχνολογίες έχουν προηγμένη ιατρική θεωρία και πρακτική, έθεσαν επίσης νέα ζητήματα. Για παράδειγμα, οι γιατροί πρέπει τώρα να ταξινομήσουν διαγνωστικά αποτελέσματα.

Οι τεχνικές αυτόματης ταξινόμησης εικόνων μπορούν να κατανοήσουν σε κάποιο βαθμό το περιεχόμενο της εικόνας. για παράδειγμα, μπορούν να εντοπίσουν αποτελεσματικά τις περιοχές βλάβης στις ιατρικές εικόνες και να βοηθήσουν τους γιατρούς στην εκτέλεση αποτελεσματικών διαγνώσεων [48].

Δηλαδή, η αποτελεσματική ταξινόμηση των ιατρικών εικόνων παίζει σημαντικό ρόλο στην υποβοήθηση της κλινικής περίθαλψης και θεραπείας. Για παράδειγμα, η ανίχνευση ακτίνων Χ είναι η καλύτερη προσέγγιση για τη διάγνωση της πνευμονία, η οποία είναι υπεύθυνη για το θάνατο πολλών ανθρώπων. Ωστόσο η ταξινόμηση των ακτινογραφιών θώρακα για παράδειγμα απαιτεί επαγγελματίες ακτινολόγους, οι οποίοι σε ορισμένες περιπτώσεις είναι σπάνιοι για ορισμένες περιοχές [49].

Ωστόσο, υπάρχουν πολλοί τύποι ιατρικών εικόνων και η διάκριση των κατηγορηματικών πληροφοριών σε αυτές τις ιατρικές εικόνες απαιτεί συχνά διαφορετικές μεθόδους επεξεργασίας και ανάλυσης [50]. Μέχρι σήμερα, η ταξινόμηση των ιατρικών εικόνων βασίζεται κυρίως σε μεθόδους αναγνώρισης προτύπων, στις οποίες έχει εκπαιδευτεί ένα μοντέλο ταξινόμησης για τον εντοπισμό και τη διάκριση ιατρικών εικόνων. Η ταξινόμηση της ιατρικής εικόνας μπορεί γενικά να διαιρεθεί σε επιβλεπόμενες μεθόδους ταξινόμησης και μεθόδους ταξινόμησης χωρίς επίβλεψη (Εικ.3.7).



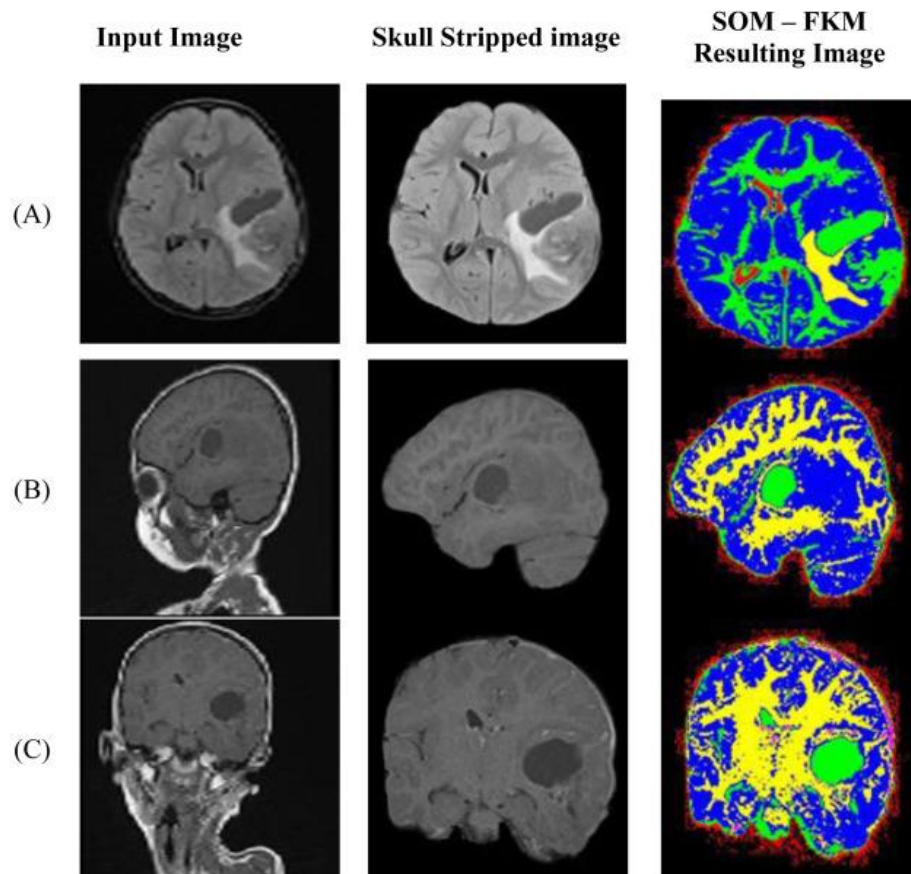
Εικόνα 3.7. Είδη ταξινόμησης εικόνων ιατρικής απεικόνισης [51].

Κατά συνέπεια, η ελεγχόμενη ταξινόμηση μπορεί να είναι σε επίπεδο: (i) εικόνας, (ii) περιοχής ή (iii) εικονοστοιχείων. Αντιθέτως, η μη επιτηρούμενη ταξινόμηση εικόνας μπορεί να είναι σε επίπεδο: (i) εικόνας ή (ii) εικονοστοιχείων.

3.2.2 Μη εποπτευόμενη ταξινόμηση

Όταν πραγματοποιείται μη εποπτευόμενη ταξινόμηση ή ομαδοποίηση σε επίπεδο εικόνας δίνει μια σειρά εικόνων των οποίων οι κατηγορίες είναι άγνωστες. Ο στόχος είναι να βρεθεί ομάδα εικόνων με παρόμοια χαρακτηριστικά που καθορίζουν τις κατηγορίες. Στην Εικ.3.8 απεικονίζεται η διαδικασία της μη επιτηρούμενης ταξινόμησης ή της ομαδοποίησης ενός συνόλου εικόνων που αποτελείται από έναν αριθμό στοιχείων.

Segmentation Results from Hybrid SOM – FKM algorithm:



Εικόνα 3.8. Μη εποπτευόμενη ταξινόμηση [52].

Η μη εποπτευόμενη ταξινόμηση σε επίπεδο εικονοστοιχείων χρησιμοποιείται για την ομαδοποίηση των εικονοστοιχείων εικόνας με βάση παρόμοια χαρακτηριστικά, προκειμένου να βρεθούν ομοιόμορφες περιοχές εικόνας. Αυτό είναι ένα σημαντικό βήμα για την κατάτμηση μιας εικόνας. Γενικότερα, η ομαδοποίηση εικόνας χωρίς εποπτεία σε επίπεδο εικονοστοιχείων καθορίζει την κατάτμηση της εικόνας [52].

Επομένως, οι μη επιβλεπόμενες μέθοδοι ταξινόμησης με βάση τις ομοιότητες μεταξύ των δειγμάτων, χωρίς να απαιτούν προηγούμενα δείγματα, διαχωρίζουν αυτόματα τις διαφορετικές κατηγορίες. Δηλαδή, η μη εποπτευόμενη μάθηση είναι ουσιαστικά μια διαδικασία ομαδοποίησης. Οι τυπικές μη επιτηρούμενες μέθοδοι περιλαμβάνουν τη K-means clustering [53], τη συσσωμάτωση των ασαφών μέσω C (fuzzy C-means clustering) [54] και την ανάλυση βασικών συστατικών (principal component analysis, PCA) [55].

Για να επιλυθεί το πρόβλημα ότι τα περιγράμματα των ‘βλαβών’ δεν μπορούν να βρεθούν με ακρίβεια σε εικόνες, το 2010, ο Χουάν πρότεινε έναν αλγόριθμο κατακερματισμού χρώματος

με μια τεχνική ομαδοποίησης K-μέσων που προσέθεσε μια λειτουργία κατάτμησης με βάση το χρώμα [53].

Το 2012, οι Singh και Kaur για την αυτόματη ταξινόμηση των μαγνητικών τομογραφιών και των φυσικών εικόνων, πρότειναν μια μέθοδο που βασίζεται σε PCA. Τα πειράματά τους έδειξαν ότι η ακρίβεια ταξινόμησης έφτασε το 91% [55]. Αν και αυτή η μέθοδος είναι σχετικά απλή στην εφαρμογή, έχει επιτύχει συγκεκριμένα αποτελέσματα στις εφαρμογές αναγνώρισης εικόνας και ταξινόμησης. Ωστόσο, η ταξινομητική επίδραση αυτού του είδους της μεθόδου έχει μεγάλες διαφορές κατά την ταξινόμηση και αναγνώριση διαφορετικών ιατρικών εικόνων και δεν μπορεί να ταξινομή και να αναγνωρίζει προσαρμοστικά τις ιατρικές εικόνες βάσει των χαρακτηριστικών τους.

3.2.3 Εποπτευόμενη ταξινόμηση

Στην εποπτευόμενη ταξινόμηση σε επίπεδο εικόνας, δίνεται ένα σύνολο εικόνων κάθε μία από τις οποίες ανήκει σε μια συγκεκριμένη κατηγορία που αντιστοιχεί στο σετ εκπαίδευσης. Ο ταξινομητής μαθαίνει τις κατηγορίες εικόνων χρησιμοποιώντας το σετ εκπαίδευσης. Τότε, είναι σε θέση να γενικεύσει την ταξινόμηση κατηγοριοποιώντας εικόνες που δεν περιλαμβάνονται στο σετ εκπαίδευσης [52].

Η εποπτευόμενη ταξινόμηση σημαίνει ότι τα επεξεργασμένα δείγματα έχουν επισημανθεί εκ των προτέρων και το μοντέλο ταξινόμησης εκπαιδεύεται χρησιμοποιώντας την ετικέτα των χαρακτηριστικών της εικόνας, και οι κατηγορίες ταξινόμησης είναι συνήθως προκαθορισμένες.

Οι εποπτευόμενες μέθοδοι ταξινόμησης περιλαμβάνουν κυρίως τον K-πλησιέστερο αλγόριθμο γειτονίας [56], Bayesian μοντέλα [57], λογιστική παλινδρόμηση [58], νευρωνικά δίκτυα [59] και μηχανήματα υποστήριξης διανυσμάτων [60]. Η χρήση των παραδοσιακών μεθόδων μηχανικής μάθησης, όπως οι μέθοδοι φορέα διανυσμάτων, στην ταξινόμηση των ιατρικών εικόνων, άρχισε εδώ και καιρό. Ωστόσο, οι μέθοδοι αυτές έχουν τα ακόλουθα μειονεκτήματα: η απόδοση απέχει πολύ από το πρακτικό πρότυπο και η ανάπτυξή τους τα τελευταία χρόνια είναι αρκετά αργή. Επίσης, η εξαγωγή και η επιλογή χαρακτηριστικών είναι χρονοβόρες διαδικασίες και ποικίλλουν ανάλογα με διαφορετικά αντικείμενα.

Μεταξύ λοιπόν των προαναφερόμενων μεθόδων, οι μέθοδοι ταξινόμησης ιατρικών εικόνων που βασίζονται σε νευρωνικά δίκτυα τείνουν να αποδίδουν καλύτερα [61]. Για παράδειγμα, ο Abbass [62] πρότεινε ένα νευρωνικό δίκτυο βασισμένο στον αλγόριθμο διαφορικής

εξέλιξης Pareto για την ταξινόμηση του καρκίνου του μαστού στη Βάση Δεδομένων του Καρκίνου του Μαστού Wisconsin (WBCD) που απέκτησε καλύτερη απόδοση ταξινόμησης από εκείνη των παραδοσιακών νευρωνικών δικτύων.

Οι Karabatak και Ince [63] χρησιμοποίησαν κανόνες σύνδεσης για τη μείωση των διαστάσεων, μειώνοντας τα εννέα χαρακτηριστικά σε τέσσερα και στη συνέχεια τα ταξινόμησαν χρησιμοποιώντας ένα τεχνητό νευρωνικό δίκτυο. Η ακρίβεια της διπλής εγκυρότητας στον καρκίνο του μαστού στο WBCD έφτασε το 90%. Ωστόσο, αυτό το είδος της μεθόδου δεν μπορεί να προσαρμοστεί στις χαρακτηριστικές πληροφορίες στις ίδιες τις ιατρικές εικόνες, γεγονός που οδηγεί σε μεγάλες διαφορές στην ταξινόμηση σε διαφορετικές ιατρικές εικόνες.

Τα βαθιά νευρωνικά δίκτυα (Deep Neural Networks), ειδικά τα συνελκτικά νευρωνικά δίκτυα, χρησιμοποιούνται ευρέως στις εργασίες ταξινόμησης εικόνων και έχουν επιτύχει σημαντικές επιδόσεις από το 2012 [64]. Ορισμένες έρευνες για την ταξινόμηση ιατρικών εικόνων από CNN έχουν επιτύχει επιδόσεις που ανταγωνίζονται ιατρικούς εμπειρογνώμονες. Για παράδειγμα, το CheXNet, ένα CNN με 121 στρώματα που εκπαιδεύτηκε σε σύνολο δεδομένων με περισσότερες από 100.000 ακτίνες X στο στήθος (ChestX-ray 14), επέτυχε καλύτερη απόδοση από τη μέση απόδοση τεσσάρων ακτινολόγων [64].

Επιπλέον, οι Kermany et al. [65] προτείνουν ένα σύστημα εκμάθησης μεταφοράς για ταξινόμηση 108.309 εικόνων Οπτικής Ομοιογενούς Συνοχής (OCT) και το σταθμισμένο μέσο σφάλμα είναι ίσο με τη μέση απόδοση 6 εμπειρογνομώνων ιατρών.

Οι Li et al. [66] παρουσίασαν ένα εξατομικευμένο CNN με shallow ConvLayer για να ταξινομήσουν τις επιφάνειες της εικόνας με ασθένειες των πνευμόνων. Οι ερευνητές διαπίστωσαν ότι το σύστημα μπορεί να γενικευθεί σε άλλα σύνολα δεδομένων ιατρικών εικόνων.

Επιπλέον, αυτοί σε μια άλλη έρευνά τους, διαπίστωσαν επίσης ότι το σύστημα που βασίζεται στο CNN μπορεί να εκπαιδευτεί από μεγάλο σύνολο δεδομένων ακτίνων X (CXR) και από τεχνολογία αιχμής με υψηλή ακρίβεια και ευαισθησία στο σύνολο δεδομένων, όπως το Stanford Normal Radiology Diagnostic Dataset που περιέχει περισσότερα από 400.000 CXR και μια νέα βάση δεδομένων CXR (ChestX-ray8), τα οποία αποτελούνται από 108.948 CXR frontal-view [67].

Όπως έχει προαναφερθεί η χρήση περιορισμένων δεδομένων καθιστά δύσκολη τη διαδικασία εκπαίδευσης ενός μοντέλου. Συνεπώς, η εκμάθηση μεταφοράς του CNN χρησιμοποιείται αρκετά σε εργασίες ταξινόμησης ιατρικών εικόνων.

Οι Kermany et al.[65] χρησιμοποίησαν το InceptionV3 με το εκπαιδευμένο ImageNet και μετέφεραν τη μάθηση σε ένα σύνολο δεδομένων ιατρικών εικόνων που περιείχε 108.312 εικόνες οπτικής συνοπτικής τομογραφίας (OCT). Έλαβαν μέση ακρίβεια 96,6%, με ευαισθησία 97,8% και ακρίβεια 97,4%. Οι ερευνητές συνέκριναν επίσης τα αποτελέσματα με αυτά που έδωσαν έξι ειδικούς. Οι περισσότεροι από τους ειδικούς παρείχαν ανάλυση υψηλής ευαισθησία, αλλά χαμηλή ακρίβεια. Ενώ το σύστημα με βάση το CNN έλαβε υψηλές τιμές τόσο για την ευαισθησία όσο και για την ακρίβεια.

Επιπλέον, με βάση τη μέση βαρύτητα σφάλματος, το σύστημα που βασίζεται στο CNN υπερβαίνει τους δύο εμπειρογνώμονες. Οι ερευνητές επίσης επαλήθευσαν το σύστημά τους σε ένα σύνολο δεδομένων για την πνευμονία, συμπεριλαμβανομένων περίπου πέντε χιλιάδων εικόνων, και πέτυχαν μια μέση ακρίβεια 92,8%, με ευαισθησία 93,2% και ακρίβεια 90,1%. Τα συνελκτικά επομένως νευρωνικά δίκτυα, μπορούν τελικά να βοηθήσουν στην επιτάχυνση της διάγνωσης και της παραπομπής των ασθενών και ως εκ τούτου να εισαγάγει την έγκαιρη θεραπεία, με αποτέλεσμα την αύξηση του ποσοστού θεραπείας.

Γενικότερα οι ιατρικές εικόνες είναι δύσκολο να συγκεντρωθούν, καθώς η συλλογή και επισήμανση (labelling) των ιατρικών δεδομένων έρχεται αντιμέτωπη τόσο με τις ανησυχίες περί προστασίας της ιδιωτικής ζωής όσο και με την απαίτηση για χρονοβόρες εξηγήσεις εμπειρογνομώνων.

Στις δύο γενικές κατευθύνσεις επίλυσης, είναι η συλλογή περισσότερων δεδομένων, όπως το crowdsourcing [68] ή η διερεύνηση στις υπάρχουσες κλινικές αναφορές [69]. Ένας άλλος τρόπος είναι να μελετηθούν τρόποι αύξησης της απόδοσης ενός μικρού συνόλου δεδομένων, το οποίο είναι πολύ σημαντικό επειδή οι γνώσεις που επιτυγχάνονται από την έρευνα μπορούν να εφαρμοστούν σε μεγάλα σύνολα δεδομένων. Μέχρι σήμερα το σημαντικότερο δημοσιευμένο σύνολο δεδομένων εικόνων ακτίνων X του θώρακα (ChestX-ray 14) εξακολουθεί να είναι πολύ μικρότερο από το μεγαλύτερο σύνολο δεδομένων εικόνας ImageNet το οποίο έφτασε τις 14.197.122 περιπτώσεις το 2010 [70].

Οι μέθοδοι που βασίζονται σε CNN έχουν διάφορες στρατηγικές για την αύξηση της απόδοσης της ταξινόμησης εικόνων σε μικρά σύνολα δεδομένων: Μια μέθοδος είναι η

αύξηση των δεδομένων [71]. Οι Wang και Perez [72] ερεύνησαν την αποτελεσματικότητα της αύξησης των δεδομένων στην ταξινόμηση εικόνων. Οι συγγραφείς διαπίστωσαν ότι η παραδοσιακή αύξηση δεδομένων με βάση τα μετασχηματισμένα κελιά έχει καλύτερη απόδοση από το γενετικό ανταγωνιστικό δίκτυο (GAN) και άλλες μεθόδους βασισμένες στο νευρωνικό δίκτυο.

Μια άλλη μέθοδος είναι η εκμάθηση μεταφοράς. Οι Kermany et al. [65] πέτυχαν 92% ακρίβεια σε ένα σύνολο δεδομένων για εικόνα πνευμονίας με ακτίνες X μέσω της μεταφοράς δεδομένων. Η τρίτη μέθοδος είναι το capsule network. Οι Sabour et al. [73] δημιούργησαν ένα νέο δίκτυο νευρωνικού δικτύου δομής-κάψουλας, το οποίο επιτυγχάνει σύγχρονες επιδόσεις στην βάση δεδομένων του τροποποιημένου Εθνικού Ινστιτούτου Προτύπων και Τεχνολογίας (MNIST), η οποία θεωρείται και η καλύτερη απόδοση σε άλλα μικρά σύνολα δεδομένων [74]. Οι Afshar et al. [75] χρησιμοποίησαν την ίδια μέθοδο για την ανίχνευση όγκων στον εγκέφαλο, με ακρίβεια 86,56%.

Οι Parveen και Sathik [43] ερεύνησαν την ανίχνευση της πνευμονίας από ακτινογραφίες X. Οι ερευνητές εξήγαγαν τα χαρακτηριστικά με διακριτό μετασχηματισμό κύματος (discrete wavelet transform, DWT), μετασχηματισμό πλαισίου κύματος (wavelet frame transform, WFT), μετασχηματισμό πακέτων wavelet (wavelet packet transform, WPT) και χρησιμοποιημένα μέσα Fuzzy C.

Οι Caicedo et al. [44] χρησιμοποίησαν μετασχηματισμό μεταβλητών χαρακτηριστικών κλίμακας (scale-invariant feature transform, SIFT) ως τοπικό περιγραφικό χαρακτηριστικό και ταξινομητές μηχανών φορέα υποστήριξης (support vector machines, SVM) για ταξινόμηση ιατρικών εικόνων. Οι εικόνες των ακτινογραφιών ταξινομήθηκαν με απόδοση ακρίβειας κατά 67%

Ωστόσο, το SIFT είναι αλγόριθμος ευρεσιτεχνίας. Έτσι, οι Rublee et al. [45] πρότειναν ένα ελεύθερο, ταχύτερο και πιο τοπικό δυαδικό αξιόπιστο ανεξάρτητο στοιχειώδες χαρακτηριστικό γνώρισμα (oriented fast and rotated binary robust, ORB), το οποίο έχει τις ίδιες επιδόσεις με το SIFT και ακόμη καλύτερη απόδοση από το SIFT, υπό ορισμένες προϋποθέσεις. Το SVM είναι επίσης ένας αλγόριθμος κατηγοριοποίησης υψηλής απόδοσης, που χρησιμοποιείται ευρέως σε διάφορες εργασίες ταξινόμησης ιατρικών εικόνων και από άλλους ερευνητές και επιτυγχάνει εξαιρετική απόδοση [46].

3.3 Ταξινόμηση εικόνων ακτινολογικών δεδομένων

Η ταξινόμηση των εικόνων με ανατομικά ή παθολογικά χαρακτηριστικά αποτελεί θεμελιώδη γνωστική λειτουργία στην διαγνωστική ακτινολογία. Παρόλο που οι υπολογιστές δεν έχουν πλέον τη δυνατότητα να αναπαράγουν την πλήρη αλυσίδα αιτιολογίας που απαιτείται για την ερμηνεία των ιατρικών εικόνων, ο αυτοματισμός της βασικής ταξινόμησης εικόνων αποτελεί το επίκεντρο της έρευνας.

Μια ψηφιακή εικόνα, όπως αναφέρθηκε αναλυτικά στο Κεφάλαιο 2 μπορεί να θεωρηθεί ως ένας πίνακας αριθμών που κωδικοποιούν τη φωτεινότητα και το χρώμα των επιμέρους εικονοστοιχείων. Ένας αλγόριθμος ταξινόμησης εικόνων τυπικά μειώνει αυτή τη μήτρα σε έναν απλούστερο φορέα των χαρακτηριστικών της εικόνας όπως οι άκρες, οι καμπύλες, οι κηλίδες και οι υφές. Αυτά τα χαρακτηριστικά με τη σειρά τους μπορούν να συνδυαστούν για να κωδικοποιήσουν ευρύτερα χαρακτηριστικά όπως την ταυτότητα, το σχήμα, το προσανατολισμός και το περιβάλλον των αντικειμένων [76].

Μέχρι πρόσφατα, οι βελτιώσεις στην αυτοματοποιημένη ταξινόμηση εικόνων βασίζονταν σε μεγάλο βαθμό στη σχεδίαση χειροποίητων χαρακτηριστικών εικόνας για τη διάκριση των κατηγοριών ενδιαφέροντος που αφορούν τις εικόνες. Ο κλάδος λοιπόν της μηχανικής μάθησης που ονομάζεται "βαθιά εκμάθηση" έχει προσφέρει πρόσφατα σημαντικές βελτιώσεις στην απόδοση σε διάφορα καθήκοντα, όπως και την ταξινόμηση των εικόνων [77].

Ένα σύστημα βαθιάς μάθησης συνήθως χρησιμοποιεί ένα τεχνητό νευρωνικό δίκτυο πολλαπλών στρώσεων, μια διάταξη μαθηματικά διασυνδεδεμένων κόμβων εμπνευσμένων από βιολογικά νευρωνικά δίκτυα. Τα "βαθιά" νευρωνικά δίκτυα διαθέτουν ιεραρχικές αρχιτεκτονικές πολλαπλών στρώσεων που τους επιτρέπουν να μαθαίνουν όχι μόνο τις αντιστοιχίσεις χαρακτηριστικών δεδομένων σε κατηγορίες αλλά και τις ίδιες τις λειτουργίες [76].

Επομένως το βασικό καθήκον της διαγνωστικής ακτινολογίας είναι η ταυτοποίηση των παθολογικών ανωμαλιών σε εικόνες υψηλής ανάλυσης που ποικίλουν στην εμφάνιση λόγω της παθολογίας, του προσανατολισμού του ασθενούς, της φυσιολογικής ανατομικής διακύμανσης, των προηγούμενων ιατρικών επεμβάσεων (π.χ. τοποθέτηση του βηματοδότη) ή από τις διαφορές στις τεχνικές λήψης εικόνων ή στις μηχανές [25].

Ένα στοιχείο της αρχικής ερμηνείας εικόνας εκτελεί ένα σύνολο εργασιών ταξινόμησης για τον εντοπισμό αντικειμένων σε μια εικόνα, για την οποία η μηχανική μάθηση είναι κατάλληλη. Αυτό υποδηλώνει ότι η μηχανική μάθηση μπορεί να είναι ιδιαίτερα χρήσιμη για να βοηθήσει στην αναγνώριση των χαρακτηριστικών στις ακτινογραφίες.

Ωστόσο, παρά τις δεκαετίες προσπαθειών που βοηθούν στην αυτοματοποίηση της ερμηνείας των εικόνων σχετικά λίγοι αλγόριθμοι [78, 79] για τη διάγνωση με υπολογιστή έχουν ενσωματωθεί σε ευρεία κλινική χρήση.

Η επανάσταση στον τομέα της μηχανικής όρασης μπορεί να ειπωθεί ότι επήλθε το 2012, όταν αξιοποιήθηκαν τα συνθετικά νευρωνικά δίκτυα και παρουσίασαν τα μισά ποσοστά σφάλματος σε διαδικασία ταξινόμησης εικόνας [80]. Η νέα προσέγγιση επέτρεψε στα συστήματα υπολογιστικής εκμάθησης να λειτουργούν σε ακατέργαστα δεδομένα παρά σε χειροκίνητα χαρακτηριστικά ή τμήματα εικόνας, μαθαίνοντας πολλαπλά επίπεδα αφηρημένης αναπαράστασης μιας εικόνας [27].

Μόλις το σύστημα εκπαιδευτεί, πραγματοποιείται βελτιστοποιημένη για γρήγορη και αποτελεσματική ταξινόμηση των εικόνων. Ωστόσο, οι αλγόριθμοι υψηλής απόδοσης έχουν εκατοντάδες εκατομμύρια παραμέτρους που πρέπει να αντληθούν από τα δεδομένα, απαιτώντας μεγάλο αριθμό εικόνων για αποτελεσματική εκπαίδευση.

Πρόσφατα οι Baltruschat et al. [81] με στόχο την καλύτερη κατανόηση διαφορετικών προσεγγίσεων στην ταξινόμηση των ακτίνων X στο στήθος, διερεύνησαν λεπτομερώς μια ισχυρή αρχιτεκτονική δικτύου: το ResNet-50. Βασιζόμενη σε προηγούμενη εργασία σε αυτόν τον τομέα, αυτοί θεώρησαν την εκμάθηση μεταφοράς με και χωρίς τελειοποίηση, καθώς και την εκπαίδευση ενός αποκλειστικού δικτύου ακτίνων X από την αρχή.

Επίσης για να μεγιστοποιήσουν τη χωρική ανάλυση των δεδομένων ακτίνων X, συμπεριέλαβαν επίσης μια εκτεταμένη αρχιτεκτονική ResNet-50 και ένα δίκτυο που ενσωματώνει στη διαδικασία ταξινόμησης δεδομένα μη εικόνας (ηλικία ασθενούς, φύλο και τύπο απόκτησης). Σε ένα τελικό πείραμα, διερεύνησαν επίσης πολλαπλά βάθη ResNet (δηλαδή ResNet-38 και ResNet-101).

Στη συνέχεια σε μια συστηματική αξιολόγηση, με τη χρήση της 5πλής δειγματοληψίας και της απώλειας πολλαπλών ετικετών, συνέκριναν την απόδοση των διαφορετικών προσεγγίσεων για την ταξινόμηση της παθολογίας με στατιστικές ROC και ανέλυσαν τις διαφορές μεταξύ των ταξινομητών χρησιμοποιώντας συσχέτιση βαθμών.

Συνολικά αυτοί παρατήρησαν σημαντική εξάπλωση στην επιτευχθείσα απόδοση και κατέληξαν στο συμπέρασμα ότι τα ειδικά δεδομένα ακτίνων X ResNet-38, που ενσωματώνουν δεδομένα μη εικόνας, αποφέρουν τα καλύτερα συνολικά αποτελέσματα. Στις επόμενες υποενότητες αναφέρονται δυο σημαντικές διαδικασίες στην ταξινόμηση των εικόνων ιατρικών απεικονίσεων.

3.3.1 Η ανάγκη μετατροπής από rgb σε grayscale των ιατρικών απεικονίσεων στην ταξινόμηση τους

Η επιλογή εκκίνησης από ένα προ-εκπαιδευμένο μοντέλο έχει συνέπειες για τον τελικό σχεδιασμό ενός συνελκτικού νευρωνικού δικτύου. Προκειμένου να ληφθεί ένα ConvNet που έχει προ-εκπαιδευτεί σε φυσικές εικόνες και να τελειοποιηθεί σε ιατρικές εικόνες, οι ιατρικές εικόνες πρέπει να προεπεξεργαστούν για να είναι σύμφωνες με το σχήμα και τη δομή των αρχικών έγχρωμων εικόνων που χρησιμοποιούνται για την εκπαίδευση του δικτύου [82].

Επειδή οι ιατρικές εικόνες περιέχουν συχνά μόνο ένα κανάλι, αυτό συνήθως περιλαμβάνει τη στοίβαξη κάθε εικόνας σε κλίμακα του γκρι σε μια ψευδο-έγχρωμη εικόνα 3 καναλιών για να μιμείται τη δομή RGB φυσικών εικόνων. Ωστόσο, η εικόνα σε μορφή τρίτου καναλιού σε μορφή στοιβάδας δεν περιέχει πληροφορίες χρώματος. Ως εκ τούτου, δεν είναι σαφές εάν τα φίλτρα που μαθαίνονται από έγχρωμες εικόνες χρησιμοποιούνται πλήρως στη μεταφορά της εκμάθησης, ειδικά για φίλτρα στο πρώτο και στο δεύτερο στρώμα του ConvNet, τα οποία αντιπροσωπεύουν χαρακτηριστικά χαμηλότερου επιπέδου όπως χρώματα και άκρα [82].

Τα νευρωνικά δίκτυα αναπτύσσονται συχνά σε «φυσικές» εικόνες καθώς οι περισσότερες εικόνες λαμβάνονται με κανονικές κάμερες RGB. Πρόσφατα, τα πιο εξελιγμένα νευρωνικά δίκτυα έχουν επίσης προσαρμοστεί και αναπτυχθεί με μεγάλη επιτυχία για την ανίχνευση ανωμαλιών στις ιατρικές εικόνες, όπως στις ακτίνες X, CT και MRI. Ωστόσο, υπάρχουν κάποιες βασικές διαφορές μεταξύ των δύο τομέων: οι ιατρικές εικόνες περιέχουν συχνά ποσοτικές πληροφορίες και τα αντικείμενα δεν έχουν κανονικό προσανατολισμό. Επομένως η συνεκτίμηση αυτών των διαφορών και η προσαρμογή των αλγορίθμων μπορεί να ενισχύσει σημαντικά την απόδοση.

Σήμερα η έρευνα σχετικά με τα νευρωνικά δίκτυα και την ανάλυση εικόνων για φυσικές και ιατρικές εικόνες έχει παραδοσιακά χωριστεί. Η φυσική ανάλυση εικόνας συχνά αναφέρεται σε προβλήματα όπως ανίχνευση αντικειμένων, αναγνώριση προσώπου και 3D ανακατασκευή, χρησιμοποιώντας εικόνες από κανονικές κάμερες RGB. Ωστόσο η ανάλυση ιατρικών εικόνων

συνεπάγεται την ανίχνευση ασθενειών σε εικόνες ακτίνων X, ποσοτικοποίηση ανωμαλιών στη μαγνητική τομογραφία, τοποθέτηση οργάνων σε CT σαρώσεις κλπ [39].

Μεταξύ φυσικών και ιατρικών δεδομένων υπάρχουν πολλές βασικές διαφορές. Το πιο προφανές είναι η μορφή των δεδομένων: οι φυσικές εικόνες είναι συνήθως 2 (χωρικές) διαστάσεις RGB, ενώ τα ιατρικά δεδομένα μπορούν να έχουν οποιαδήποτε μορφή, όπως 2D σε γκρι, 2D με 4 κανάλια, 3D όγκους. Λαμβάνοντας υπόψη αυτά τα στοιχεία η απόδοση σε κλινικά χρησιμοποιήσιμα προϊόντα μπορεί να βελτιωθεί. Ωστόσο, υπάρχουν και πιο λεπτές διαφορές. Είναι γνωστό ότι το χρώμα παρέχει ισχυρές και σημαντικές διακριτικές πληροφορίες για πολλές εφαρμογές της μηχανικής όρασης σε υπολογιστές, όπως η ταξινόμηση, η κατάτμηση, η παρακολούθηση, η ευρετηρίαση και η ανάκτηση [83]. Το χρώμα ενός εικονοστοιχείου (pixel) μπορεί να περιγραφεί σε δύο συνιστώσες: τη χρωμικότητα και τη φωτεινότητα.

Η χρωμικότητα [84] ορίζει τα ‘συστατικά’ του χρώματος της εικόνας. Η φωτεινότητα καθορίζει την ένταση που λαμβάνεται από τα σταθμισμένα μέσα των καναλιών εικόνας κόκκινου, πράσινου και μπλε (RGB). Ο κύριος ‘χώρος’ χρώματος RGB είναι ο πιο δημοφιλής και οι άλλοι χώροι χρώματος όπως HSV και YUV, λαμβάνονται με γραμμικούς ή μη γραμμικούς μετασχηματισμούς του χώρου RGB [85].

Σε πολλές εφαρμογές ηλεκτρονικών οραμάτων, διαφορετικοί χρωματικοί χώροι μπορεί να είναι πιο κατάλληλοι για να αποκτήσουν λογικά αποτελέσματα. Για παράδειγμα, ο χώρος CIE $L^* u^* v^*$ θεωρείται πιο κατάλληλος από RGB για πολλές εφαρμογές κατάτμησης εικόνας [86, 87]. Παρόλο που οι έγχρωμες αναλύσεις εικόνων έχουν χρησιμοποιηθεί ευρέως σε εφαρμογές μηχανικής όρασης, εξακολουθούν να υπάρχουν πολλοί αλγόριθμοι μηχανικής όρασης που απαιτούν μετασχηματισμό χρώματος-γκρι [88, 89].

Οι μέθοδοι μετατροπής χρώματος σε γκρι χρησιμοποιούν τη λειτουργία μείωσης διαστάσεων τυπικά μετασχηματίζοντας τις τιμές RGB tristimulus $(R, G, B) \in \mathbb{R}^3$ σε τιμές έντασης $(I) \in \mathbb{R}^1$. Τα τελευταία χρόνια έχουν προταθεί πολλές προσεγγίσεις μετατροπής χρώματος προς γκρι. Μια βασική προσέγγιση είναι να χρησιμοποιηθούν τα μεμονωμένα κανάλια έγχρωμων εικόνων, όπως η συνιστώσα αξίας (V) του χώρου χρώματος HSV ή το κόκκινο, πράσινο ή μπλε συστατικό χρώματος του χώρου χρώματος RGB. Ωστόσο, αυτές οι προσεγγίσεις μπορεί να οδηγήσουν σε απώλεια διακριτικής πληροφορίας λόγω της εκμετάλλευσης μόνο ενός καναλιού εικόνας.

Μια άλλη προσέγγιση είναι η δημιουργία μιας εικόνας σε κλίμακα του γκρι με την ίδια συμβολή των καναλιών χρώματος R, G και B:

$$I = 0,33 \times R + 0.33 \times G + 0.33 \times B \quad (3.1)$$

, η οποία δεν έχει χρησιμοποιηθεί συνήθως στις εφαρμογές μηχανικής όρασης επειδή κάθε κανάλι διαθέτει διαφορετική φασματική ισχύ [90].

Οι συντελεστές της επιτροπής εθνικών τηλεοπτικών προτύπων (National Television Standards Committee, NTSC) [90] θεωρούν την άνιση συμβολή των διαύλων χρώματος για τη μετατροπή RGB σε κλίμακα του γκρι:

$$I = 0,299 \times R + 0.587 \times G + 0.114 \times B \quad (3.2)$$

Παρόλο που η μέθοδος NTSC έχει γίνει αποδεκτή ως κοινή προσέγγιση για τη μετατροπή από RGB σε κλίμακα του γκρι, πρόσφατα έχουν προταθεί πολλές προηγμένες μέθοδοι μετατροπής χρώματος-γκρι.

Η μετατροπή χρώματος σε γκρι έγινε με τη διατήρηση των άκρων χρώματος [91] ή των χαρακτηριστικών εικόνας [92]. Οι Smith et al. [93] πρότειναν μια προσέγγιση μετατροπής η οποία συνδυάζει την παγκόσμια χαρτογράφηση και την τοπική βελτίωση για να αποκτήσει τις εικόνες σε κλίμακα του γκρι, σε δύο βήματα. Οι Grundl και Dodgson [94] πρότειναν τον αλγόριθμο Decoralize ο οποίος συνδυάζει γραμμικά τα κανάλια χρωματισμού και φωτεινότητας για να παράγει την καλύτερη ενίσχυση αντίθεσης. Οι Goach et al. [95] περιέγραψαν έναν αλγόριθμο μετατροπής της κλίμακας του γκρι που στη βιβλιογραφία είναι γνωστός ως Color2Gray. Οι Kim et al. [96] για τη διατήρηση των πληροφοριών χρώματος μειώνοντας παράλληλα τη διάσταση ανέπτυξαν έναν γρήγορο μη γραμμικό αλγόριθμο παγκόσμιας χαρτογράφησης για μετατροπή χρώματος σε κλίμακα του γκρι.

Όλες αυτές οι συμβατικές μέθοδοι μπορεί να παράγουν καλά αποτελέσματα για οπτική εμφάνιση. Ωστόσο, μια εικόνα με ικανοποιητική οπτική εμφάνιση μπορεί να μην είναι αρκετά καλή, για παράδειγμα, για προβλήματα ταξινόμησης. Επομένως, για παράδειγμα στην ταξινόμηση, η μετατροπή χρώματος σε γκρι πρέπει να γίνει για τη μεγιστοποίηση των αποστάσεων των εικόνων μεταξύ διαφορετικών κλάσεων.

Επειδή τα κανάλια χρώματος μπορεί να έχουν διαφορετικό επίπεδο συνεισφοράς από το επίπεδο διανομής που ορίζεται στην προσέγγιση πρότυπου και ισότιμου συντελεστή NTSC. Για παράδειγμα, εξετάζοντας τους NTSC συντελεστές, το στοιχείο G συμβάλλει στο 0,587

της ενέργειας των εικόνων σε κλίμακα του γκρι (Εξίσωση 3.2) και αυτή η συνεισφορά μπορεί να μην είναι κατάλληλη για την απόκτηση διακριτικών εικόνων για ταξινόμηση.

Από αυτήν την άποψη, ο Lu και ο Plataniotis [97] πρότειναν μια μέθοδο μετατροπής χρώματος σε γκρι για αναγνώριση προσώπου και δήλωσαν ότι οι τυπικοί συντελεστές NTSC δεν είναι βέλτιστοι για τη διαδικασία αναγνώρισης προσώπου. Χρησιμοποιώντας τα δύο συστήματα ανίχνευσης προσώπου βασισμένα στο Adaboost, αυτοί βελτίωσαν την ακρίβεια ταξινόμησης κατά 2,85 και 3,58% για κάθε σύστημα ενισχύοντας τη συμβολή του κόκκινου καναλιού.

Οι Jones και Abbott [98] επίσης δήλωσαν ότι η κοινή προσέγγιση μετατροπής ίσων βαρών που λαμβάνεται λαμβάνοντας τον μέσο όρο των διαύλων χρώματος R, G, B μπορεί να μην είναι αποτελεσματική για προσεγγίσεις προσανατολισμένες σε εργασίες όπως η αναγνώριση προσώπου. Προκειμένου να βελτιωθεί η απόδοση αναγνώρισης προσώπου, αυτοί πρότειναν το πλαίσιο μετατροπής χρώματος-γκρι-βασισμένο στο Karhunen-Loeve (KL).

Με τη χρήση της ανάλυσης KL, ο βελτιστοποιημένος συντελεστής χρωμάτων RGB w^T που λαμβάνεται από τον ιδιοδιανύτη $U = [u_1, u_2, u_3]$ έχει τη μεγαλύτερη ιδιοτιμή της μήτρας συνδιακύμανσης των R, G, B διανομή χρώματος. Οι Jones και Abbott χρησιμοποίησαν τη γνωστή μέθοδο eigenvalue για πειράματα αναγνώρισης προσώπου και βελτίωσαν την απόδοση αναγνώρισης αυξάνοντας την απόδοση από το 4 έως 14% ανάλογα με τον αριθμό των χρησιμοποιούμενων ιδιοσκευών [98].

3.3.2 Επεξεργασία των ιατρικών απεικονίσεων με αλγόριθμο ιστογραμμάτων κατευθυνόμενων κλίσεων (Histogram of oriented gradients , HOG)

Στον τομέα 2D εικόνας, ο περιγραφέας HOG οργανώνει τις διαβαθμίσεις της εικόνας προς επεξεργασία σε ιστογράμματα. Ως το πρώτο βήμα, με την κατάλληλη μάσκα φίλτρου η εικόνα διαβάθμισης υπολογίζεται με τη συνέλιξη της εικόνας εισόδου. Στη συνέχεια κατασκευάζεται ένα πλέγμα ιστογραμμάτων, όπου κάθε ιστόγραμμα οργανώνει τις αντίστοιχες βαθμίδες σε bins ανάλογα με τον προσανατολισμό τους.

Για να διατηρηθεί η τοπική ανάλυση, για κάθε κελί υπολογίζεται ένα ιστόγραμμα σε ένα ομοιόμορφα διαχωρισμένο πλέγμα. Κατά συνέπεια, κάθε κελί περιέχει τον ίδιο αριθμό βαθμίδων (ανάλογα με το μέγεθος του κελιού) και παίρνει ακριβώς ένα ιστόγραμμα. Τα ίδια

τα κελιά στη συνέχεια οργανώνονται σε ορθογώνια μπλοκ, τα οποία μπορεί να αλληλεπικαλύπτονται. Οι τιμές ιστογράμματος όλων των κελιών μέσα σε ένα μπλοκ συνδυάζονται για να σχηματίσουν ένα φορέα. Ο φορέας κάθε μπλοκ στη συνέχεια ομαλοποιείται και στη συνέχεια, η συνένωση όλων αυτών των διανυσμάτων αποδίδει τον τελικό διάνυσμα των χαρακτηριστικών της εικόνας [99].

Συνοπτικά, ένας περιγραφέας εικόνων είναι ένας αλγόριθμος που εξάγει ένα διάνυσμα του οποίου τα συστατικά περιγράφουν το περιεχόμενο ενός συνεχόμενου μέρους μιας εικόνας. Η απλούστερη μορφή ενός περιγραφικού είναι εκείνη που εξάγει τις εντάσεις των εικονοστοιχείων μιας περιοχής εικόνας [100]. Ωστόσο, αυτή η απλή περιγραφή έχει δύο μείζονα μειονεκτήματα: (1) δεν είναι υπολογιστικά αποτελεσματική λόγω της σχετικά μεγάλης διάστασης του φορέα, (2) ο περιγραφέας δεν είναι ανθεκτικός στις στρεβλώσεις εικόνας.

Αξίζει να αναφερθεί, ενώ ένας τέτοιος περιγραφικός κατάλογος έχει νόημα σε μερικά σενάρια, οι υπερσύγχρονοι περιγραφείς αντιπροσωπεύουν πιο εξελιγμένες λύσεις στις οποίες δεν παραβλέπονται ιδιότητες υπολογιστικής απόδοσης, ευρωστίας και ανακατανομής.

Για το σχεδιασμό αποδοτικών και αποτελεσματικών περιγραφικών δεικτών έχουν χρησιμοποιηθεί διαφορετικές τεχνικές. Μια από τις πιο κοινές προσεγγίσεις είναι η χρήση ιστογραμμάτων για την παρουσίαση ορισμένων χαρακτηριστικών που σχετίζονται με την εμφάνιση ή το σχήμα [100].

Οι πιο δημοφιλείς περιγραφείς όπως ο μετασχηματισμός μεταβλητών ιδιοτήτων κλίμακας (Scale invariant Feature Transform, SIFT) [101], τα τοπικά δυαδικά πρότυπα (Local binary patterns, LBP) [102] ή το ιστόγραμμα προσανατολισμένων διαβαθμίσεων (HOG) [103] είναι στρατηγικές που βασίζονται στη διανομή. Σε ορισμένες περιπτώσεις, όπως στο SIFT, το ιστόγραμμα απεικονίζει διαφορικές πληροφορίες, όπου είναι 'χτισμένο' ένα τρισεπίδη ιστόγραμμα σημείων κλίσης και προσανατολισμών για να περιγράψει την κατανομή άκρων.

Μια άλλη κοινή στρατηγική βασίζεται στην ανάλυση χωρικών συχνοτήτων των περιοχών εικόνας. Παραδείγματα από αυτή την κατηγορία είναι οι περιγραφικοί δείκτες που κατασκευάζονται από τις τράπεζες φίλτρων Gabor ή το μετασχηματισμό wavelet. Οι τεχνικές που βασίζονται σε στιγμές είναι επίσης συχνές στο σχεδιασμό περιγραφικών δεικτών.

Σε αυτή την περίπτωση, οι περιγραφείς έχουν σχεδιαστεί για να επωφελούνται από τις ιδιότητες invariance των στιγμών σε σχέση με τους γεωμετρικούς μετασχηματισμούς. Για

παράδειγμα, υπάρχουν ορισμένοι περιγραφείς σχήματος που βασίζονται σε στιγμές Zernike, οι οποίες είναι αμετάβλητες στην περιστροφή και στιβαρή στο θόρυβο [104].

Γενικά ο σχεδιασμός αλγορίθμων με σκοπό την παροχή συνοπτικής περιγραφής των περιοχών εικόνας (ή ολόκληρης της εικόνας) είναι ένα παραγωγικό και κεντρικό ερευνητικό θέμα στους τομείς της εικόνας των υπολογιστών και της ανάλυσης εικόνας [51]. Οι αλγόριθμοι αυτοί, γνωστοί ως περιγραφικοί, έχουν χρησιμοποιηθεί με επιτυχία σε ένα ευρύ φάσμα εφαρμογών όπως η αναγνώριση αντικειμένων (class) [52] και η ανάκτηση εικόνων [53].

Λόγω της επιτυχημένης εφαρμογής του σε περίπλοκα περιβάλλοντα, ο HOG έχει προσελκύσει μεγάλο ενδιαφέρον ως μέθοδος εξαγωγής χαρακτηριστικών και χρησιμοποιείται συχνά στη βιβλιογραφία. Το Oriented Gradient (HOG), το οποίο χρησιμοποιείται συνήθως για την αναγνώριση αντικειμένων και προτύπων, παρέχει υψηλή απόδοση κάτω από διαφορετικές συνθήκες [105].

Το HOG εξάγει πληροφορίες σχετικά με την κατανομή των άκρων στόχων (target edges) στις τοπικές περιοχές ως μέσο αντιπροσώπευσης του σχήματος στόχου. Η χρήση της μεθόδου HOG, η οποία μπορεί επίσης να χαρακτηριστεί ως ο χαρακτηρισμός των τιμών προσανατολισμού (θ) και του μεγέθους (m) των εικονοστοιχείων στην εικόνα, έχει μεγάλο ενδιαφέρον για πολλά πεδία. Ο κύριος σκοπός αυτής της μεθόδου είναι να ορίσει μια εικόνα, ως μια ομάδα τοπικών ιστογραμμάτων. Αυτά τα ιστογράμματα περιλαμβάνουν τον αριθμό των προσανατολισμών των κλίσεων σε μια τοπική περιοχή της εικόνας [105, 106].

Τα κύρια βήματα της διαδικασίας για την εφαρμογή των περιγραφέων HOG έχουν ως εξής [107]:

1. **Υπολογισμός της κλίσης της εικόνας:** Κατ' αρχάς, οι εικόνες υφής επιπέδου χρώματος μειώνονται σε γκριζο επίπεδο. Η οριζόντια κλίση f_a και η κατακόρυφη κλίση f_b υπολογίζονται χρησιμοποιώντας παράγωγο μέσο (derivative mask) της εικόνας γκριζου επιπέδου. Η χρήση ενός πιο σύνθετου παράγωγου μέσου μειώνει την απόδοση του συστήματος, έτσι χρησιμοποιείται ένα απλό συνήθως. Οι οριζόντιες διαβαθμίσεις εικόνας f_a και οι κάθετες f_b are υπολογίζονται ως εξής:

$$f_a(\alpha, b) = I(\alpha + 1, b) - I(\alpha - 1, b) \quad \forall \alpha \quad (3.1)$$

$$f_b(\alpha, b) = I(\alpha, b+1) - I(\alpha, b-1) \quad \forall \alpha, b \quad (3.2)$$

Σε αυτές τις εξισώσεις το $I(\alpha, b)$ αντιπροσωπεύει την πυκνότητα των pixel στο σημείο (α, b) .

2. **Υπολογισμός προσανατολισμού των ιστογραμμάτων για συγκεκριμένες θέσεις:** Χρησιμοποιώντας τις οριζόντιες κλίσεις εικόνας f_a και τις κατακόρυφες f_b , το μέγεθος κλίσης (m) και προσανατολισμού (θ) υπολογίζονται ως εξής:

$$m(\alpha, b) = \sqrt{f_a(\alpha, b)^2 + f_b(\alpha, b)^2} \quad (3.3)$$

$$\theta(\alpha, b) = \tan^{-1} \left(\frac{f_b(\alpha, b)}{f_a(\alpha, b)} \right) \quad (3.4)$$

3. **Καμπύλη ιστογράμματος κατεύθυνσης:** στην πράξη, οι συντελεστές κλίσης που είναι μικρότεροι από 0 αθροίζονται με 180, δεδομένου ότι δεν πρέπει να σημαίνονται οι επιθυμητές τιμές κατεύθυνσης. Συνεπώς, οι μη σημειωμένοι προσανατολισμοί κλίσης της αποκτηθείσας νέας εικόνας υπολογίζονται ως εξής:

$$\check{\theta}(a, b) = \begin{cases} \theta(\alpha, b) + \pi & \text{if } \theta(\alpha, b) < 0 \\ \theta(\alpha, b) & \text{else if} \end{cases} \quad (3.5)$$

Αφού υπολογιστεί ο προσανατολισμός κλίσης, η εικόνα κλίσης διαιρείται από τα κελιά στο μέγεθος εικονοστοιχείου $[u \ a \ v]$. Η γωνία προσανατολισμού $\check{\theta}(a, b)$ διαιρείται από τη γωνία κλίσης (m) σε διαφορετικό διάστημα γωνίας για να παραχθεί το ιστόγραμμα του προσανατολισμού κάθε κελιού. Το ιστόγραμμα προσανατολισμού δημιουργείται διαιρώντας εξίσου το εύρος $0^\circ - 180^\circ$.

4. **Ομαλοποίηση των μπλοκ:** Η ομαλοποίηση των μπλοκ δείχνει το ιστόγραμμα προσανατολισμού για κάθε κελί και αφαιρούνται τα μεγαλύτερα μπλοκ $[n \times m]$. Δεδομένου ότι κάθε κελί έχει προσανατολισμούς k , το μέγεθος ιδιότητας για κάθε μπλοκ είναι $[n \times m \times k]$. Το χαρακτηριστικό διάνυσμα v αντιπροσωπεύει το μη κανονικοποιημένο ιστόγραμμα στο σημείο (i, j) του μπλοκ $p(i, j)$. Το χαρακτηριστικό διάνυσμα ομαλοποιείται με τη μέθοδο του προτύπου L2 ως εξής:

$$p'(ij) = \frac{p(ij)}{\sqrt{|v|^2 + \epsilon}} \quad (\epsilon=1) \quad (3.6)$$

, όπου ϵ είναι μια πολύ μικρή σταθεροποίηση κανονικοποίησης που χρησιμοποιείται για να αποφευχθεί η διαίρεση με το μηδέν.

5. Στη συνέχεια της ανάλυσης των βημάτων πολλές εργασίες χρησιμοποιούν γραμμικό μετασχηματισμό (Linear Discriminant Analysis, LDA) για να επιτύχουν το μετασχηματισμό από υψηλή διάσταση σε χαμηλότερη διάσταση με τη χρήση της εφαρμογής. Το κριτήριο στον αλγόριθμο LDA είναι να αποκτηθούν οι διαστάσεις που έχουν τα πιο ιδιαίτερα χαρακτηριστικά από τις προεπιλεγμένες κατηγορίες. Με αυτή τη μέθοδο, επιδιώκεται να αποκτηθεί η υψηλότερη διαφορά, φέρνοντας το δείκτη κατανομής στην τάξη της διακλάδωσης διανομής στο υψηλότερο επίπεδο.
6. Σε αυτό το στάδιο της ανάλυσης το τεχνητό νευρωνικό δίκτυο είναι απαραίτητο για την ταξινόμηση αυτού του συνόλου δεδομένων αφού οι πληροφορίες εικόνας μεγάλων διαστάσεων έχουν μειωθεί σε δεδομένα με μικρότερες διαστάσεις.

Η μέθοδος HOG με βάση τη βιβλιογραφία έχει χρησιμοποιηθεί σε πολλές μελέτες για την ερμηνεία ακτινολογικών δεδομένων. Με βάση τους Raghavendra et al. [108] που μελέτησαν μαστογραφίες, η HOG είναι σε θέση να καταγράφει τις ελάχιστες εσωτερικές διακυμάνσεις στα θερμογραφήματα. Προκειμένου να μειωθούν οι διαστάσεις των εξαγόμενων περιγραφών της HOG, χρησιμοποιούνται οι περιγραφείς της διατήρησης της προβολής (kernel locality preserving projection, KLPP) του πυρήνα. Τα προκύπτοντα χαρακτηριστικά KLPP ταξινομούνται στη συνέχεια για να διαμορφώσουν ένα αποδοτικό μοντέλο ταξινόμησης. Για την επικύρωση της προτεινόμενης μεθόδου χρησιμοποιούνται διάφοροι αλγόριθμοι μηχανικής μάθησης.

Σύμφωνα με τα αποτελέσματά τους η μεθόδός τους παρουσιάζει πολλά υποσχόμενη απόδοση με μέση ακρίβεια, ευαισθησία, ειδικότητα και περιοχή κάτω από την καμπύλη 98%, 96,66%, 100% και 0,98 αντίστοιχα. Αυτοί ανέπτυξαν επίσης δείκτη κινδύνου καρκίνου του μαστού (BCRI) χρησιμοποιώντας σημαντικά χαρακτηριστικά KLPP που μπορούν να διακρίνουν τις δύο κατηγορίες χρησιμοποιώντας έναν ενιαίο ολοκληρωμένο δείκτη. Αυτό μπορεί να βοηθήσει τους ακτινολόγους να διακρίνουν τις κανονικές και κακοήθεις μορφές κατά τη διάρκεια της εξέτασης για να επικυρώσουν τα ευρήματά τους.

Επίσης οι Farshbaf et al. [109] παρουσίασαν μια νέα μέθοδο εντοπισμού κεφαλαιομετρικών ορόσημων με βάση το συνδυασμό δύο αποτελεσμάτων ταξινομητή (classifier). Αρχικά, ένας ταξινομητής βασισμένος σε ιστογράμματα προσανατολισμένων κλίσεων κάνει μια πρώτη εκτίμηση των δυνητικών παραθύρων και στη συνέχεια ένας δεύτερο ταξινομητής, με βάση τα

ιστογράμματα του γκρι προφίλ, ταξινομεί τα εντοπισμένα παράθυρα. Συνδυάζοντας τα αποτελέσματα αυτών των δύο ταξινομητών, λαμβάνεται η τελική απόφαση σχετικά με τη θέση παραθύρου ορόσημου.

Τα χαρακτηριστικά HOG συγκεντρώνουν τα προφίλ των άκρων στην εικόνα. Ωστόσο η λήψη αποφάσεων για το πιο κατάλληλο παράθυρο σε ορισμένα παράθυρα ανίχνευσης απαιτεί περισσότερες πληροφορίες από απλά προφίλ ακμών. Με την προσθήκη των χαρακτηριστικών γκρίζου προφίλ στο σύστημα και τη σωστή συσχέτιση των αποτελεσμάτων, η απόδοση ανίχνευσης αυξάνεται σημαντικά για μια σειρά από δύσκολα και εύκολα σημεία αναφοράς.

Οι Yao et al. [110] χρησιμοποιούν ένα μοντέλο παραμορφώσιμου τμήματος βασισμένο σε ένα 2D ιστόγραμμα προσανατολισμένων διαβαθμίσεων (HOG) για τη μελέτη των σπονδύλων. Αποδίδουν το κέντρο των σπονδύλων τμήμα-τμήμα και εκτελούν nonmaxima καταστολή για την εξαγωγή ανιχνεύσεων σπονδύλων σε ένα κεντρικό τμήμα. Επίσης δε πραγματοποιείται 3D εκτίμηση του πλαισίου οριοθέτησης.

Οι Daenzer et al. [111] αναφέρουν τη μέθοδο VolHOG (ιστογράμματα προσανατολισμένων διαβαθμίσεων για ογκομετρικά δεδομένα) για 3D ανίχνευση τραχηλικού σπονδύλου σε εικόνες MR. Αυτοί πρότειναν επίσης μια μέθοδο μηχανικής μάθησης με στόχο την ανίχνευση τραχηλικών σπονδύλων βασισμένη σε νέα χαρακτηριστικά, σε συνδυασμό με μια γραμμική μηχανή φορέα υποστήριξης με στόχο την ταξινόμηση. Αυτοί παρέχουν έναν αλγόριθμο για την παραγωγή διγραμμικού προσανατολισμού ιστογράμματος κλίσης από τρισδιάστατα δεδομένα ράστερ εικόνας που τους επιτρέπει να περιγράψουν τρισδιάστατα αντικείμενα χρησιμοποιώντας τα διμεταβλητά ιστογράμματα από άλλους ερευνητές. Το VolHOG έχει τεράστιο δυναμικό για πολλούς τύπους ακτινολογικών εικόνων λόγω της τρισδιάστατης φύσης του.

Κατά την απεικόνιση της καρδιάς, χρησιμοποιώντας έναν 2D ανιχνευτή υπερήχων, μπορούν να εκδηλωθούν διαφορετικές απόψεις ανάλογα με τη θέση και τις γωνίες του 'καθετήρα' του ανιχνευτή. Ορισμένες από αυτές τις απόψεις έχουν χαρακτηριστεί ως τυπικές απόψεις, λόγω της παρουσίας και της ευκολίας αξιολόγησης των βασικών καρδιακών δομών σε αυτές.

Ωστόσο οι Agarwal et al. [112] χρησιμοποιώντας το ιστόγραμμα των προσανατολισμένων διαβαθμίσεων (HOG) ως χαρακτηριστικό διάκρισης που κωδικοποιεί την χωρική διάταξη ακμών/κλίσεων στις εικόνες ερμήνευσαν αρκετές εικόνες καρδιάς. Η λειτουργία HOG υπολογίζεται στα δεδομένα εικόνας που έχουν μετατραπεί πριν από τη σάρωση στο χώρο

δέσμης υπερήχων. Σε μια αρκετά μεγάλη βάση δεδομένων με 703 εικόνες, με έναν ταξινομητή φορέα υποστήριξης, λαμβάνεται η κάθε εικόνα με ακρίβεια περίπου 98%.

Αξίζει να σημειωθεί ότι οι Dalal et al. [113] ήταν αυτοί που εισήγαγαν τα χαρακτηριστικά HOG για να ανιχνεύουν την εμφάνιση και το σχήμα του τοπικού αντικειμένου χρησιμοποιώντας την ένταση της τοπικής κλίσης. Τα χαρακτηριστικά HOG υπολογίζονται από το μεγαλύτερο αντιπροσωπευτικό 'κομμάτι' του ζωτικού οργάνου προς εξέταση. Ο μέσος όρος, η διακύμανση και η τυπική απόκλιση των χαρακτηριστικών HOG χρησιμοποιούνται στη συνέχεια σε προτεινόμενα συστήματα ανάκτησης και μεγαλύτερη ακρίβεια στις εικόνες.

4 Ανάπτυξη αλγορίθμου για ταξινόμηση ακτινολογικών ιατρικών απεικονίσεων

4.1 Μετατροπή ακτινολογικών δεδομένων

Όπως προαναφέρθηκε, ένας περιγραφέας χαρακτηριστικών είναι μια αναπαράσταση μιας εικόνας ή μιας ετικέτας εικόνας που απλοποιεί την εικόνα με την εξαγωγή χρήσιμων πληροφοριών και την απόσπαση εξωγενών πληροφοριών. Συνήθως, ένας περιγραφέας χαρακτηριστικών μετατρέπει μια εικόνα μεγέθους πλάτους x ύψους x 3 (κανάλια) σε ένα διάνυσμα χαρακτηριστικών/συστοιχία μήκους n .

Στην περίπτωση του περιγραφέα χαρακτηριστικών HOG, η εικόνα εισόδου έχει μέγεθος $64 \times 128 \times 3$ και το διάνυσμα χαρακτηριστικών εξόδου είναι μήκους 3780. Πρέπει να ληφθεί υπόψη ότι ο περιγραφέας HOG μπορεί να υπολογιστεί για άλλα μεγέθη, αλλά για το συγκεκριμένο παράδειγμα θα δοθούν συγκεκριμένα νούμερα.

Σε μια τέτοια διαδικασία είναι πολύ χρήσιμο να είναι γνωστό, τι είναι χρήσιμο και για ποιο σκοπό είναι χρήσιμο. Σαφώς, το διάνυσμα χαρακτηριστικών δεν είναι χρήσιμο για την προβολή της εικόνας, αλλά, είναι πολύ χρήσιμο για εργασίες όπως η αναγνώριση εικόνων και η ανίχνευση αντικειμένων.

Το διάνυσμα χαρακτηριστικών που παράγεται από αυτούς τους αλγορίθμους όταν τροφοδοτούνται σε αλγόριθμους ταξινόμησης εικόνων όπως η μηχανή φορέα υποστήριξης (SVM) παράγει καλά αποτελέσματα. Αλλά, το ερώτημα εντοπίζεται στο τί είδους "χαρακτηριστικά" είναι χρήσιμα για τις εργασίες ταξινόμησης.

Αν υποθεθεί ότι ο στόχος είναι το χτίσιμο ενός ανιχνευτή αντικειμένων που να ανιχνεύει κουμπιά πουκάμισων και παλτών, τότε ένα κουμπί είναι κυκλικό (μπορεί να φαίνεται ελλειπτικό σε μια εικόνα) και συνήθως έχει μερικές τρύπες για το ράψιμο. Επομένως εύκολα μπορεί να τρέξει ένας ανιχνευτή άκρων στην εικόνα ενός κουμπιού και να γίνει γνωστό εύκολα αν πρόκειται για ένα κουμπί απλά κοιτάζοντας την εικόνα του άκρου μόνο. Στην περίπτωση αυτή, οι πληροφορίες ακμής είναι "χρήσιμες" και οι πληροφορίες χρώματος δεν είναι.

Επιπλέον, τα χαρακτηριστικά πρέπει να έχουν επίσης διακριτική ισχύ. Για παράδειγμα, τα καλά χαρακτηριστικά που εξάγονται από μια εικόνα θα πρέπει να είναι σε θέση να διακρίνουν τη διαφορά μεταξύ κουμπιών και άλλων κυκλικών αντικειμένων, όπως νομίσματα

και ελαστικά αυτοκινήτων. Στον περιγραφέα χαρακτηριστικών HOG, χρησιμοποιούνται ως χαρακτηριστικά γνωρίσματα διανομής (ιστογράμματα) διευθύνσεων κλίσεων (προσανατολισμένες κλίσεις).

Σημειώνεται ότι οι διαβαθμίσεις (παράγωγα x και y) μιας εικόνας είναι χρήσιμες επειδή το μέγεθος των διαβαθμίσεων είναι μεγάλο γύρω από τις άκρες και τις γωνίες (περιοχές με απότομες μεταβολές έντασης) και οι άκρες και οι γωνίες σε σύγκριση με τις επίπεδες περιοχές είναι χρήσιμες για την εξαγωγή πολύ περισσότερων πληροφοριών σχετικά με το σχήμα του αντικειμένου.

Ο υπολογισμός του HOG περιγραφέα περιγράφεται στην παρούσα υποενότητα και πραγματοποιείται σύμφωνα με τα παρακάτω βήματα [114]:

Βήμα 1^ο: Επιλογή Εικόνων και μετατροπή σε Dicom. Οι εικόνες είναι αποτελέσματα X-Ray και επιλέχτηκαν από το [115]. Αυτές οι εικόνες είναι συμπιεσμένες σε μορφή jpeg, και μετατράπηκαν σε εικόνες Dicom με βοηθητικό πρόγραμμα δεδομένου ότι δεν υπάρχουν τόσες πολλές εικόνες της μορφής αυτής. Ενδεικτικά παρουσιάζονται τρεις εικόνες Dicom ως δείγμα μαζί τις αντίστοιχες επικεφαλίδες (headers) που περιέχουν συνοδευτικές πληροφορίες (π.χ. στοιχεία ασθενούς, είδος εξέτασης κ.ά.).

Δείγμα πρώτο εικόνας Dicom (X-Ray κάτω άκρους)



Επικεφαλίδα – Header του πρώτου δείγματος Dicom

Filename: 'C:\Users\John\Desktop\Matlab_Dicom_Image_Classification\Training Data\XRay Dicom\ANKLE-FRACTURE-FIXED-2.dcm'

Rows: 682

Columns: 600

BitsStored: 8

Δείγμα δεύτερο εικόνας Dicom (Ultrasound κοιλιακής χώρας)



Επικεφαλίδα – Header του δεύτερου δείγματος Dicom

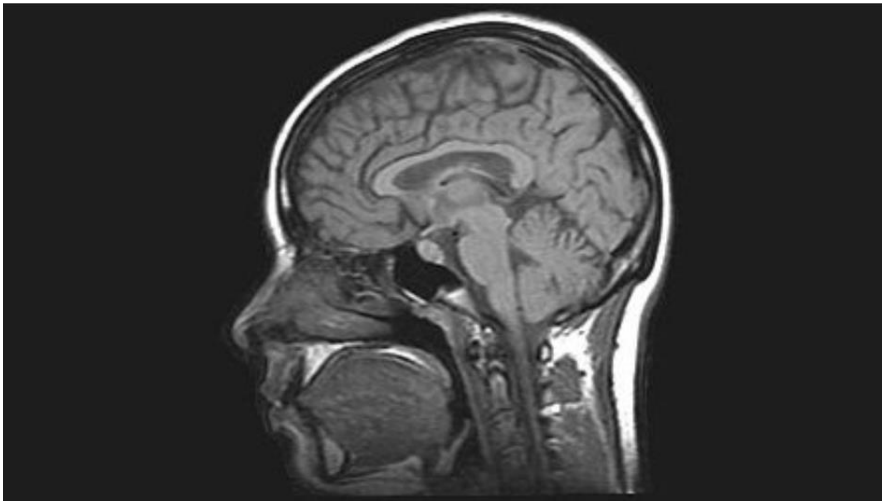
Filename: 'C:\Users\John\Desktop\Matlab_Dicom_Image_Classification\Training Data\Ultrasound Dicom\ultrasound-training-4..dcm'

Rows: 280

Columns: 522

BitsStored: 8

Δείγμα τρίτο εικόνας Dicom (MRI εγκεφάλου)



Επικεφαλίδα – Header του τρίτου δείγματος Dicom

Filename: 'C:\Users\John\Desktop\Matlab_Dicom_Image_Classification\Training Data\Others Dicom\MRI_brain_scan_node_full_image_2.dcm'

Rows: 251

Columns: 477

BitsStored: 8

Οι παραπάνω επικεφαλίδες – headers εμφανίζουν πολλές πληροφορίες αλλά οι πιο χρήσιμες πληροφορίες μίας τέτοιας εικόνας είναι οι παραπάνω τιμές. Οι πληροφορίες των επικεφαλίδων – headers μπορούν να εξαχθούν με την εντολή `dicominfo` του προγράμματος Matlab και συντάσσεται ως εξής:

```
I = dicomread ('path\name file');  
info = dicominfo ('path\name file').
```

Βήμα 2^ο: Προεπεξεργασία. Εκτέλεση του αρχείου HOG στο Matlab. Οι εικόνες φορτώνονται και επεξεργάζονται μία προς μία στο Matlab από φάκελο με την παρακάτω μορφή κώδικα:

FolderImages=dir(strcat(cd, '\Test Data', '*')); , όπου Test Data ο φάκελος από όπου φορτώνονται οι εικόνες Dicom.

Η επεξεργασία περιλαμβάνει την ανάγνωση μιας εικόνας και τον έλεγχο αν αποτελεί εικόνα RGB ή όχι . Αν η εικόνα είναι έγχρωμη εικόνα RGB τότε μετατρέπεται σε εικόνα σε κλίμακα του γκρι. Στη συνέχεια, αυτή η εικόνα που βρίσκεται σε κλίμακα του γκρι αναπροσαρμόζεται σε κανονικές διαστάσεις μεγέθους 512x512, όπως φαίνεται στην συνέχεια με την μορφή κώδικα:

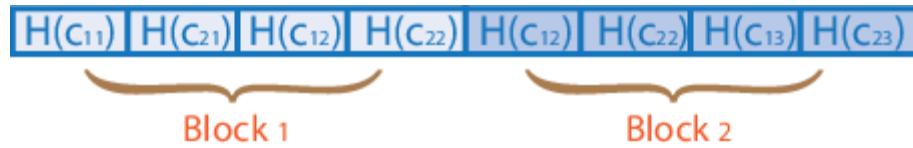
```
for j_fold=3:length(FolderImages);  
  
    info = dicominfo(strcat(cd, '\Test Data\', FolderImages(j_fold)).name  
    Y = dicomread(info);  
    if(length(size(Y))<=3)  
        Y1=Y;  
        if(size(Y1,3)==3);  
            Y1=rgb2gray(Y1);  
        end;  
    Y1=imresize(Y1, [512,512]);  
    Y1=uint8(Y1); .
```

Έπειτα ακολουθεί η εξαγωγή χαρακτηριστικών βασισμένη στο HOG η οποία εφαρμόζεται στην εικόνα χρησιμοποιώντας μια τιμή μεγέθους κελιού 128x128 η οποία εξάγει τιμές 324 χαρακτηριστικών από κάθε εικόνα, όπως δηλώνεται με το σημείο του κώδικα:

```
Feature_Val=compute_HOG(Y1,128,2,9);  
  
    Pred_Val=net(Feature_Val);  
    Pred_Val=vec2ind(Pred_Val);  
    figure  
    imshow(Y1, []);title(Classes_Names{Pred_Val});.
```

Αυτές με τη σειρά τους αποθηκεύονται σε ένα πίνακα μαζί με το πραγματικό label της κλάσης αυτής της εικόνας.

Υπολογισμός βαθμιδωτής απεικόνισης. Για τον υπολογισμό ενός HOG περιγραφέα, πρέπει πρώτα να υπολογιστούν οι οριζόντιες και κάθετες κλίσεις και τελικά, να υπολογιστεί το ιστόγραμμα των κλίσεων. Η Εικ.4.1 δείχνει το διάνυσμα χαρακτηριστικών HOG με αλληλεπικάλυψη κελιών 1 προς 1 μεταξύ των μπλοκ.

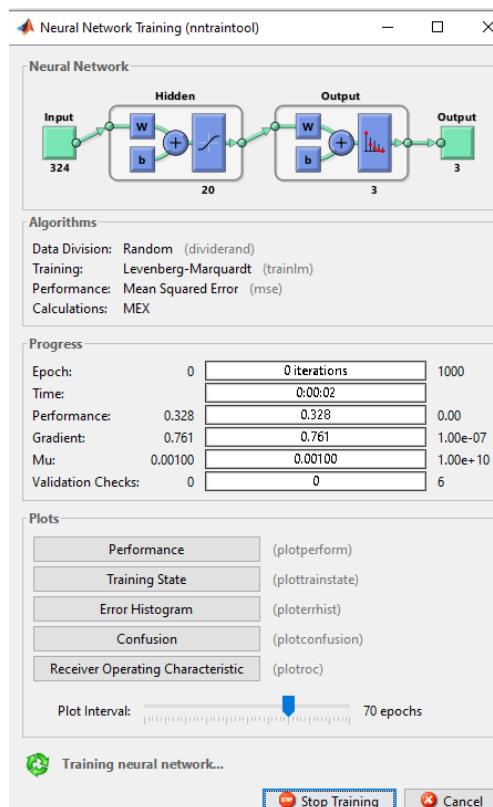


Εικόνα 4.1. HOG στοιχείο.

Βήμα 3^ο: Δημιουργία νευρωνικού δικτύου. Αφού ολοκληρωθεί αυτή η διαδικασία έχουμε πλέον τις τιμές χαρακτηριστικών για όλες τις εικόνες. Στη συνέχεια δημιουργούμε ένα νευρωνικό δίκτυο βασισμένο σε patternnet με 20 hidden layers και επιλέγουμε να ρυθμίσουμε ως τύπο δικτύου την backpropagation Levenberg-Marquardt. Επίσης, ορίζουμε τυχαία κατανομή δεδομένων με αναλογίες 70% για training, 15% για validation και 15% για testing των δεδομένων. Επίσης, ως λειτουργία μέτρησης απόδοσης ορίζεται το Mean Squared Error.

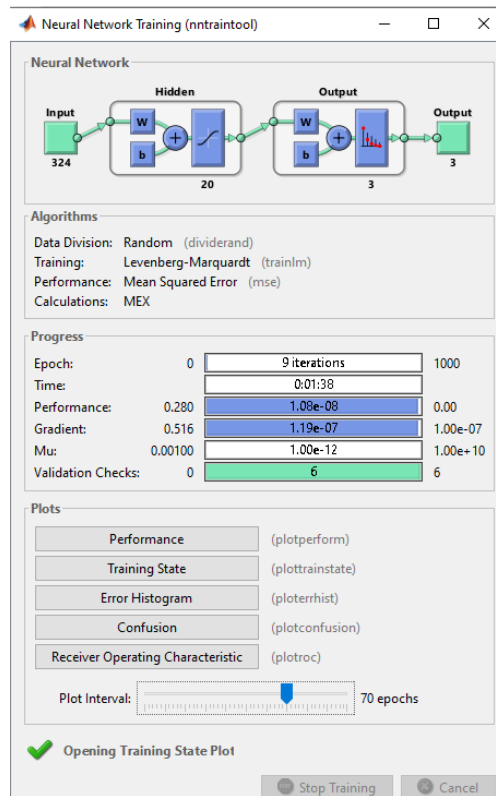
Βήμα 4^ο: Εκπαίδευση νευρωνικού δικτύου. Στη συνέχεια, το δίκτυο εκπαιδεύεται με τα χαρακτηριστικά εκπαίδευσης που έχουμε εξάγει από όλες τις εικόνες και τα αποτελέσματα μετά την εκπαίδευση εμφανίζονται στο παράθυρο εντολών Matlab και σε άλλα παράθυρα GUI.

Το παρακάτω module (Εικ.4.2) χρησιμοποιείται για τη διαμόρφωση των παραμέτρων εκπαίδευσης. Αυτές οι παράμετροι εκπαίδευσης ελέγχουν τον αριθμό των επαναλήψεων, του χρόνου και των τιμών των συντελεστών. Παρουσιάζει επίσης τα γραφήματα απόδοσης και σφάλματος.



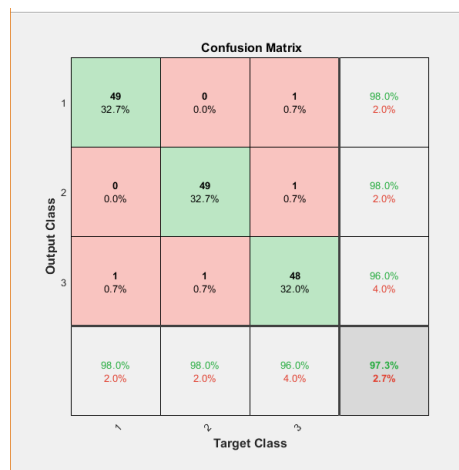
Εικόνα 4.2. Παράθυρο εντολών για εκπαίδευση νευρωνικού δικτύου.

Η εκπαίδευση γίνεται για 9 επαναλήψεις. Το output της εκπαίδευσης φαίνεται στην Εικ.4.3 και επίσης φαίνονται τα γραφήματα της διαδικασίας κατάρτισης. Το νευρωνικό δίκτυο ακολουθεί το μοντέλο εκπαίδευσης του δεδομένου προβλήματος. Σημειώνεται ότι υπάρχει ένα κρυφό στρώμα το οποίο λαμβάνει τροποποιημένους συντελεστές εισόδου και εκπαίδευσης και παράγει την έξοδο. Μπορούμε να δούμε ότι 324 εισροές ταξινομούνται σε 3 ομάδες στο τέλος.



Εικόνα 4.3. Εκπαίδευση νευρωνικών δικτύων εννιά επαναλήψεις.

Στην Εικ.4.4 απεικονίζεται ο πίνακας σύγχυσης για την περίληψη της απόδοσης του αλγόριθμου ταξινόμησης. Είναι γνωστό ότι η ακρίβεια ταξινόμησης από μόνη της μπορεί να είναι παραπλανητική σε περίπτωση άνισου αριθμού παρατηρήσεων σε κάθε κατηγορία ή εάν υπάρχουν περισσότερες από δύο κατηγορίες στο σύνολο δεδομένων.

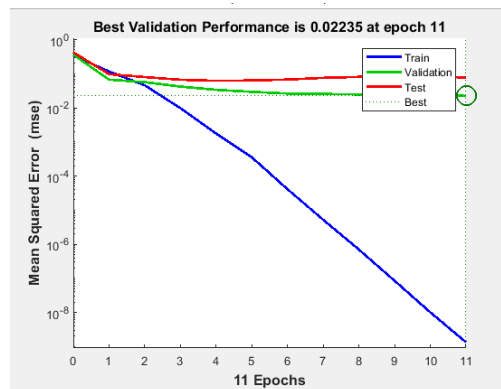


Εικόνα 4.4. Πίνακας σύγχυσης.

Στη συνέχεια επικαλείται ο κώδικας του αρχείου HOG ο οποίος κάνει τα εξής:

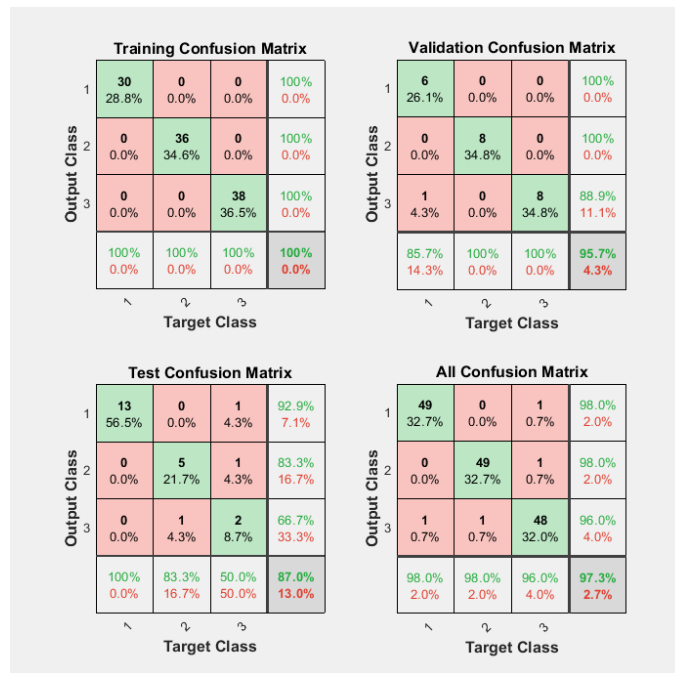
Αρχικά δημιουργείται ένας νέος φάκελος "Classified Data" με τρεις υποφακέλους "X-Ray", "Ultrasound" και "Others" για να αποθηκευτούν τα δεδομένα των ταξινομημένων αποτελεσμάτων. Στη συνέχεια, φορτώνει τις 150 εικόνες Dicom μία προς μία και ελέγχει αν όλες οι εικόνες είναι εικόνα ή βίντεο. Οι 150 εικόνες Dicom χωρίζονται σε τρεις ομάδες των 50 εικόνων η καθεμία. Οι ομάδες είναι οι "X-Ray", "Ultrasound" και "Others". Αν κάποια εικόνα αποτελεί βίντεο τότε θα βάλει αυτό το βίντεο απευθείας στο φάκελο "Others" επειδή το σύστημα είναι εκπαιδευμένο να αναγνωρίζει μόνο εικόνες. Επίσης το εκπαιδευμένο μοντέλο αποθηκεύεται ως αρχείο mat στον ίδιο φάκελο έτσι ώστε να μπορεί να φορτωθεί και να χρησιμοποιηθεί για την πρόβλεψη των άγνωστων εικόνων dicom.

Στην Εικ.4.5 απεικονίζεται η απόδοση της εκτίμησης λειτουργίας του νευρωνικού δικτύου.

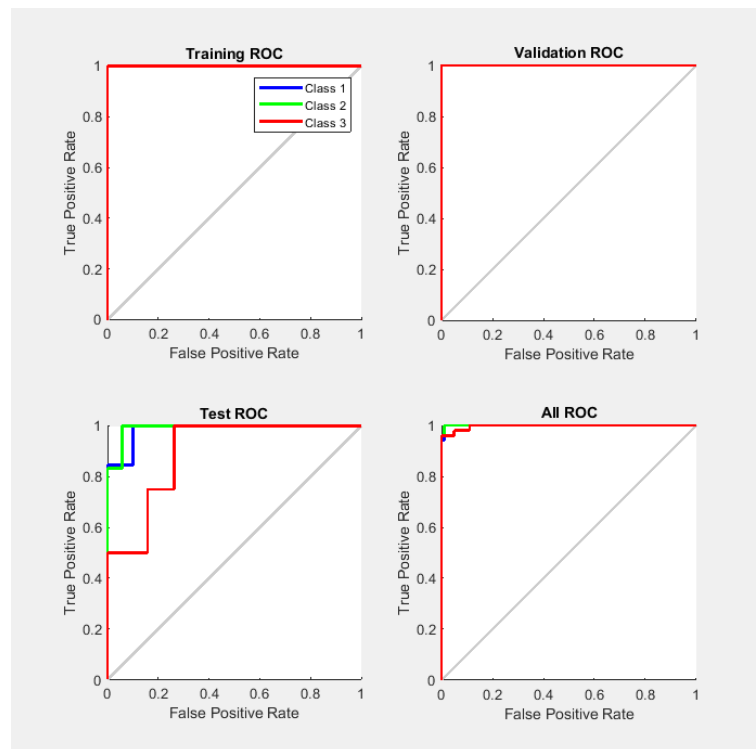


Εικόνα 4.5. Απόδοση νευρωνικού δικτύου.

Οι γραφικές παραστάσεις απόδοσης, το confusion matrix και το γράφημα ROC curves μπορούν να εμφανιστούν πατώντας το κουμπί στο παράθυρο training του νευρωνικού δικτύου μετά την ολοκλήρωση της εκπαίδευσης. Επίσης, τα αποτελέσματα για την εκπαίδευση διαφέρουν λίγο κάθε φορά που επαναλαμβάνεται η διαδικασία της εκπαίδευσης, καθώς η επιλογή των δεδομένων για το training κλπ. είναι τυχαία και διαφορετική σε κάθε νέα επανάληψη.



Εικόνα 4.6. Κελιά κατά την εκπαίδευση του νευρωνικού δικτύου.



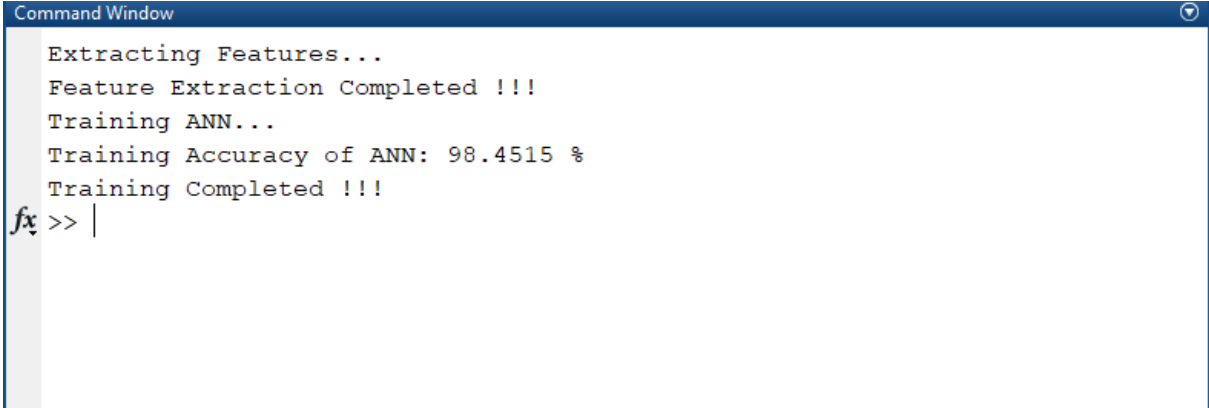
Εικόνα 4.7. Διάγραμμα ROC νευρωνικού δικτύου.

Έπειτα, φορτώνει το εκπαιδευμένο μοντέλο νευρωνικού δικτύου στο χώρο εργασίας από το φάκελο έτσι ώστε να μπορεί να χρησιμοποιηθεί για την πρόβλεψη labels των άγνωστων εικόνων. Ακολουθώς, ελέγχει αν υπάρχει έγχρωμη εικόνα RGB, και εάν υπάρχει τότε θα μετατρέψει την εικόνα σε μία νέα σε κλίμακα του γκρι και θα τροποποιήσει το μέγεθός της

στο τυπικό μέγεθος εικόνας 512x512. Επίσης θα μετατρέψει αυτή την εικόνα σε τύπο uint8 και στη συνέχεια θα εξαγάγει τα χαρακτηριστικά HOG από την εικόνα και θα προβλέψει το label της χρησιμοποιώντας το εκπαιδευμένο μοντέλο νευρωνικού δικτύου.

Μόλις ληφθεί το προβλεπόμενο label, θα εμφανιστεί η αρχική εικόνα dicom μαζί με την ετικέτα της ως τίτλος και ανάλογα με το label που θα τις έχει αποδοθεί θα αποθηκευθεί στον αντίστοιχο φάκελο αποτελεσμάτων "Classified Data". Αφού εξετασθούν, ταξινομηθούν και αποθηκευτούν όλες οι εικόνες, μπορούμε να μεταβούμε στον φάκελο "Classified Data" και μέσα στους αντίστοιχους υποφακέλους να δούμε την ταξινόμηση των εικόνων μας.

Στην Εικ. 4.8 φαίνεται η ακρίβεια της εκτέλεσης εκπαίδευσης του νευρωνικού δικτύου.



```
Command Window
Extracting Features...
Feature Extraction Completed !!!
Training ANN...
Training Accuracy of ANN: 98.4515 %
Training Completed !!!
fx >> |
```

Εικόνα 4.8. Παράθυρο εκτέλεσης εκπαίδευσης νευρωνικού δικτύου με πρόγραμμα Matlab και ακρίβεια.

Η υλοποίηση του παραπάνω διαδικασίας στο πρόγραμμα Matlab με την μορφή κώδικα παρουσιάζεται ως εξής:

```
disp('Extracting Features...');
SubFolders=dir('Training Data');%% διαβάζει όλα τα δεδομένα από την κυρίως
βάση δεδομένων του φακέλου Test Data %%
Feature_Vector=[];ind=1;%% αρχικοποίηση των μεταβλητών %%
for i=3:length(SubFolders);%% φορτώνονται όλοι οι υποφάκελοι από την κυρίως
βάση δεδομένων ένας προς ένας %%
    Classes_Names{i-2}=SubFolders(i).name;%% αποθήκευση του ονόματος της
κλάσης από το όνομα του φακέλου %%
    FolderImages=dir(strcat(cd,'\Training
Data\',SubFolders(i).name,'\*'));%% διαβάζονται όλες τις εικόνες από τον
φάκελο Training Data %%
    for j=3:length(FolderImages);%% φορτώνονται όλες οι εικόνες από τον
φάκελο Training Data μία προς μία %%
        Infi=dicominfo(strcat(cd,'\Training
Data\',SubFolders(i).name,'\',FolderImages(j).name));%% διαβάζεται η εικόνα
%%
        I=dicomread(Infi);
        if(size(I,3)==3);%% εάν η εικόνα είναι RGB μορφής %%
            I=rgb2gray(I);%% να μετατραπεί σε grayscale εικόνα %%
        end;
```



```

        I=imresize(I,[512,512]);%% αλλαγή μεγέθους της εικόνας σε 512x512
μέγεθος %%
        Feature_Vector(ind,1:324)=compute_HOG(I,128,2,9);%% εξαγωγή HOG
χαρακτηριστικών από εικόνα και αποθήκευση σε χαρακτηριστικό διάνυσμα %%
        Feature_Vector(ind,325)=i-2;%% αποθηκεύει την κλάση της εικόνας με
χαρακτηριστικές τιμές%%
        ind=ind+1;%% ενημέρωση του ευρετηρίου από χαρακτηριστικά διανύσματα
%%
%           imshow(I,[]);%% εικόνα οθόνης εκπαίδευσης %%
%           pause(0.2);%% παύση για 0.2 δευτερόλεπτα %%
%           close;%% κλείσιμο εικόνας %%
    end;
end;
disp('Feature Extraction Completed !!!');
disp('Training ANN...');
randval=randperm(size(Feature_Vector,1));%% δημιουργία τυχαίων αναμειγμένων
δείκτων create random shuffle indices %%
Feature_Vector=Feature_Vector(randval,:);%% τυχαία ανάμιξη δεδομένων
εκπαίδευσης %%
x=Feature_Vector(:,1:324)';%% εξαγωγή χαρακτηριστικών τιμών από
χαρακτηριστικό διάνυσμα %%
t=Feature_Vector(:,325)';%% εξαγωγή ετικετών των χαρακτηριστικών τιμών από
χαρακτηριστικό διάνυσμα %%
org_labels=t;%% αποθήκευση πραγματικής ετικέτας των τιμών για σύγκριση με
τις προβλεπόμενες εικέτες %%
t=ind2vec(t);%% μετατροπή των τιμών των ετικετών σε δείκτες %%
net = patternnet(20);%% δημιουργία νευρωνικού δικτύου με προς τα εμπρός
τροφοδότηση %%
net.trainFcn = 'trainlm';%% ορισμός δικτύου τύπου οπισθοδιάδοσης Levenberg-
Marquardt %%
net.divideParam.trainRatio = 70/100;%% ορισμός 70% δεδομένων που πρόκειται
να χρησιμοποιηθούν για εκπαίδευση %%
net.divideParam.valRatio = 15/100;%% ορισμός 15% δεδομένων που πρόκειται να
χρησιμοποιηθούν για επικύρωση %%
net.divideParam.testRatio = 15/100;%% ορισμός 15% δεδομένων που πρόκειται
να χρησιμοποιηθούν για δοκιμή %%
net.performFcn = 'mse'; %% ορισμός της λειτουργίας μέτρησης απόδοσης ως
Mean Squared Error %%
net.trainParam.showWindow = 1; %% παράθυρο οθόνης νευρωνικού δικτύου %%
net = train(net,x,t);%% εκπαίδευση δικτύου %%
view(net)%% προβολή δικτύου %%
y = net(x);%% test δικτύου %%
y_pr=y;
y=vec2ind(y);%% μετατροπή των προβλεπόμενων τιμών σε δείκτες υπό την μορφή
διανύσματος %%
perf = perform(net,ind2vec(org_labels),y_pr);%% υπολογισμός εκτέλεσης του
δικτύου %%
Accuracy=1-perf ;%% υπολογισμός ακρίβειας του δικτύου %%
S=sprintf('Training Accuracy of ANN: %s %%',num2str(Accuracy*100));%% οθόνη
Accuracy (ακρίβειας) στο Command window %%
disp(S);
save('net.mat','net','Classes_Names');%% αποθήκευση του εκπαιδευμένου
νευρωνικού δικτύου %%
disp('Training Completed !!!');%% μορφή οθόνης %%

```

4.1.1 Διαδικασία testing του DICOM image classification – Συμπεράσματα

Η ακρίβεια της διαδικασίας testing ενός συστήματος τεχνητών νευρωνικών δικτύων (ANN) δεν μπορεί ποτέ να είναι επιτυχής 100%, και θα πρέπει να υπάρχει κάποια διαφορά στη δοκιμή 97% επαρκώς εκπαιδευμένων δεδομένων. Σε αυτή την περίπτωση το δίκτυο ανιχνεύει με επιτυχία τις εικόνες στις περισσότερες περιπτώσεις, αλλά υπάρχουν και ορισμένες αρνητικές περιπτώσεις όπως φαίνεται στα παραδείγματα που παρατίθενται παρακάτω:

Positive X-ray:

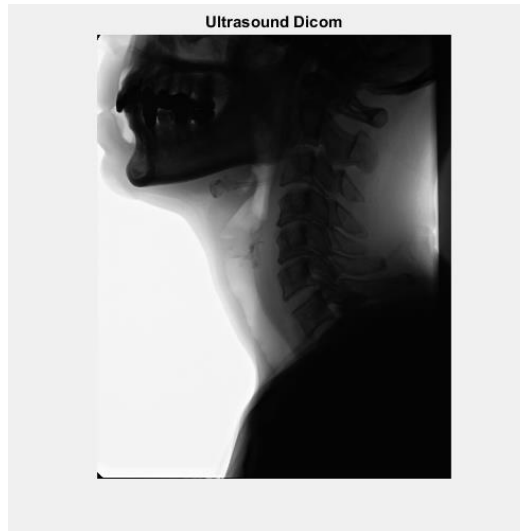
Η εικόνα ήταν μια εικόνα ακτίνων X και το δίκτυο την αναγνώρισε επίσης ως εικόνα ακτίνων X, επομένως πρόκειται για μια θετική περίπτωση (Εικ.4.9).



Εικόνα 4.9. Θετικό αποδεχόμενο X-Ray dicom.

Negative X-Ray:

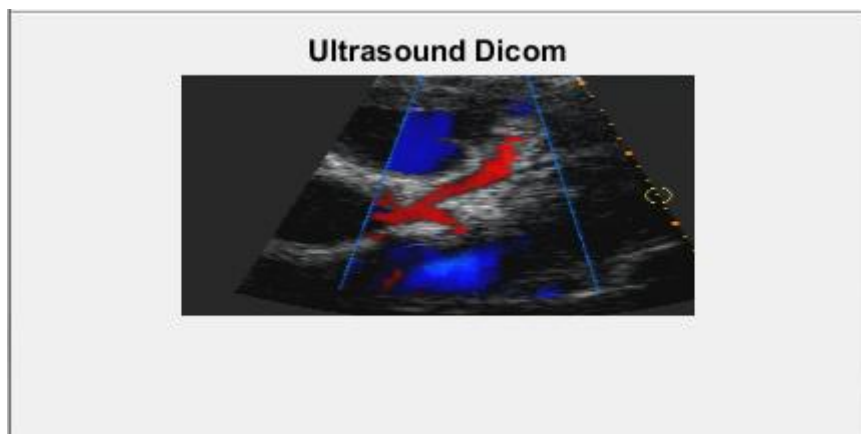
Πρόκειται για μια εικόνα ακτίνων X, αλλά το δίκτυο θεώρησε ότι είναι υπερηχογράφημα. Πρόκειται για αρνητική περίπτωση ακτινογραφίας (Εικ.4.10).



Εικόνα 4.10. Αρνητικά αποδεχόμενο DICOM.

Positive ultrasound:

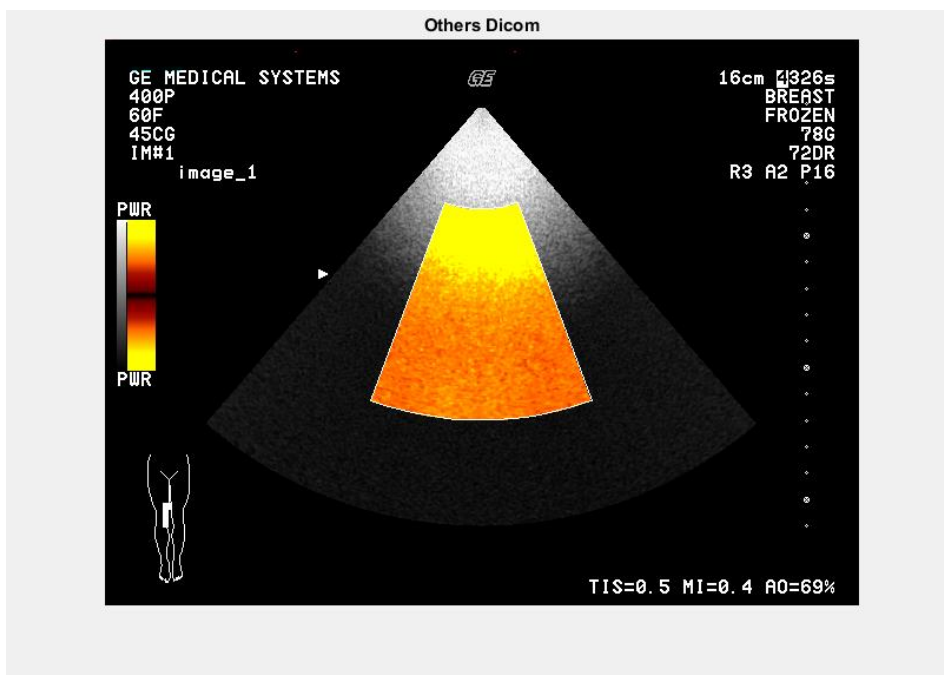
Πρόκειται για την εικόνα ενός υπερήχου και το δίκτυο το ταξινόμησε επίσης ως υπερηχογράφημα (Εικ.4.11).



Εικόνα 4.11. Εικόνα υπερήχου που ταξινομήθηκε ως υπερηχογράφημα θετικά αποδεχόμενη.

Negative ultrasound:

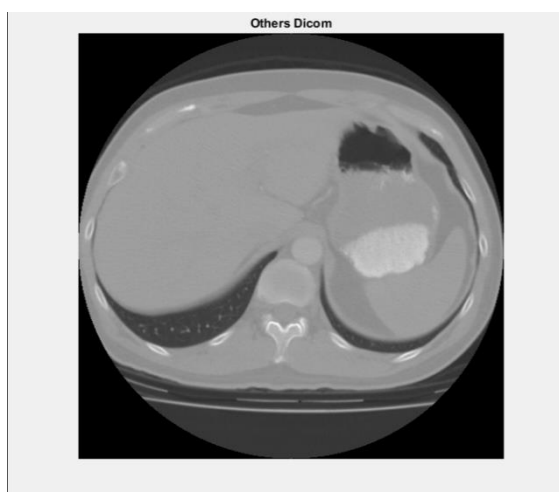
Πρόκειται για ένα υπερηχογράφημα, το οποίο όμως το δίκτυο δεν το ταξινόμησε ως ούτε ακτίνες X ούτε ως υπέρηχο (Εικ.4.12).



Εικόνα 4.12. Αρνητικά αποδεχόμενη εικόνα υπερηχογραφήματος.

Others:

Η εικόνα που δεν είναι ούτε ακτίνες X ούτε υπερηχογράφημα ταξινομήθηκε ως "Others" (Εικ.4.13).



Εικόνα 4.13. Ταξινόμηση εικόνας στην κατηγορία άλλο.

4.2 Μελλοντικές προκλήσεις

Στην περίπτωση του τομέα της ακτινολογικής απεικόνισης η εφαρμογή τεχνικών βαθιάς εκμάθησης σε ακτινολογικές εικόνες αναπτύσσεται ταχέως, με την έλλειψη ακτινολογικών δεδομένων να αποτελεί μία πρόκληση στην εφαρμογή νευρωνικών δικτύων σε ακτινολογικές εικόνες.

Η σημερινή πρόκληση στον τομέα αυτό έγκειται στο γεγονός ότι η κατάρτιση βαθιών δικτύων βασίζεται σε μεγάλα σύνολα δεδομένων. Το υπάρχον σύνολο δεδομένων απεικόνισης στον ιατρικό κόσμο είναι πιο περιορισμένο σε σύγκριση με τα ευρύτερα μη ιατρικά σύνολα δεδομένων στην περίπτωση της μηχανικής όρασης του υπολογιστή. Η βάση δεδομένων ImageNet η οποία περιέχει περισσότερες από 14 εκατομμύρια σχολιασμένες εικόνες και η βάση δεδομένων CIFAR-10, η οποία περιέχει 60.000 σχολιασμένες εικόνες αποτελούν παραδείγματα δύο μεγάλων μη ιατρικών βάσεων δεδομένων.

Σήμερα υπάρχει μεγάλη έλλειψη ακτινολογικών δεδομένων. Οι μελέτες που έχουν χρησιμοποιήσει τη βαθιά εκμάθηση στην ακτινολογία βασίζονται είτε σε ένα σύνολο ιδιωτικών δεδομένων ή σε ένα σύνολο διαθέσιμων στο κοινό δεδομένων. Η χρήση μιας δημόσιας βάσης δεδομένων προσφέρει την ευκαιρία συνεργασίας μεταξύ ερευνητών στην ακαδημαϊκή κοινότητα. Επιπλέον, τα σύνολα δημόσιων δεδομένων σχολιάζονται για να επιτρέπουν την ομοιόμορφη ετικέτα για την αξιολόγηση της απόδοσης του αλγορίθμου. Μόλις αποκτηθούν τα δεδομένα, αυτά χωρίζονται στη συνέχεια σε σύνολα εκπαίδευσης και δοκιμών.

Σήμερα είναι δύσκολο για τους ερευνητές να έχουν πρόσβαση σε εικόνες εκτός του θεσμικού τους οργάνου λόγω των προβλημάτων προστασίας της ιδιωτικής ζωής των ασθενών. Παρόλο που ένα σύστημα αρχειοθέτησης και επικοινωνίας εικόνων (PACS) για ένα μόνο ίδρυμα μπορεί να περιέχει αρκετές εκατομμύρια μελέτες, ο αριθμός των μελετών συγκεκριμένου τύπου είναι πολύ μικρότερος και μπορεί να μην είναι εφικτό να αποκτηθούν αρκετές εικόνες για την πλήρη κατάρτιση ενός δικτύου [28].

Αντίθετα, οι μη ραδιολογικές εικόνες είναι ευρέως διαθέσιμες σε πολύ μεγαλύτερο αριθμό. Οι τεχνικές που εκμεταλλεύονται τη διαθεσιμότητα μη ραδιολογικών εικόνων για να βοηθήσουν στην κατασκευή δικτύων για ταξινόμηση με ακτινολογική εικόνα μπορούν συνεπώς να επιτρέψουν τη δημιουργία μεγαλύτερων και ακριβέστερων δικτύων από ό, τι θα ήταν δυνατή χρησιμοποιώντας μόνο ακτινολογικές εικόνες [28].

Η αύξηση δεδομένων είναι μια τεχνική που χρησιμοποιείται για να ξεπεραστεί το εμπόδιο ενός περιορισμένου συνόλου δεδομένων εκπαίδευσης. Η επέκταση υλοποιείται για την τεχνητή αύξηση του αριθμού των εικόνων εκπαίδευσης. Αυτό επιτυγχάνεται με τη χρήση διαφόρων συνδυασμών πολλαπλών μετασχηματισμών που περιλαμβάνουν τεχνικές όπως περιστροφή εικόνας και αναστροφή εικόνας.

Αναφέρεται ότι ένα σημαντικό εμπόδιο στην εκμάθηση μηχανών είναι η υπερφόρτωση, με το οποίο ένα μοντέλο δεν είναι σε θέση να γενικεύσει μοτίβα πέρα από το σύνολο εκπαίδευσης. Με την αύξηση των δεδομένων, δημιουργούνται περισσότερα δεδομένα και έτσι το μοντέλο γίνεται πιο ισχυρό για ανεξάρτητες πληροφορίες σχετικά με το σετ δοκιμών.

Βιβλιογραφία

- [1] Müller B, J Reinhardt, M T Strickland, Neural networks: an introduction, Springer Science & Business Media, New York, 2012.
- [2] Shutterstock, Diagram of neuron, 2019. <https://image.shutterstock.com/image-vector/simple-neuron-diagram-2d-labeled-260nw-1341038063.jpg>. (Accessed 01/09 2019).
- [3] Gurney K, An Introduction to Neural Networks, CRC Press, New York, 2018.
- [4] Clinical gate, Electrical signaling in neurons, 2019. https://clinicalgate.com/wp-content/uploads/2015/05/B9780443103216000060_f006-001-9780443103216.jpg. (Accessed 01/09 2019).
- [5] McCulloch W S, W Pitts, A logical calculus of the ideas immanent in nervous activity, The bulletin of mathematical biophysics 5(4) (1943) 115-133.
- [6] Rosenblatt F, Principles of neurodynamics: perceptrons and the theory of brain mechanisms, Spartan Books 1962.
- [7] Minsky M, S Papert, Perceptrons: an introduction to computational geometry, MIT Press, Cambridge, Massachusetts (1969).
- [8] Rojas R, Weighted networks-the perceptron, in: Springer-Verlag (Ed.), Berlin, 1996.
- [9] Hopfield J J, Neural networks and physical systems with emergent collective computational abilities, Proceedings of the national academy of sciences 79(8) (1982) 2554-2558.
- [10] Werbos P, Beyond regression: "new tools for prediction and analysis in the behavioral sciences", Ph. D. dissertation, Harvard University (1974).
- [11] Rumelhart D E, Parallel distributed processing: Explorations in the microstructure of cognition, Learning internal representations by error propagation 1 (1986) 318-362.
- [12] Anderson J A, E Rosenfeld, A Pellionisz, Neurocomputing, MIT press 1988.
- [13] Kalogirou S A, Artificial neural networks in renewable energy systems applications: a review, Renewable and Sustainable Energy Reviews 5(4) (2001) 373-401.
- [14] Josh, Everything You Need to Know About Artificial Neural Networks, 2015. <https://medium.com/technology-invention-and-more/everything-you-need-to-know-about-artificial-neural-networks-57fac18245a1>. (Accessed 08/09 2019).
- [15] Josh, Everything you need to know about artificial neural networks, 2015. <https://medium.com/technology-invention-and-more/everything-you-need-to-know-about-artificial-neural-networks-57fac18245a1>. (Accessed 09/09 2019).
- [16] Gershenson C, Artificial neural networks for beginners, arXiv preprint cs/0308031 (2003).
- [17] Whitley D, An overview of evolutionary algorithms: practical issues and common pitfalls, Information and software technology 43(14) (2001) 817-831.
- [18] Xie J, A brief review on evolutionary computation, Control Decis 12(1) (1997) 1-7.

- [19] Yao W-s, Q Wan, Z-q Chen, J-g Wang, The researching overview of evolutionary neural networks, *Comput Sci* 31(3) (2004) 125-129.
- [20] Yao X, Y Liu, A new evolutionary system for evolving artificial neural networks, *IEEE transactions on neural networks* 8(3) (1997) 694-713.
- [21] Pianykh O S, *Digital Imaging and Communications in Medicine (DICOM): A Practical Introduction and Survival Guide*, Springer Berlin Heidelberg, Germany, 2008.
- [22] Suapang P, K Dejhan, S Yimmun, Medical image compression and DICOM-format image archive, 2009 ICCAS-SICE, IEEE, 2009, pp. 1945-1949.
- [23] Clunie D, DICOM Compression 2002, David Clunie's Medical Image Format Site.
- [24] Devi S S, K Vidhya, Development of medical image compression techniques, *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, IEEE, 2007, pp. 97-101.
- [25] Cleary K, T M Peters, Image-guided interventions: technology review and clinical applications, *Annual review of biomedical engineering* 12 (2010) 119-142.
- [26] Nagel M, G Schmidt, R Petzold, W A Kalender, A navigation system for minimally invasive CT-guided interventions, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2005, pp. 33-40.
- [27] Sennst D-A, M Kachelriess, C Leidecker, B Schmidt, O Watzke, W A Kalender, An extensible software-based platform for reconstruction and evaluation of CT images, *Radiographics* 24(2) (2004) 601-613.
- [28] Erickson B J, P Korfiatis, Z Akkus, T L Kline, Machine learning for medical imaging, *Radiographics* 37(2) (2017) 505-515.
- [29] Powles J, H Hodson, Google DeepMind and healthcare in an age of algorithms, *Health and technology* 7(4) (2017) 351-367.
- [30] Quer G, E D Muse, N Nikzad, E J Topol, S R Steinhubl, Augmenting diagnostic vision with AI, *The Lancet* 390(10091) (2017) 221.
- [31] Soffer S, A Ben-Cohen, O Shimon, M M Amitai, H Greenspan, E Klang, Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide, *Radiology* 290(3) (2019) 590-606.
- [32] Schlegl T, S M Waldstein, W-D Vogl, U Schmidt-Erfurth, G Langs, Predicting semantic descriptions from medical images with convolutional neural networks, *International Conference on Information Processing in Medical Imaging*, Springer, 2015, pp. 437-448.
- [33] Khan S, S-P Yong, A deep learning architecture for classifying medical images of anatomy object, 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), IEEE, 2017, pp. 1661-1668.
- [34] Saha S, *A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way*, 2020. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>. (Accessed 06/04 2020).

- [35] Ali S, K A Smith, On learning algorithm selection for classification, *Applied Soft Computing* 6(2) (2006) 119-138.
- [36] Jain A K, R P W Duin, J Mao, Statistical pattern recognition: A review, *IEEE transactions on pattern analysis and machine intelligence* 22(1) (2000) 4-37.
- [37] Lashari S A, R Ibrahim, A Framework for Medical Images Classification Using Soft Set, *Procedia Technology* 11 (2013) 548-556.
- [38] Bhavsar H, A Ganatra, A comparative study of training algorithms for supervised machine learning, *International Journal of Soft Computing and Engineering (IJSCE)* 2(4) (2012) 2231-2307.
- [39] Kassner A, R Thornhill, Texture analysis: a review of neurologic MR imaging applications, *American Journal of Neuroradiology* 31(5) (2010) 809-816.
- [40] Sunil R, Commonly used Machine Learning Algorithms, 2015. <https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>. 08/04/20).
- [41] Suguna N, K Thanushkodi, An improved k-nearest neighbor classification using genetic algorithm, *International Journal of Computer Science Issues* 7(2) (2010) 18-21.
- [42] Latifoğlu F, K Polat, S Kara, S Güneş, Medical diagnosis of atherosclerosis from Carotid Artery Doppler Signals using principal component analysis (PCA), k-NN based weighting pre-processing and Artificial Immune Recognition System (AIRS), *Journal of Biomedical Informatics* 41(1) (2008) 15-23.
- [43] Czarnetzki S, S Kerner, P Szcypior, Learning from demonstration-automatic generation of extended behavior networks for autonomous robots from an expert's demonstration, *International Conference on Agents and Artificial Intelligence, SCITEPRESS*, 2011, pp. 394-400.
- [44] Rajini N H, R Bhavani, Classification of MRI brain images using k-nearest neighbor and artificial neural network, 2011 *International Conference on Recent Trends in Information Technology (ICRTIT)*, IEEE, 2011, pp. 563-568.
- [45] Sengur A, An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnosis of valvular heart diseases, *Computers in biology and medicine* 38(3) (2008) 329-338.
- [46] Selvaraj H, S T Selvi, D Selvathi, L Gewali, Brain MRI slices classification using least squares support vector machine, *International Journal of Intelligent Computing in Medical Sciences & Image Processing* 1(1) (2007) 21-33.
- [47] Song Y, W Cai, H Huang, Y Zhou, D D Feng, Y Wang, M J Fulham, M Chen, Large margin local estimate with applications to medical image classification, *IEEE transactions on medical imaging* 34(6) (2015) 1362-1377.
- [48] Verma J, M Nath, P Tripathi, K Saini, Analysis and identification of kidney stone using K th nearest neighbour (KNN) and support vector machine (SVM) classification techniques, *Pattern Recognition and Image Analysis* 27(3) (2017) 574-580.
- [49] Organization W H, 2001, Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children,

- [50] Esteva A, B Kuprel, R A Novoa, J Ko, S M Swetter, H M Blau, S Thrun, Dermatologist-level classification of skin cancer with deep neural networks, *nature* 542(7639) (2017) 115-118.
- [51] Amelio L, A Amelio, Classification methods in image analysis with a special focus on medical analytics, *Machine Learning Paradigms*, Springer2019, pp. 31-69.
- [52] Vishnuvarthanan G, M P Rajasekaran, P Subbaraj, A Vishnuvarthanan, An unsupervised learning method with a clustering approach for tumor identification and tissue segmentation in magnetic resonance brain images, *Applied Soft Computing* 38 (2016) 190-212.
- [53] Juang L-H, M-N Wu, MRI brain lesion image detection based on color-converted K-means clustering segmentation, *Measurement* 43(7) (2010) 941-949.
- [54] Zhang D-Q, S-C Chen, A novel kernelized fuzzy c-means algorithm with application in medical image segmentation, *Artificial intelligence in Medicine* 32(1) (2004) 37-50.
- [55] Singh D, K Kaur, Classification of abnormalities in brain MRI images using GLCM, PCA and SVM, *International Journal of Engineering and Advanced Technology (IJEAT)* 1(6) (2012) 243-248.
- [56] Ramteke R, Y K Monali, Automatic medical image classification and abnormality detection using k-nearest neighbour, *International Journal of Advanced Computer Research* 2(4) (2012) 190-196.
- [57] Corso J J, E Sharon, S Dube, S El-Saden, U Sinha, A Yuille, Efficient multilevel brain tumor segmentation with integrated bayesian model classification, *IEEE transactions on medical imaging* 27(5) (2008) 629-640.
- [58] Rao A, Y Lee, A Gass, A Monsch, Classification of Alzheimer's Disease from structural MRI using sparse logistic regression with optional spatial regularization, 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2011, pp. 4499-4502.
- [59] Niwas S I, P Palanisamy, K Sujathan, Complex wavelet as nucleus descriptors for automated cancer cytology classifier system using ANN, 2010 IEEE International Conference on Computational Intelligence and Computing Research, IEEE, 2010, pp. 1-5.
- [60] Camlica Z, H R Tizhoosh, F Khalvati, Medical image classification via svm using lbp features from saliency-based folded data, 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), IEEE, 2015, pp. 128-132.
- [61] Zhang Y D, S Chen, S H Wang, J F Yang, P Phillips, Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine, *International Journal of Imaging Systems and Technology* 25(4) (2015) 317-327.
- [62] Abbass H A, An evolutionary artificial neural networks approach for breast cancer diagnosis, *Artificial intelligence in Medicine* 25(3) (2002) 265-281.
- [63] Karabatak M, M C Ince, An expert system for detection of breast cancer based on association rules and neural network, *Expert systems with Applications* 36(2) (2009) 3465-3469.
- [64] Rawat W, Z Wang, Deep convolutional neural networks for image classification: A comprehensive review, *Neural computation* 29(9) (2017) 2352-2449.

- [65] Kermany D S, M Goldbaum, W Cai, C C Valentim, H Liang, S L Baxter, A McKeown, G Yang, X Wu, F Yan, Identifying medical diagnoses and treatable diseases by image-based deep learning, *Cell* 172(5) (2018) 1122-1131. e9.
- [66] Li Q, W Cai, X Wang, Y Zhou, D D Feng, M Chen, Medical image classification with convolutional neural network, 2014 13th international conference on control automation robotics & vision (ICARCV), IEEE, 2014, pp. 844-848.
- [67] Wang X, Y Peng, L Lu, Z Lu, M Bagheri, R M Summers, Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2097-2106.
- [68] Cardoso M J, T Arbel, S Lee, V Cheplygina, S Balocco, D Mateus, G Zahnd, L Maier-Hein, S Demirci, E Granger, *Intravascular Imaging and Computer Assisted Stenting, and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis, CVII-STENT and Second International Workshop, LABELS*, Springer, 2017.
- [69] Wang X, Y Peng, L Lu, Z Lu, R M Summers, Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9049-9058.
- [70] Deng J, W Dong, R Socher, L-J Li, K Li, L Fei-Fei, Imagenet: A large-scale hierarchical image database, 2009 IEEE conference on computer vision and pattern recognition, Ieee, 2009, pp. 248-255.
- [71] Frid-Adar M, I Diamant, E Klang, M Amitai, J Goldberger, H Greenspan, GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification, *Neurocomputing* 321 (2018) 321-331.
- [72] Perez L, J Wang, The effectiveness of data augmentation in image classification using deep learning, *arXiv preprint arXiv:1712.04621* (2017).
- [73] Sabour S, N Frosst, G E Hinton, Dynamic routing between capsules, *Advances in neural information processing systems*, 2017, pp. 3856-3866.
- [74] LeCun Y, Cortes C, B C., Mnist handwritten digit database, at&t labs. (2010).
- [75] Afshar P, A Mohammadi, K N Plataniotis, Brain tumor type classification via capsule networks, 2018 25th IEEE International Conference on Image Processing (ICIP), IEEE, 2018, pp. 3129-3133.
- [76] Kato H, M Kanematsu, X Zhang, M Saio, H Kondo, S Goshima, H Fujita, Computer-aided diagnosis of hepatic fibrosis: preliminary evaluation of MRI texture analysis using the finite difference method and an artificial neural network, *American Journal of Roentgenology* 189(1) (2007) 117-122.
- [77] Greenspan H, B Van Ginneken, R M Summers, Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique, *IEEE transactions on medical imaging* 35(5) (2016) 1153-1159.
- [78] Toriwaki J-I, Y Suenaga, T Negoro, T Fukumura, Pattern recognition of chest X-ray images, *Computer Graphics and Image Processing* 2(3-4) (1973) 252-271.

- [79] Meyers P H, C M Nice Jr, H C Becker, W J Nettleton Jr, J W Sweeney, G R Meckstroth, Automated computer analysis of radiographic images, *Radiology* 83(6) (1964) 1029-1034.
- [80] LeCun Y, Y Bengio, G Hinton, Deep learning, *nature* 521(7553) (2015) 436-444.
- [81] Baltruschat I M, H Nickisch, M Grass, T Knopp, A Saalbach, Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification, *Scientific Reports* 9(1) (2019) 6381.
- [82] Xie Y, D Richmond, Pre-training on grayscale imagenet improves medical image classification, *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 0-0.
- [83] Li Y, H Lu, L Zhang, S Yang, S Serikawa, Color image segmentation using fast density-based clustering method, *Future Communication, Computing, Control and Management, Springer2012*, pp. 593-598.
- [84] Olsen R I, J Gates, D L Sato, Digital cameras with direct luminance and chrominance detection, *Google Patents*, 2014.
- [85] Liu C, J Yang, ICA color space for pattern recognition, *IEEE transactions on neural networks* 20(2) (2008) 248-257.
- [86] Wu D, D-W Sun, Colour measurements by computer vision for food quality control—A review, *Trends in Food Science & Technology* 29(1) (2013) 5-20.
- [87] Menesatti P, C Angelini, F Pallottino, F Antonucci, J Aguzzi, C Costa, RGB color calibration for quantitative image analysis: The “3D Thin-Plate Spline” warping approach, *Sensors* 12(6) (2012) 7063-7079.
- [88] Lissner I, J Preiss, P Urban, M S Lichtenauer, P Zolliker, Image-difference prediction: From grayscale to color, *IEEE Transactions on Image Processing* 22(2) (2012) 435-446.
- [89] Hua X, Human computer interactions for converting color images to gray, *Neurocomputing* 85 (2012) 1-5.
- [90] Pascale D, A review of rgb color spaces... from xyy to r'g'b', *Babel Color* 18 (2003) 136-152.
- [91] Bala R, R Eschbach, Spatial color-to-grayscale transform preserving chrominance edge information, *Color and Imaging Conference, Society for Imaging Science and Technology*, 2004, pp. 82-86.
- [92] Rasche K, R Geist, J Westall, Detail preserving reproduction of color images for monochromats and dichromats, *IEEE Computer Graphics and Applications* 25(3) (2005) 22-30.
- [93] Smith K, P E Landes, J Thollot, K Myszkowski, Apparent greyscale: A simple and fast conversion to perceptually accurate images and video, *Computer Graphics Forum, Wiley Online Library*, 2008, pp. 193-200.
- [94] Grundland M, N A Dodgson, The decolorize algorithm for contrast enhancing, color to grayscale conversion, *Technical Report—University of Cambridge. Computer Laboratory*, XX, XX, vol. UCAM-CL-TR 649 (2005) 1-15.
- [95] Gooch A A, S C Olsen, J Tumblin, B Gooch, Color2gray: salience-preserving color removal, *ACM Transactions on Graphics (TOG)*, ACM, 2005, pp. 634-639.

- [96] Kim Y, C Jang, J Demouth, S Lee, Robust color-to-gray via nonlinear global mapping, *ACM Transactions on Graphics (TOG)* 28(5) (2009) 161.
- [97] Lu J, K N Plataniotis, On conversion from color to gray-scale images for face detection, 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, IEEE, 2009, pp. 114-119.
- [98] Jones C F, A L Abbott, Optimization of color conversion for face recognition, *EURASIP Journal on Advances in Signal Processing* 2004(4) (2004) 948790.
- [99] Serag A, G Macnaught, F C Denison, R M Reynolds, S I Semple, J P Boardman, Histograms of oriented 3D gradients for fully automated fetal brain localization and robust motion correction in 3 T magnetic resonance images, *BioMed research international* 2017 (2017).
- [100] Mikolajczyk K, C Schmid, A performance evaluation of local descriptors, (2005).
- [101] Lowe D G, Distinctive image features from scale-invariant keypoints, *International journal of computer vision* 60(2) (2004) 91-110.
- [102] Ojala T, M Pietikäinen, T Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis & Machine Intelligence* (7) (2002) 971-987.
- [103] Dalal N, B Triggs, Histograms of oriented gradients for human detection, 2005.
- [104] Teague M R, Image analysis via the general theory of moments, *JOSA* 70(8) (1980) 920-930.
- [105] Dalal N, B Triggs, Histograms of oriented gradients for human detection, 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), IEEE, 2005, pp. 886-893.
- [106] Onishi K, T Takiguchi, Y Arika, 3D human posture estimation using the HOG features from monocular image, 2008 19th International Conference on Pattern Recognition, IEEE, 2008, pp. 1-4.
- [107] Korkmaz S A, H Bínol, A Akçiçek, M F Korkmaz, A expert system for stomach cancer images with artificial neural network by using HOG features and linear discriminant analysis: HOG_LDA_ANN, 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY), IEEE, 2017, pp. 000327-000332.
- [108] Raghavendra U, U Rajendra Acharya, E Y K Ng, J-H Tan, A Gudigar, An integrated index for breast cancer identification using histogram of oriented gradient and kernel locality preserving projection features extracted from thermograms, *Quantitative InfraRed Thermography Journal* 13(2) (2016) 195-209.
- [109] Farshbaf M, A A Pouyan, Landmark detection on cephalometric radiology images through combining classifiers, 2010 17th Iranian Conference of Biomedical Engineering (ICBME), 2010, pp. 1-4.
- [110] Yao J, B Glocker, T Klinder, S Li, Recent advances in computational methods and clinical applications for spine imaging, (2015).

- [111] Daenzer S, S Freitag, S von Sachsen, H Steinke, M Groll, J Meixensberger, M Leimert, VolHOG: a volumetric object recognition approach based on bivariate histograms of oriented gradients for vertebra detection in cervical spine MRI, *Medical physics* 41(8Part1) (2014) 082305.
- [112] Agarwal D, K S Shriram, N Subramanian, Automatic view classification of echocardiograms using Histogram of Oriented Gradients, 2013 IEEE 10th International Symposium on Biomedical Imaging, 2013, pp. 1368-1371.
- [113] Dalal N, B Triggs, C Schmid, Human detection using oriented histograms of flow and appearance, *European conference on computer vision*, Springer, 2006, pp. 428-441.
- [114] Satya M, Histogram of oriented gradients, 2016. <https://www.learnopencv.com/histogram-of-oriented-gradients/>. (Accessed 28/09 2019).
- [115] National Library of Medicine, Tuberculosis Chest X-ray Image Data Sets, 2019. <https://ceb.nlm.nih.gov/repositories/tuberculosis-chest-x-ray-image-data-sets/>. (Accessed 22/12 2019).
- [116] When training AI to classify chest x-rays, is more data always better? <https://www.radiologybusiness.com/topics/artificial-intelligence/ai-chest-x-rays-convolutional-neural-network>