



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ

# Εντοπισμός Καρδιακών Ανωμαλιών σε Χρονοσειρές Χρησιμοποιώντας Τεχνητά Νευρωνικά Δίκτυα

[Ταξινόμηση Καρδιακών Αρρυθμιών από ECG μέσω GAF  
Εικόνων και Συνελκτικών Νευρωνικών Δικτύων]

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ της

Μανωλοπούλου Ευανθίας

Επιβλέπων Καθηγητής:

Τασουλής Σωτήριος

Αναπληρωτής Καθηγητής

Λαμία, Ιούνιος 2025





# Εντοπισμός Καρδιακών Ανωμαλιών σε Χρονοσειρές Χρησιμοποιώντας Τεχνητά Νευρωνικά Δίκτυα

[Ταξινόμηση Καρδιακών Αρρυθμιών από ECG μέσω GAF  
Εικόνων και Συνελικτικών Νευρωνικών Δικτύων]

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ της

Μανωλοπούλου Ευανθίας

Επιβλέπων:

Τασουλής Σωτήριος

Αναπληρωτής Καθηγητής

Εγκρίθηκε από την Τριμελή Εξεταστική Επιτροπή, Ιούνιος 2025.

---

(Υπογραφή)

Τασουλής Σωτήριος

Επιβλέπων Καθηγητής

---

(Υπογραφή)

---

(Υπογραφή)

Λαμία, Ιούνιος 2025



Copyright © – All rights reserved  
Με την επιφύλαξη παντός δικαιώματος

Μανωλοπούλου Ευανθία, 2025

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της επιτροπής που την ενέκρινε.

#### ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ ΚΑΙ ΑΝΑΛΗΨΗΣ ΠΡΟΣΩΠΙΚΗΣ ΕΥΘΥΝΗΣ

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ενυπογράφως ότι είμαι αποκλειστικός συγγραφέας της παρούσας Πτυχιακής Εργασίας, για την ολοκλήρωση της οποίας κάθε βοήθεια είναι πλήρως αναγνωρισμένη και αναφέρεται λεπτομερώς στην εργασία αυτή. Έχω αναφέρει πλήρως και με σαφείς αναφορές, όλες τις πηγές χρήσης δεδομένων, απόψεων, θέσεων και προτάσεων, ιδεών και λεκτικών αναφορών, είτε κατά κυριολεξία είτε βάσει επιστημονικής παράφρασης.

Αναλαμβάνω την προσωπική και ατομική ευθύνη ότι σε περίπτωση αποτυχίας στην υλοποίηση των ανωτέρω δηλωθέντων στοιχείων, είμαι υπόλογος έναντι λογοκλοπής, γεγονός που σημαίνει αποτυχία στην Πτυχιακή μου Εργασία και κατά συνέπεια αποτυχία απόκτησης του Τίτλου Σπουδών, πέραν των λοιπών συνεπειών του νόμου περί πνευματικών δικαιωμάτων. Δηλώνω, συνεπώς, ότι αυτή η Πτυχιακή Εργασία προετοιμάστηκε και ολοκληρώθηκε από εμένα προσωπικά και αποκλειστικά και ότι αναλαμβάνω πλήρως όλες τις συνέπειες του νόμου στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής άλλης πνευματικής ιδιοκτησίας.

---

(Υπογραφή)

Μανωλοπούλου Ευανθία Λαμία, Ιούνιος 2025

---

# 1 Περίληψη

Τα τελευταία χρόνια, η εμφάνιση των νευρωνικών δικτύων και της μηχανικής μάθησης στον τομέα της ιατρικής έχει επιφέρει καθοριστικές αλλαγές στην ανάλυση, πρόβλεψη και διάγνωση διαφόρων παθήσεων. Μερικές από τις πιο σημαντικές εφαρμογές τους περιλαμβάνουν την ανάλυση ιατρικών εικόνων για εξαγωγή διαγνωστικών συμπερασμάτων, την υποστήριξη ιατρικών ρομποτικών και χειρουργικών διαδικασιών μέσω αυτοματισμού υψηλής ακρίβειας, καθώς και την ανίχνευση και επεξεργασία βιολογικών σημάτων, όπως τα ηλεκτροκαρδιογραφήματα (ECG) και τα ηλεκτροεγκεφαλογράφηματα (EEG).

Αντικείμενο της παρούσας εργασίας είναι η ανάπτυξη ενός μοντέλου ταξινόμησης καρδιακών παλμών μέσω της χρήσης Συνελικτικού Νευρωνικού Δικτύου (CNN), με στόχο τον εντοπισμό παθολογικών μοτίβων σε σήματα ECG. Η προσέγγιση βασίζεται στη μετατροπή των καρδιογραφικών χρονοσειρών σε δισδιάστατες εικόνες με χρήση του μετασχηματισμού Gramian Angular Field (GAF), προκειμένου να καταστούν κατάλληλες για επεξεργασία από το CNN. Το προτεινόμενο pipeline περιλαμβάνει εντοπισμό R-peaks, RR-based segmentation, δημιουργία GAF εικόνων, εκπαίδευση του μοντέλου και αξιολόγηση της απόδοσής του.

Κατά την εκπαίδευση του μοντέλου δεν εφαρμόστηκαν τεχνικές εξισορρόπησης (όπως SMOTE ή data augmentation), παρά την άνιση κατανομή των κατηγοριών στο σύνολο δεδομένων, γεγονός που ενισχύει τον στόχο αξιολόγησης της εγγενούς ικανότητας του μοντέλου. Για τη βελτιστοποίηση της απόδοσης χρησιμοποιήθηκαν Grid Search και Optuna για επιλογή υπερπαραμέτρων και κατάλληλων κατωφλίων πρόβλεψης, χωρίς να εισαχθεί επιπλέον πολυπλοκότητα στην αρχιτεκτονική.

Τα πειραματικά αποτελέσματα δείχνουν ότι ένα σχετικά απλό και υπολογιστικά αποδοτικό CNN μπορεί να επιτύχει υψηλό βαθμό ακρίβειας και ισορροπίας, ακόμη και χωρίς προηγμένες τεχνικές ενίσχυσης των μειονοτικών κλάσεων. Η απόδοση του μοντέλου κρίνεται ικανοποιητική τόσο σε συνολικό accuracy όσο και σε metrics όπως F1-score και recall, ιδίως για τις κατηγορίες υψηλού ενδιαφέροντος, δηλαδή για όσες εμφανίζονται λιγότερο.

Συμπερασματικά, τα ευρήματα καταδεικνύουν ότι η μετατροπή σημάτων ECG σε εικόνες, σε συνδυασμό με CNN και επιλογή υπερπαραμέτρων, μπορεί να προσφέρει μια αποτελεσματική και πρακτική λύση για την αυτόματη ανίχνευση καρδιολογικών ανωμαλιών, χωρίς την ανάγκη πολύπλοκων μοντέλων ή πρόσθετης επεξεργασίας των δεδομένων.

**Λέξεις-κλειδιά:** Ηλεκτροκαρδιογράφημα, ECG, CNN, GAF, AAMI, MIT-BIH, Ταξινόμηση Αρρυθμιών

---

## 2 Abstract

In recent years, the emergence of neural networks and machine learning in the field of medicine has led to transformative changes in the analysis, prediction, and diagnosis of various diseases. Some of their most significant applications include the analysis of medical images to extract diagnostic insights, the support of robotic and surgical procedures through high-precision automation, and the detection and processing of biological signals such as electrocardiograms (ECG) and electroencephalograms (EEG).

The objective of this study is the development of a heartbeat classification model using a Convolutional Neural Network (CNN), aimed at identifying pathological patterns in ECG signals. The proposed approach is based on the transformation of ECG time series into two-dimensional images using the Gramian Angular Field (GAF) transformation, in order to make them suitable for processing by the CNN. The overall pipeline involves R-peak detection, RR-based segmentation, GAF image generation, model training, and performance evaluation.

No class balancing techniques (such as SMOTE or data augmentation) were applied during training, despite the imbalanced nature of the dataset. This was done intentionally to assess the model's inherent capacity to handle class imbalance. To improve performance, Grid Search and Optuna were employed for hyperparameter tuning and optimal threshold selection, without introducing additional complexity into the network architecture.

The experimental results demonstrate that a relatively simple and computationally efficient CNN can achieve a high level of accuracy and balance, even without advanced techniques for boosting minority classes. The model's performance is considered satisfactory, both in terms of overall accuracy and metrics such as F1-score and recall, especially for clinically significant categories.

In conclusion, the findings suggest that transforming ECG signals into images, combined with a well-designed CNN and careful hyperparameter tuning, can provide an effective and practical solution for the automatic detection of cardiac anomalies—without requiring complex models or additional preprocessing steps.

**Keywords:** Electrocardiogram, ECG, CNN, GAF, AAMI, MIT-BIH, Arrhythmia Classification

---

### 3 Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον καθηγητή κ. Τασουλή Σωτήριο, για τη διδασκαλία των μαθημάτων του και των projects, τα οποία με ενέπνευσαν να ασχοληθώ με το πεδίο της Μηχανικής Μάθησης.

Επίσης, ευχαριστώ θερμά την Δρ. Τσακμάκη Παρασκευή για την καθοδήγησή της και την άριστη συνεργασία που είχαμε καθ' όλη τη διάρκεια της εκπόνησης της παρούσας πτυχιακής εργασίας.

Λαμία, Ιούνιος 2025  
Μανωλοπούλου Ευανθία

## Περιεχόμενα

|       |                                                         |    |
|-------|---------------------------------------------------------|----|
| 1     | Περίληψη                                                | 2  |
| 2     | Abstract                                                | 3  |
| 3     | Ευχαριστίες                                             | 4  |
| 4     | Εισαγωγή                                                | 7  |
| 5     | Σχετικές έρευνες                                        | 7  |
| 6     | Θεωρία                                                  | 8  |
| 6.1   | Βασικές έννοιες και προσεγγίσεις                        | 8  |
| 6.1.1 | Ηλεκτροκαρδιογράφημα (ECG) και καρδιολογικές διαταραχές | 8  |
| 6.1.2 | Κατηγορία N – Φυσιολογικοί (Μη-εκτοπικοί) Παλμοί        | 10 |
| 6.1.3 | Κατηγορία S – Υπερκοιλιακά Έκτοπα                       | 10 |
| 6.1.4 | Κατηγορία V – Κοιλιακοί Έκτακτοι Παλμοί                 | 10 |
| 6.1.5 | Κατηγορία F – Συγχωνευμένοι Παλμοί                      | 11 |
| 6.1.6 | Κατηγορία Q – Άγνωστοι ή Μη-Ταξινομήσιμοι Παλμοί        | 11 |
| 6.1.7 | Gramian Angular Field (GAF)                             | 12 |
| 6.2   | Αρχιτεκτονική και λειτουργία των CNNs                   | 13 |
| 6.2.1 | Αρχιτεκτονική                                           | 13 |
| 6.2.2 | Activation Functions                                    | 14 |
| 6.2.3 | Pooling Layers                                          | 15 |
| 6.2.4 | Fully Connected Layers                                  | 15 |
| 6.2.5 | Batch Normalization                                     | 15 |
| 6.3   | Εκπαίδευση Δικτύου και Συνάρτηση Απώλειας               | 16 |
| 6.4   | Βελτιστοποίηση και υπερπαραμέτροι                       | 16 |
| 6.4.1 | Grid Search                                             | 16 |
| 6.4.2 | Optuna                                                  | 16 |
| 6.4.3 | Threshold Tuning                                        | 17 |
| 6.5   | Κανονικοποίηση και προεπεξεργασία                       | 18 |
| 6.5.1 | Min-Max Normalization                                   | 18 |
| 6.5.2 | Resampling                                              | 18 |
| 6.5.3 | Zero-padding                                            | 19 |
| 6.6   | Μετρικές Αξιολόγησης                                    | 20 |
| 6.6.1 | Accuracy                                                | 20 |
| 6.6.2 | Precision                                               | 21 |
| 6.6.3 | Recall                                                  | 21 |

|           |                                                               |           |
|-----------|---------------------------------------------------------------|-----------|
| 6.6.4     | F1-score . . . . .                                            | 22        |
| <b>7</b>  | <b>Μεθοδολογία</b>                                            | <b>22</b> |
| 7.1       | Δεδομένα – MIT-BIH Arrhythmia Database . . . . .              | 23        |
| 7.2       | Προεπεξεργασία δεδομένων . . . . .                            | 24        |
| 7.2.1     | Χρήση AAMI . . . . .                                          | 24        |
| 7.2.2     | Ανίχνευση R-peaks και RR-based Segmentation . . . . .         | 25        |
| 7.2.3     | Μετασχηματισμός GAF . . . . .                                 | 26        |
| 7.2.4     | Δημιουργία και επεξεργασία εισόδου για το νευρωνικό . . . . . | 28        |
| 7.3       | Αρχιτεκτονική του Νευρωνικού Δικτύου (CNN) . . . . .          | 29        |
| 7.4       | Παραλλαγές Μοντέλων και Βελτιστοποιήσεις . . . . .            | 30        |
| 7.4.1     | Υλοποίηση προσθήκης Optuna για βάρη . . . . .                 | 30        |
| 7.4.2     | Υλοποίηση grid search για κατώφλια . . . . .                  | 30        |
| 7.5       | Overfitting . . . . .                                         | 31        |
| 7.5.1     | L1 Regularization . . . . .                                   | 31        |
| 7.5.2     | L2 Regularization . . . . .                                   | 32        |
| <b>8</b>  | <b>Αποτελέσματα</b>                                           | <b>32</b> |
| 8.1       | Πειραματικές ρυθμίσεις . . . . .                              | 32        |
| 8.2       | CNN μοντέλο χωρίς βάρη και κατώφλι με φίλτρο . . . . .        | 33        |
| 8.3       | CNN μοντέλο χωρίς βάρη και κατώφλι και χωρίς φίλτρο . . . . . | 34        |
| 8.4       | CNN μοντέλο με βάρη και φίλτρο . . . . .                      | 34        |
| 8.5       | CNN μοντέλο με βάρη χωρίς φίλτρο . . . . .                    | 35        |
| 8.6       | CNN μοντέλο με fine tuning βαρών και φίλτρο . . . . .         | 36        |
| 8.7       | CNN μοντέλο με fine tuning βαρών χωρίς φίλτρο . . . . .       | 36        |
| 8.8       | CNN μοντέλο με βάρη, κατώφλι και φίλτρο . . . . .             | 37        |
| 8.9       | CNN μοντέλο με βάρη, κατώφλι χωρίς φίλτρο . . . . .           | 38        |
| 8.10      | Σύγκριση αποτελεσμάτων με και χωρίς φίλτρο . . . . .          | 38        |
| 8.11      | Σύγκριση αποτελεσμάτων των 8 πειραμάτων . . . . .             | 39        |
| 8.12      | Σύγκριση αποτελεσμάτων με άλλες έρευνες . . . . .             | 40        |
| <b>9</b>  | <b>Συζήτηση</b>                                               | <b>41</b> |
| <b>10</b> | <b>Συμπεράσματα και μελλοντικές έρευνες</b>                   | <b>43</b> |
|           | <b>Βιβλιογραφία</b>                                           | <b>44</b> |

---

## 4 Εισαγωγή

Οι καρδιακές αρρυθμίες αποτελούν διαταραχές του φυσιολογικού καρδιακού ρυθμού, οι οποίες μπορεί να εκδηλωθούν ως επιτάχυνση (ταχυκαρδία), επιβράδυνση (βραδυκαρδία) ή ακανόνιστος ρυθμός. Αν και ορισμένες είναι καλοήθειες, άλλες μπορεί να συνδέονται με σοβαρές παθολογίες όπως καρδιακή ανεπάρκεια, αγγειακά εγκεφαλικά επεισόδια ή αιφνίδιο καρδιακό θάνατο [1, 2].

Η έγκαιρη ανίχνευση των αρρυθμιών είναι ζωτικής σημασίας για την πρόληψη επιπλοκών και την αύξηση του προσδόκιμου επιβίωσης. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, οι καρδιαγγειακές παθήσεις παραμένουν η κυριότερη αιτία θανάτου παγκοσμίως, προκαλώντας περισσότερους από 17 εκατομμύρια θανάτους ετησίως [3].

Το ηλεκτροκαρδιογράφημα (ECG) είναι μια μη επεμβατική μέθοδος καταγραφής της ηλεκτρικής δραστηριότητας της καρδιάς με τη χρήση επιφανειακών ηλεκτροδίων [4]. Ο εντοπισμός ανωμαλιών στο ΗΚΓ (Ηλεκτροκαρδιογράφημα) αφορά τον εντοπισμό περιοχών που παρεκκλίνουν από τον φυσιολογικό ρυθμό και μπορεί να σχετίζονται με παθολογικές καταστάσεις.

Σκοπός της παρούσας εργασίας είναι η αξιοποίηση ενός συνελκτικού νευρωνικού δικτύου (CNN) με σχετικά απλή αρχιτεκτονική, για την ταξινόμηση καρδιακών παλμών σε πέντε κατηγορίες σύμφωνα με το πρότυπο AAMI.

Τα δεδομένα του MIT-BIH Database που χρησιμοποιούνται, παρουσιάζουν ανισορροπία στις κλάσεις και παρότι στην πλειοψηφία των σχετικών μελετών γίνεται χρήση τεχνικών εξισορρόπησης, όπως το data augmentation ή μεθόδων υπερδειγματοληψίας (π.χ. SMOTE), στην παρούσα εργασία δεν εφαρμόστηκαν τεχνητές ενισχύσεις στις υποεκπροσωπούμενες κατηγορίες. Ο στόχος ήταν να εξεταστεί η απόδοση του μοντέλου χωρίς παρέμβαση στα δεδομένα εκπαίδευσης, αξιοποιώντας μόνο τις ενσωματωμένες ιδιότητες του νευρωνικού δικτύου και τεχνικές επιλογής υπερπαραμέτρων.

Για την αξιολόγηση της επίδοσης του μοντέλου, χρησιμοποιήθηκαν αποκλειστικά μέθοδοι βελτιστοποίησης υπερπαραμέτρων και προσαρμογής κατωφλίων πρόβλεψης.

Η δομή της εργασίας περιλαμβάνει το θεωρητικό υπόβαθρο, τη μεθοδολογική προσέγγιση, τα πειραματικά αποτελέσματα τη συζήτησή τους και τα συμπεράσματα.

## 5 Σχετικές έρευνες

Τα τελευταία χρόνια, η τεχνητή νοημοσύνη και τα νευρωνικά δίκτυα, όπως τα Convolutional Neural Networks (CNNs), οι Autoencoders και τα LSTM, έχουν ενισχύσει σημαντικά την ακρίβεια στην αυτόματη ανίχνευση καρδιολογικών ανωμαλιών από ECG [5]. Για παράδειγμα, η μελέτη των Hannun et al. [5] κατάφερε cardiologist-level απόδοση σε τα-

---

ξινόμηση αρρυθμιών μέσω deep learning σε 24ωρα ECG.

Σημαντική πρόοδος έχει σημειωθεί και με την απεικόνιση των σημάτων ECG σε δισδιάστατη μορφή. Στη μελέτη των Ahmad et al. [6], χρησιμοποιούνται δύο σχήματα συνδυασμού πληροφοριών (fusion frameworks): το Multimodal Image Fusion (MIF) και το Multimodal Feature Fusion (MFF). Οι παλμοί μετατρέπονται σε εικόνες GAF, RP και MTF, οι οποίες είτε συγχωνεύονται σε τρισδιάστατες εικόνες (MIF) είτε επεξεργάζονται ξεχωριστά με AlexNet και συνενώνονται μέσω Gated Fusion Network (MFF). Η μέθοδος MFF πέτυχε με χρήση SMOTE για εξισορρόπηση των κατηγοριών, accuracy 99.7%.

Παρόμοια, η μελέτη των Zhou et al. [7] εφαρμόζει πολυτροπική απεικόνιση (GAF, RP, MTF) σε συνδυασμό με Residual CNN ενισχυμένο με μηχανισμό Frequency Channel Attention (FCA). Η χρήση Borderline-SMOTE ενισχύει την ταξινόμηση σε πέντε κατηγορίες AAMI, με accuracy 99.6%, ενώ τα μεγαλύτερα σφάλματα παρατηρούνται στις κατηγορίες N, S και F λόγω παρόμοιων χαρακτηριστικών..

Επιπλέον, η μελέτη των Zhang et al. [8] προτείνει μια αυτοεπιβλεπόμενη προσέγγιση (SMC), συνδυάζοντας τροποποιημένο MAE με Gaussian Mixture Models για ομαδοποίηση. Η μετατροπή των παλμών σε GAF εικόνες επιτρέπει υψηλή ακρίβεια χωρίς ανάγκη επισημασμένων δεδομένων (precision 84.25%, recall 65.51%).

Ακόμη και απλοποιημένα μοντέλα έχουν δείξει εντυπωσιακά αποτελέσματα: η εργασία των Izci et al. [9] εφαρμόζει 2D CNN σε grayscale εικόνες 128×128, χωρίς προκατεργασία, επιτυγχάνοντας accuracy 97.42% και αναδεικνύοντας την πρακτικότητα τέτοιων συστημάτων για εφαρμογές πραγματικού χρόνου.

Πιο σύνθετες προσεγγίσεις, όπως αυτή των Zhang et al. [10], αξιοποιούν Relative Position Matrix (RPM) και attention μηχανισμό (GAM) σε μοντέλο ResNet18, καταγράφοντας accuracy 99.30% και recall 98.90%, ενώ η μελέτη των Qin et al. [11] χρησιμοποιεί συνδυασμό WGAN-GP και SE-ResNet1D, ενισχύοντας τις σπάνιες κατηγορίες και καταφέρνοντας F1-score 96.27%.

## 6 Θεωρία

### 6.1 Βασικές έννοιες και προσεγγίσεις

#### 6.1.1 Ηλεκτροκαρδιογράφημα (ECG) και καρδιολογικές διαταραχές

Η καρδιά είναι το κύριο όργανο του καρδιακού συστήματος και αποτελείται κυρίως από μυϊκό ιστό (μυοκάρδιο), αλλά και από άλλους ιστούς όπως αγγειακό, νευρικό και συνδετικό ιστό και είναι υπεύθυνη για την διοχέτευση αίματος στο ανθρώπινο σώμα, κάτι που επιτυγχάνεται μέσω της συστολής-διαστολής της. Η καρδιά είναι όργανο-αντλία και προκειμένου να συσπασθεί απαιτεί μικρό ηλεκτρικό ερέθισμα. Το ερέθισμα αυτό παράγεται από ένα σύνολο ειδικών κυττάρων, τον φλεβόκομβο, ο οποίος καθορίζει και τον ρυθμό της σύσπασης της καρ-

---

διάς. Το ηλεκτρικό σήμα ακολουθεί μία συγκεκριμένη διαδρομή στην καρδιά και όταν φτάσει στις χαμηλότερες κοιλότητες, τότε γίνεται η σύσπαση των κοιλιών και πραγματοποιείται η τροφοδότηση του σώματος με αίμα. Το Ηλεκτρικό Καρδιογράφημα (ΗΚΓ) καταγράφει αυτά τα ηλεκτρικά δυναμικά όπως φτάνουν στην επιφάνεια του σώματος και η μέτρησή του γίνεται στη μονάδα μέτρησης Volt με τον ηλεκτροκαρδιογράφο (εδικό μηχάνημα για παραγωγή ΗΚΓ) [12].

Ένας καρδιακός κύκλος (heartbeat segment) ορίζεται ως η αλληλουχία γεγονότων για έναν καρδιακό παλμό, όπου καρδιακός παλμός είναι μία διαστολή και μία συστολή της καρδιάς. Τα ΗΚΓ έχουν ιδιαίτερη σημασία καθώς δίνουν χρήσιμες πληροφορίες για το προσανατολισμό της καρδιάς, τα μεγέθη των κοιλοτήτων, τον καρδιακό ρυθμό, την παραγωγή και αγωγή διέγερσης αλλά και για καρδιακές διαταραχές, ανεξάρτητα από το αν οφείλονται σε ανατομικές, μηχανικές, μεταβολικές ή κυκλοφορικές ατέλειες.

Κατά την καταγραφή ενός φυσιολογικού ηλεκτροκαρδιογραφήματος (ΗΚΓ), αποτυπώνονται διαδοχικά κύματα τα οποία αντιστοιχούν σε διαφορετικές φάσεις της ηλεκτρικής δραστηριότητας της καρδιάς. Το P κύμα είναι το πρώτο στοιχείο του καρδιακού κύκλου και παράγεται από την ηλεκτρική διέγερση που προκαλείται κατά την εκπόλωση των κόλπων, δηλαδή την ενεργοποίηση των κόλπων πριν από τη συστολή τους. Η εκπόλωση ξεκινά από τον φλεβοκόμβο και εξαπλώνεται σταδιακά στους κόλπους, προκαλώντας τη δημιουργία αυτού του κύματος [13].

Το QRS σύμπλεγμα είναι το ισχυρότερο και πιο σύνθετο μέρος του ΗΚΓ και αντιστοιχεί στην εκπόλωση των κοιλιών, δηλαδή την ηλεκτρική ενεργοποίηση των κατώτερων κοιλοτήτων της καρδιάς πριν από τη συστολή τους. Το QRS σύμπλεγμα προκύπτει από τη γρήγορη μετάδοση του ηλεκτρικού σήματος μέσα στις κοιλίες, καθώς αυτό διαδίδεται μέσω ειδικών οδών αγωγής. Η μετάδοση αυτή ενεργοποιεί ταυτόχρονα μεγάλο μέρος του καρδιακού μυός στις κοιλίες, προκαλώντας τη συστολή τους. Μαζί με το P κύμα, αποτελούν φάσεις εκπόλωσης, καθώς καταγράφονται κατά τη μετάδοση του ηλεκτρικού ερεθίσματος που προκαλεί σύσπαση [14].

Αντίθετα, το T κύμα εκφράζει την επαναπόλωση των κοιλιών, δηλαδή την επαναφορά των κυττάρων τους στην αρχική ηλεκτρική κατάσταση μετά την ενεργοποίηση. Πρόκειται για τη φάση ανάκαμψης του καρδιακού μυός, όπου οι κοιλίες επανέρχονται στη δυναμική ηρεμίας. Το T κύμα συχνά έχει χαμηλότερη τάση, είναι παρατεταμένο χρονικά και μπορεί να επικαλύπτεται ή να συγχέεται με το QRS σύμπλεγμα, ειδικά σε παθολογικές περιπτώσεις [13].

Στη συγκεκριμένη εργασία γίνεται χρήση AAMI για ταξινόμηση των αρρυθμιών σε 5 βασικές κατηγορίες, οι οποίες περιγράφονται παρακάτω.

### 6.1.2 Κατηγορία N – Φυσιολογικοί (Μη-εκτοπικοί) Παλμοί

Η κατηγορία N (Normal) περιλαμβάνει τους φυσιολογικούς καρδιακούς παλμούς, δηλαδή εκείνους που ξεκινούν από τον φλεβόκομβο, και διαδίδονται κανονικά μέσα στην καρδιά [15]. Σε αυτή την κατηγορία ανήκουν παλμοί που έχουν φυσιολογικό ρυθμό και μορφή, καθώς και παλμοί με ήπιες αποκλίσεις, όπως αυτοί που εμφανίζουν καθυστέρηση στη μετάδοση του σήματος στα ηλεκτρικά καλώδια της καρδιάς (δεξί ή αριστερό σκέλος), χωρίς όμως να θεωρούνται παθολογικοί. Συγκεκριμένα, στην κατηγορία N περιλαμβάνονται οι φυσιολογικοί παλμοί με αποκλεισμό αριστερού σκέλους (LBBB), αποκλεισμό δεξιού σκέλους (RBBB), καθώς και οι κοιλικοί παλμοί διαφυγής και οι παλμοί διαφυγής από τον κολποκοιλιακό κόμβο [16].

Ένα φυσιολογικό ΗΚΓ παρουσιάζει τα εξής βασικά χαρακτηριστικά: ένα αρχικό κύμα (P), ακολουθούμενο από ένα στενό QRS και στη συνέχεια ένα T. Αυτή η σειρά αντικατοπτρίζει τη σωστή λειτουργία των ηλεκτρικών σημάτων στην καρδιά, που οδηγεί σε φυσιολογικές συστολές. Επιπλέον, χαρακτηριστικά ενός φυσιολογικού ΗΚΓ είναι οι τακτικές αποστάσεις μεταξύ των R-peaks, που υποδηλώνουν σταθερό καρδιακό ρυθμό παρουσία P κύματος πριν από κάθε QRS σύμπλεγμα, στοιχείο που φανερώνει φυσιολογική αγωγή από τους κόλπους προς τις κοιλίες και συμμετρικότητα και τυπική μορφολογία των QRS και T κυμάτων. Σε περιπτώσεις καρδιακών διαταραχών, το ΗΚΓ μπορεί να παρουσιάζει αλλοιώσεις όπως ακανόνιστες αποστάσεις μεταξύ των R-peaks, απουσία ή παραμόρφωση του P κύματος, ή ασυνήθιστο πλάτος και σχήμα στο QRS και στο T κύμα. Τέτοιες αλλαγές συχνά αποτελούν ενδείξεις αρρυθμιών ή άλλων παθολογικών καταστάσεων της καρδιάς.

### 6.1.3 Κατηγορία S – Υπερκοιλιακά Έκτοπα

Η κατηγορία S (Supraventricular) περιλαμβάνει παλμούς που δεν ξεκινούν από το φυσιολογικό σημείο έναρξης του καρδιακού παλμού (φλεβόκομβο), αλλά από άλλα σημεία της καρδιάς, όπως οι κόλποι ή περιοχές κοντά σε αυτούς (ο κολποκοιλιακός κόμβος) [15].

Οι παλμοί κατηγορίας S εμφανίζονται πριν την αναμενόμενη στιγμή, και μπορεί να έχουν ελαφρώς διαφορετική μορφή στο ΗΚΓ, κυρίως στο πρώτο κύμα (το P κύμα) που αφορά την ηλεκτρική δραστηριότητα των κόλπων. Το υπόλοιπο του παλμού (QRS) συνήθως παραμένει φυσιολογικό, επειδή η ηλεκτρική ώση ακολουθεί κανονικά τη διαδρομή προς τις κοιλίες [17].

Στην κατηγορία S κατατάσσονται οι πρόωρες κοιλικές συστολές (AP), οι αποκλίνουσες πρόωρες κοιλικές (aAP), οι πρόωρες συστολές από τον κολποκοιλιακό κόμβο (NP) και οι υπερκοιλιακές πρόωρες συστολές (SP) [16].

### 6.1.4 Κατηγορία V – Κοιλιακοί Έκτακτοι Παλμοί

Η κατηγορία V (Ventricular) περιλαμβάνει παλμούς που ξεκινούν απευθείας μέσα από τις κοιλίες της καρδιάς, και όχι από το φυσιολογικό σύστημα που κατευθύνει τον ρυθμό [15]. Αυτοί οι παλμοί ονομάζονται κοιλιακές έκτακτες συστολές και εμφανίζονται πρόωρα, δηλαδή

---

νωρίτερα από τον φυσιολογικό παλμό.

Στο ΗΚΓ, οι κοιλιακοί παλμοί φαίνονται διαφορετικοί από τους φυσιολογικούς, κυρίως επειδή έχουν πλατύτερη και πιο ακανόνιστη μορφή στο τμήμα QRS. Αυτό συμβαίνει επειδή η ηλεκτρική ώση δεν ακολουθεί τη συνηθισμένη, γρήγορη διαδρομή μέσα από το αγωγίμο σύστημα, αλλά εξαπλώνεται πιο αργά μέσα από τον μυϊκό ιστό της καρδιάς [18].

Σε αυτήν την κατηγορία ανήκουν οι κοιλιακές συστολές (PVC) και οι κοιλιακοί παλμοί διαφυγής [16].

#### **6.1.5 Κατηγορία F – Συγχωνευμένοι Παλμοί**

Η κατηγορία F (Fusion) περιλαμβάνει παλμούς που προκύπτουν όταν συγχρόνως δύο διαφορετικές ηλεκτρικές εντολές ενεργοποιούν την καρδιά: μία φυσιολογική (από τον φλεβόκομβο) και μία έκτοπη (από τις κοιλίες) [15]. Το αποτέλεσμα είναι ένας παλμός με μικτά χαρακτηριστικά, δηλαδή με μορφή που δεν είναι καθαρά φυσιολογική ούτε καθαρά παθολογική.

Στο ΗΚΓ, αυτοί οι παλμοί έχουν ενδιάμεση μορφή: το QRS σύμπλεγμά τους μοιάζει λίγο με το φυσιολογικό και λίγο με τον κοιλιακό παλμό. Αυτό συμβαίνει γιατί μέρος της καρδιάς ενεργοποιείται από τον φυσιολογικό ρυθμό και μέρος από την έκτοπη εστία [19].

Η αναγνώριση τέτοιων παλμών είναι σημαντική, γιατί μπορεί εύκολα να μπερδευτούν με άλλους τύπους (είτε φυσιολογικούς είτε κοιλιακούς), κάτι που μπορεί να επηρεάσει την ακρίβεια της ταξινόμησης από ένα σύστημα ανάλυσης ΗΚΓ. Επιπλέον, η παρουσία τους συνήθως υποδηλώνει αυξημένη δραστηριότητα έκτοπων παλμών, δηλαδή παλμών που δεν προέρχονται από τον κανονικό καρδιακό ρυθμό.

Η κατηγορία F περιλαμβάνει μόνο τους παλμούς συγχώνευσης φυσιολογικού και κοιλιακού παλμού (fVN) [16].

#### **6.1.6 Κατηγορία Q – Άγνωστοι ή Μη-Ταξινομήσιμοι Παλμοί**

Η κατηγορία Q (Unknown) περιλαμβάνει παλμούς που δεν ταιριάζουν σε καμία από τις άλλες κατηγορίες, είτε επειδή είναι σπάνιοι, είτε επειδή έχουν ασαφή ή ασυνήθιστη μορφή, είτε επειδή το σήμα περιέχει θόρυβο ή παρεμβολές που δυσκολεύουν την αναγνώριση.

Σε αυτή την κατηγορία τοποθετούνται επίσης παλμοί που προέρχονται από βηματοδότη, δηλαδή από τεχνητή ηλεκτρική διέγερση της καρδιάς, παλμοί συγχώνευσης φυσιολογικού και βηματοδότη (fPN), καθώς και μη ταξινομήσιμοι παλμοί [16].

Οι παλμοί αυτής της κατηγορίας δεν αντιπροσωπεύουν κάποια σαφή αρρυθμία αλλά ιδιαίτερες ή τεχνητές περιπτώσεις. Η ύπαρξη της κατηγορίας Q είναι σημαντική, γιατί επιτρέπει στο σύστημα ανάλυσης ΗΚΓ να μην αναγκαστεί να ταξινομήσει λάθος έναν παλμό σε κάποια ακατάλληλη κατηγορία.

### 6.1.7 Gramian Angular Field (GAF)

Γραμμικό Γωνιακό Πεδίο (GAF) είναι η διαδικασία με την οποία χρονοσειρές όπως σήματα καρδιογραφιών, αναπαρίστανται ως εικόνες. Η μετατροπή χρησιμοποιεί γεωμετρικούς και αλγεβρικούς μετασχηματισμούς και η τελική εικόνα διατηρεί τόσο τη χρονική πληροφορία όσο και τις σχέσεις των τιμών της χρονοσειράς.

Προκειμένου να πραγματοποιηθεί η αναπαράσταση χρονοσειρών σε εικόνες, αρχικά πραγματοποιείται κανονικοποίηση των δεδομένων στην κλίμακα  $[0,1]$ , ώστε να μπορούν να μετατραπούν σε γωνίες χρησιμοποιώντας πολικές συντεταγμένες.

Στη συνέχεια, οι τιμές της χρονοσειράς προβάλλονται σε πολικές συντεταγμένες, όπου κάθε σημείο  $x_i$  μετατρέπεται σε γωνία  $\phi_i$ , μέσω της αντίστροφης συνημιτονικής συνάρτησης:

$$\phi_i = \arccos(x_i), \quad x_i \in [-1, 1] \quad (1)$$

Η χρήση γωνιακής αναπαράστασης, δηλαδή η μετατροπή των κανονικοποιημένων τιμών της χρονοσειράς σε γωνίες, επιτρέπει την αξιοποίηση τριγωνομετρικών συναρτήσεων, προσφέροντας μεγαλύτερη ευελιξία στην αποτύπωση πιο σύνθετων και κυκλικών σχέσεων μεταξύ των τιμών της χρονοσειράς, σε σχέση με τις απλές γραμμικές τιμές, οι οποίες δεν είναι εύκολα εκφράσιμες με τη χρήση γραμμικών αναπαραστάσεων. Σε αντίθεση με την παραδοσιακή καρτεσιανή αναπαράσταση, η πολική μορφή δεν διατηρεί μόνο τη σχετική απόσταση μεταξύ σημείων αλλά και τη θέση τους στον χρόνο, ενισχύοντας την ικανότητα του μοντέλου να κατανοεί τοπικές αλλά και μακροπρόθεσμες χρονικές σχέσεις [20].

Κάθε τιμή (pixel) της τελικής εικόνας αντιστοιχεί σε ένα στοιχείο του Gramian Angular πίνακα, ο οποίος προκύπτει από την εφαρμογή τριγωνομετρικής σχέσης μεταξύ των γωνιακά αναπαριστώμενων τιμών της χρονοσειράς. Συγκεκριμένα, η τιμή στη θέση  $(i,j)$  του πίνακα υπολογίζεται βάσει του τύπου:

$$G(i, j) = \cos(\phi_i + \phi_j) \quad (2)$$

$$G(i, j) = \cos(\phi_i - \phi_j) \quad (3)$$

Η επιλογή ανάμεσα σε πρόσθεση ή αφαίρεση των γωνιών καθορίζει και τη μορφή της μεθόδου: όταν χρησιμοποιείται συνάρτηση συνημιτόνου επί της πρόσθεσης γωνιών, η διαδικασία ονομάζεται Gramian Angular Summation Field (GASF), ενώ στην περίπτωση της διαφοράς γωνιών με χρήση της συνάρτησης συνημιτόνου, πρόκειται για το Gramian Angular Difference Field (GADF) [21]. Στο πλαίσιο της παρούσας εργασίας εφαρμόστηκε η προσθετική εκδοχή, δηλαδή το Gramian Angular Summation Field. Η επιλογή του Gramian Angular Summation Field (GASF) έγινε λόγω της συμμετρικής φύσης του και της ικανότητάς του να αποτυπώνει με ακρίβεια τη συσχέτιση μεταξύ χρονικών σημείων, διατηρώντας παράλληλα τη χρονική σειρά της χρονοσειράς. Η διαδικασία αυτή παράγει έναν συμμετρικό πίνακα,

---

στον οποίο κάθε στοιχείο εκφράζει τη γωνιακή συμβολή μεταξύ δύο χρονικών σημείων. Η θέση του pixel της εικόνας αναπαριστά τη χρονική σχέση των δύο σημείων (τις χρονικές στιγμές δηλαδή των σημείων στη χρονοσειρά) και η τιμή του pixel αναπαριστά την συσχέτιση των τιμών των σημείων αυτών. Όπως ήδη αναφέρθηκε, σε αντίθεση με απλές μετρικές απόστασης ή ομοιότητας, η πολική μετατροπή ενσωματώνει τη γεωμετρική πληροφορία των τιμών της χρονοσειράς, προσφέροντας μια πιο εκφραστική και οπτικά αναγνώσιμη αναπαράσταση.

Επιπλέον, παρατηρείται ότι ο GAF διατηρεί τη μορφολογία της χρονοσειράς, καθώς βασίζεται σε συσχετίσεις μεταξύ των τιμών κάθε χρονικού σημείου με όλα τα υπόλοιπα μέσω γωνιακών σχέσεων. Έτσι, ακόμη και στην περίπτωση όπου υπάρχει θόρυβος στο σήμα, η πληροφορία που σχετίζεται με τη συνολική δομή του σήματος διατηρείται και δεν παραμορφώνεται, καθώς ο τριγωνομετρικός μετασχηματισμός αποτυπώνει την εσωτερική σχέση των τιμών του σήματος, χωρίς αλλοίωσή τους.

Η προκύπτουσα εικόνα δηλαδή διατηρεί τη χρονική σειρά των δεδομένων και επιτρέπει την αξιοποίηση χωρικών μοτίβων από μοντέλα βαθιάς μάθησης, όπως τα CNNs. Συνηθώς όσο πιο φωτεινά έντονα είναι τα χρώματα τόσο πιο ισχυρή είναι η θετική συσχέτιση, και όσο πιο σκούρα έντονα τόσο πιο ισχυρή η αρνητική συσχέτιση των τιμών.

## 6.2 Αρχιτεκτονική και λειτουργία των CNNs

### 6.2.1 Αρχιτεκτονική

Τα Συνελικτικά Νευρωνικά Δίκτυα (CNNs) αποτελούν μία ειδική κατηγορία βαθιών νευρωνικών δικτύων, σχεδιασμένη για την αυτόματη εξαγωγή χαρακτηριστικών από εικόνες. Η βασική αρχή ενός CNN είναι η δυνατότητά του να αναγνωρίζει μοτίβα σε τοπικές περιοχές της εισόδου, αξιοποιώντας συνελικτικά φίλτρα (convolutional filters), τα οποία σαρώνουν την είσοδο, εντοπίζουν τοπικά μοτίβα και δημιουργούν χαρακτηριστικούς χάρτες (feature maps).

Γενικά στο δίκτυο, οι έξοδοι προηγούμενων επιπέδων χρησιμοποιούνται ως είσοδοι για τα επόμενα, επιτρέποντας στο CNN να συνθέτει πιο σύνθετα χαρακτηριστικά από απλούστερα. Έτσι, δημιουργείται μια ιεραρχία αναπαραστάσεων, από τα βασικά μοτίβα της εικόνας έως πιο αφηρημένες και εξειδικευμένες δομές που σχετίζονται με τις διαφορετικές κατηγορίες.

Κατά την εκπαίδευση του δικτύου, κάθε φίλτρο ξεκινά με τυχαίες τιμές και προσαρμόζεται μέσω της οπισθοδιάδοσης (backpropagation) με βάση τη συνάρτηση απώλειας. Για κάθε είσοδο, το δίκτυο υπολογίζει μια έξοδο, συγκρίνει την πρόβλεψη με την πραγματική ετικέτα (label) και υπολογίζει το σφάλμα μέσω της συνάρτησης απώλειας (loss function). Το σφάλμα αυτό επιστρέφει πίσω στα επίπεδα του δικτύου και χρησιμοποιείται για την ενημέρωση των βαρών των φίλτρων.

---

Στόχος είναι η προοδευτική εξειδίκευση κάθε φίλτρου σε διαφορετικά μοτίβα που σχετίζονται με συγκεκριμένες εξόδους, αποφεύγοντας την επανάληψη και ενισχύοντας τη διαφοροποίηση. Τα φίλτρα δηλαδή γίνονται συμπληρωματικά και συμβάλλουν συλλογικά στην τελική απόφαση του δικτύου, το οποίο μαθαίνει μια ιεραρχία αναπαραστάσεων που οδηγεί σε υψηλή ακρίβεια κατάταξης.

### 6.2.2 Activation Functions

Η συνέλιξη αποτελεί γραμμική πράξη, καθώς περιλαμβάνει αθροίσματα και γινόμενα μεταξύ σταθερών βαρών και τιμών εισόδου. Αυτό έχει σαν αποτέλεσμα, η εφαρμογή πολλαπλών συνελκτικών επιπέδων να ισοδυναμεί ουσιαστικά με μία σύνθεση γραμμικών συναρτήσεων, η οποία καταλήγει σε μία τελική γραμμική συνάρτηση. Ένα τέτοιο δίκτυο είναι ανεπαρκές για την εκμάθηση πολύπλοκων ή μη γραμμικά διαχωρίσιμων σχέσεων μεταξύ των εισόδων και των επιθυμητών εξόδων. Για τον λόγο αυτό, η εισαγωγή μη γραμμικότητας μετά από κάθε συνελκτικό επίπεδο είναι ουσιώδης. Συνήθως πραγματοποιείται με την εφαρμογή συναρτήσεων ενεργοποίησης όπως η ReLU (Rectified Linear Unit), οι οποίες επιτρέπουν στο δίκτυο να μοντελοποιεί σύνθετες, μη γραμμικές αλληλεπιδράσεις στα δεδομένα. Έτσι, το μοντέλο αποκτά την ικανότητα να αναπαριστά υψηλότερης πολυπλοκότητας κατανομές και να διακρίνει μεταξύ διαφορετικών κατηγοριών με μεγαλύτερη ακρίβεια. Κατά την τελική φάση ενός CNN, η έξοδος από τα συνελκτικά επίπεδα ισιώνεται (flatten), με αποτέλεσμα τη δημιουργία ενός μονοδιάστατου διάνυσματος χαρακτηριστικών. Το διάνυσμα αυτό εισάγεται σε πλήρως συνδεδεμένο (linear) επίπεδο, το οποίο εξάγει ένα score για κάθε κατηγορία εξόδου. Κάθε score προκύπτει από μια ξεχωριστή γραμμική συνάρτηση, δηλαδή έναν συνδυασμό βαρών και bias, που εφαρμόζεται πάνω στο ίδιο διάνυσμα εισόδου. Έτσι, κάθε κατηγορία ορίζεται από ένα υπερ-επίπεδο διαχωρισμού στον χώρο των χαρακτηριστικών, και η πρόβλεψη γίνεται βάσει της κατηγορίας με το μεγαλύτερο score.

$$\text{score}_k = \mathbf{w}_k^\top \cdot \mathbf{x} + b_k \quad (4)$$

Όπου:

- $\mathbf{x}$ : το διάνυσμα εισόδου
- $\mathbf{w}_k$ : το διάνυσμα βαρών για την κατηγορία  $k$
- $b_k$ : το bias για την κατηγορία  $k$
- $\text{score}_k$ : το “logit” ή το output του μοντέλου για την κατηγορία  $k$  πριν το softmax

Αν κάποιο φίλτρο δεν εντοπίσει το μοτίβο στο οποίο έχει εκπαιδευτεί να ανταποκρίνεται, τότε οι ενεργοποιήσεις του είναι αρνητικές και μηδενίζονται από τη ReLU. Αυτό έχει ως απο-

---

τέλεσμα το αντίστοιχο feature map να είναι μηδενικό και να μην επηρεάζει την έξοδο στα επόμενα επίπεδα, αν και τεχνικά εξακολουθεί να προωθείται μέσα στο δίκτυο.

### 6.2.3 Pooling Layers

Τα pooling layers παίζουν κρίσιμο ρόλο στην αρχιτεκτονική ενός CNN, καθώς συμβάλλουν στη σταδιακή μείωση των χωρικών διαστάσεων των feature maps.

Η πιο διαδεδομένη τεχνική είναι το Max Pooling, κατά την οποία μια περιοχή αντικαθίσταται από τη μέγιστη τιμή των στοιχείων της. Με αυτόν τον τρόπο διατηρούνται τα πιο ενεργοποιημένα χαρακτηριστικά της περιοχής, ενώ αγνοούνται τα λιγότερο σημαντικά ή θορυβώδη. Το αποτέλεσμα είναι συμπαγέστερες αναπαραστάσεις που διατηρούν την πληροφορική πυκνότητα και μειώνουν τον κίνδυνο υπερπροσαρμογής. Το pooling εφαρμόζεται τυπικά μετά από κάθε συνελκτικό block, συμβάλλοντας στην ιεραρχική κατανόηση των χαρακτηριστικών.

### 6.2.4 Fully Connected Layers

Τα fully connected layers αποτελούν τα τελευταία στάδια ενός CNN και είναι υπεύθυνα για τη λήψη της τελικής απόφασης. Αφού τα feature maps που έχουν παραχθεί από τα συνελκτικά και pooling επίπεδα υποβληθούν σε διαδικασία flattening (μετατροπή σε μονοδιάστατο διάνυσμα), το αποτέλεσμα περνά μέσα από ένα ή περισσότερα πλήρως συνδεδεμένα επίπεδα. Σε αυτά, κάθε νευρώνας συνδέεται με όλα τα στοιχεία του διανύσματος εισόδου, επιτρέποντας την εκμάθηση σύνθετων, μη γραμμικών σχέσεων μεταξύ των εξαγόμενων χαρακτηριστικών.

Στο τελικό πλήρως συνδεδεμένο επίπεδο, ο αριθμός των εξόδων ισούται με τον αριθμό των κατηγοριών του προβλήματος ταξινόμησης. Συχνά, το επίπεδο αυτό συνοδεύεται από τη softmax, μια συνάρτηση που μετατρέπει τις εξόδους σε πιθανότητες, επιτρέποντας την ερμηνεία της εξόδου ως βαθμό εμπιστοσύνης για κάθε κατηγορία.

### 6.2.5 Batch Normalization

Η τεχνική Batch Normalization εφαρμόζεται για τη βελτίωση της σταθερότητας και της αποδοτικότητας της εκπαίδευσης των νευρωνικών δικτύων κανονικοποιώντας τις εξόδους ενός layer έτσι ώστε να έχουν μέση τιμή κοντά στο 0 και τυπική απόκλιση κοντά στο 1.

Αυτή η διαδικασία επιλύει το πρόβλημα συνεχής αλλαγής της κατανομής των εισόδων στα ενδιάμεσα επίπεδα του δικτύου κατά την εκπαίδευση. Επιπλέον, επιτρέπει τη χρήση μεγαλύτερων learning rates και μειώνει την ανάγκη για προσεκτική αρχικοποίηση των βαρών. Τέλος, η κανονικοποίηση αυτή λειτουργεί και ως μορφή regularization, μειώνοντας το overfitting και ενισχύοντας τη γενίκευση του μοντέλου σε δεδομένα που δεν έχει ξαναδεί.

---

## 6.3 Εκπαίδευση Δικτύου και Συνάρτηση Απώλειας

Κατά την εκπαίδευση του νευρωνικού δικτύου, είναι απαραίτητος ο ορισμός μιας συνάρτησης απώλειας (loss function), η οποία μετρά την απόκλιση μεταξύ της πρόβλεψης του μοντέλου και της πραγματικής ετικέτας. Στην παρούσα εργασία, καθώς πρόκειται για πρόβλημα πολυκατηγορικής ταξινόμησης, χρησιμοποιήθηκε η συνάρτηση Cross Entropy Loss. Η συνάρτηση αυτή συγκρίνει το διάνυσμα εξόδου του μοντέλου, το οποίο περιέχει πιθανότητες για κάθε κατηγορία, με την πραγματική κατηγορία, και παράγει έναν βαθμό σφάλματος. Όσο πιο απομακρυσμένη είναι η πρόβλεψη από τη σωστή κατηγορία, τόσο μεγαλύτερη είναι η τιμή της απώλειας. Αντίθετα, όταν το μοντέλο δώσει υψηλή πιθανότητα στην ορθή κατηγορία, η απώλεια είναι μικρή. Η προσαρμογή των παραμέτρων του μοντέλου πραγματοποιείται μέσω του αλγορίθμου βελτιστοποίησης Adam (Adaptive Moment Estimation).

## 6.4 Βελτιστοποίηση και υπερπαραμέτροι

Σε ένα μοντέλο, οι υπερπαραμέτροι ορίζονται πριν ξεκινήσει η εκπαίδευσή του και ασκούν καθοριστική επίδραση στην απόδοση του μοντέλου. Η διαδικασία βελτιστοποίησης των υπερπαραμέτρων στοχεύει στην εύρεση των τιμών που δίνουν την καλύτερη δυνατή απόδοση, δωσμένου ενός κριτηρίου αξιολόγησης.

### 6.4.1 Grid Search

Το Grid Search αποτελεί μέθοδο εξαντλητικής αναζήτησης (exhaustive search). Δοκιμάζει δηλαδή όλους τους δυνατούς συνδυασμούς τιμών από ένα δοσμένο χώρο αναζήτησης στοχεύοντας στον εντοπισμό του συνδυασμού που εξυπηρετεί αποδοτικότερα τη λύση ενός προβλήματος. Εξετάζει δηλαδή εξαντλητικά κάθε πιθανή επιλογή μέχρι να βρει τη βέλτιστη. Στο πλαίσιο της παρούσας εργασίας, η μέθοδος αυτή εφαρμόστηκε για την εύρεση βέλτιστου κατωφλίου για δύο κλάσεις. Βασίζεται στον ορισμό ενός διακριτού και πεπερασμένου χώρου τιμών για τις υπερπαραμέτρους του μοντέλου και στη συνέχεια, το μοντέλο εκπαιδεύεται και αξιολογείται για κάθε πιθανό συνδυασμό των τιμών αυτών. Κατά την αξιολόγηση, ο κάθε συνδυασμός συγκρίνεται βάσει προκαθορισμένης μετρικής απόδοσης, και διατηρείται εκείνος που οδηγεί στο καλύτερο αποτέλεσμα ως προς το επιθυμητό κριτήριο.

### 6.4.2 Optuna

Η επιλογή των κατάλληλων υπερπαραμέτρων αποτελεί κρίσιμο στάδιο στην ανάπτυξη ενός νευρωνικού μοντέλου, ιδιαίτερα όταν αντιμετωπίζονται προβλήματα με ανισόρροπες κλάσεις. Για τον σκοπό αυτό, στην παρούσα εργασία αξιοποιήθηκε το Optuna, για τη βελτιστοποίηση υπερπαραμέτρων (hyperparameter optimization) μέσω αυτοματοποιημένης διαδικασίας. Στο Optuna δεν χρειάζεται να προκαθορισθεί εξ αρχής ολόκληρος ο χώρος αναζήτη-

σης των υπερπαραμέτρων. Αντίθετα, ορίζεται ένα εύρος τιμών για κάθε παράμετρο, από το οποίο επιλέγονται τιμές δυναμικά κατά την εκτέλεση των δοκιμών. Αυτή η προσέγγιση προσφέρει μεγαλύτερη ευελιξία και επιτρέπει πιο σύνθετη και προσαρμοστική λογική κατά τη διαδικασία βελτιστοποίησης. Η βασική δομή λειτουργίας του Optuna περιλαμβάνει τον ορισμό μιας συνάρτησης αντικειμενικού σκοπού (objective function), η οποία κατά την εκτέλεση κάθε trial λαμβάνει συγκεκριμένες τιμές υπερπαραμέτρων που της προτείνει το Optuna, και επιστρέφει ως έξοδο μια τιμή μετρικής απόδοσης. Κατά την εκτέλεση, το Optuna εκτελεί διαδοχικά trials, καθένα από τα οποία αντιστοιχεί σε μια ξεχωριστή δοκιμή τιμών και αξιολογείται βάσει της επιστρεφόμενης τιμής της μετρικής. Για τη διαδικασία δειγματοληψίας και επιλογής επόμενων υποψηφίων τιμών, το Optuna χρησιμοποιεί τον αλγόριθμο Tree-structured Parzen Estimator (TPE), που είναι μια μορφή Bayesian βελτιστοποίησης. Το Optuna δεν πραγματοποιεί εξαντλητική ή τυχαία αναζήτηση, αλλά επικεντρώνεται σε περιοχές του χώρου τιμών που έχουν παρουσιάσει υποσχόμενα αποτελέσματα. Η ενσωμάτωση του Optuna στο pipeline της παρούσας εργασίας επέτρεψε τον αυτοματοποιημένο και αποδοτικό προσδιορισμό των κατάλληλων βαρών για τις ανισόρροπες κλάσεις, οδηγώντας σε βελτίωση της γενικής απόδοσης του ταξινομητή και ειδικότερα της ανάκλησης στις δύσκολες κατηγορίες.

### 6.4.3 Threshold Tuning

Σε προβλήματα ταξινόμησης με νευρωνικά δίκτυα, το κατώφλι (threshold) είναι μια οριακή τιμή που εφαρμόζεται στο διανυσματικό αποτέλεσμα της softmax, δηλαδή στις προβλεπόμενες πιθανότητες του μοντέλου για κάθε κατηγορία. Συχνά η μέγιστη αυτή πιθανότητα πρέπει να υπερβαίνει ένα προκαθορισμένο κατώφλι [22] για να τοποθετηθεί η είσοδος στην κατηγορία.

Η τεχνική αυτή εξασφαλίζει ότι αν η μέγιστη πιθανότητα δεν ξεπερνά το όριο, το μοντέλο μειώνει τις ψευδώς θετικές προβλέψεις, αυξάνοντας έτσι την αξιοπιστία του, κάτι που βοηθάει σε datasets με έντονη ανισορροπία, όπως αυτά με καρδιογραφήματα.

Αυτή η προσέγγιση εφαρμόστηκε με επιτυχία σε διάφορες μελέτες, όπως από τους Nejedly et al., όπου για κάθε μία από τις 26 κατηγορίες ECG ορίστηκε διαφορετικό threshold μέσω γενετικού αλγορίθμου, επιτυγχάνοντας την πρώτη θέση στο PhysioNet Challenge 2021 [23]. Αντίστοιχα, οι Pesl et al. εφάρμοσαν Bayesian optimization για να βρουν thresholds που μεγιστοποιούν τη συνολική απόδοση του συστήματος [24].

Μαθηματικά, η απόφαση μοντελοποίησης μπορεί να διατυπωθεί ως:

$$\hat{y} = \begin{cases} \arg \max_k p_k, & \text{αν } \max_k p_k \geq \theta_k \\ y_{\text{fallback}}, & \text{διαφορετικά} \end{cases}$$

όπου  $p_k$  είναι η πιθανότητα για την κατηγορία  $k$ ,  $\theta_k$  το κατώφλι της κατηγορίας και

---

$y_{\text{fallback}}$  η fallback πρόβλεψη (π.χ. “Normal”).

## 6.5 Κανονικοποίηση και προεπεξεργασία

### 6.5.1 Min-Max Normalization

Η κλιμάκωση είναι η διαδικασία μετασχηματισμού των δεδομένων έτσι ώστε το φάσμα τιμών τους να βρίσκεται σε συγκεκριμένο εύρος ή να ακολουθεί συγκεκριμένες στατιστικές ιδιότητες. Αυτή η μέθοδος είναι ιδιαίτερα σημαντική σε χρονοσειρές και σε άλλα είδη δεδομένων καθώς επιτυγχάνει την εξασφάλιση παρόμοιας κλίμακας για όλα τα χαρακτηριστικά του συνόλου δεδομένων. Στη συγκεκριμένη εργασία έγινε χρήση Min-Max Scaling, όπου τα δεδομένα κανονικοποιούνται σε ένα συγκεκριμένο εύρος, όπως  $[0, 1]$  ή  $[-1, 1]$ . Η κανονικοποίηση με min-max scaling δίνεται από τον τύπο:

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

### 6.5.2 Resampling

Το resampling είναι μία μέθοδος εξομοίωσης του μήκους των κατακερματισμένων τμημάτων (segments), έτσι ώστε να μπορούν να χρησιμοποιηθούν ως είσοδος σε αλγορίθμους μηχανικής μάθησης. Ειδικά στα ΗΚΓ, όταν κάθε τμήμα ορίζεται μεταξύ δύο διαδοχικών R-peaks (RR interval), τα μήκη διαφέρουν ανάλογα με τον καρδιακό ρυθμό. Σε περιπτώσεις όπου ένα τμήμα είναι μεγαλύτερο από το επιθυμητό μήκος, εφαρμόζεται η τεχνική της επαναδειγματοληψίας (resampling). Η επαναδειγματοληψία είναι η διαδικασία κατά την οποία ένα σήμα με  $N$  δείγματα ανακατασκευάζεται έτσι ώστε να έχει  $M$  δείγματα, χωρίς να αλλάζει σημαντικά η μορφολογία του. Σε αντίθεση με την απλή αποκοπή (cropping) ή τη συμπλήρωση με μηδενικά (zero-padding), η επαναδειγματοληψία προσπαθεί να συμπίσει ή να επεκτείνει το σήμα χρονικά μέσω παρεμβολής στο πεδίο της συχνότητας. Στην παρούσα εργασία έχει χρησιμοποιηθεί για συμπίεση των τμημάτων. Συγκεκριμένα, κατά την εφαρμογή της επαναδειγματοληψίας με σκοπό τη μείωση του μήκους, πραγματοποιείται συμπίεση του χρονικού άξονα του σήματος, ώστε να παραχθεί ένα νέο σήμα με ακριβώς  $M$  δείγματα. Η διαδικασία αυτή βασίζεται σε φασματική παρεμβολή και περιλαμβάνει αρχικά τον υπολογισμό του Μετασχηματισμού Fourier (FFT) του αρχικού σήματος, μέσω του οποίου το σήμα μεταφέρεται στο πεδίο της συχνότητας

Το σήμα  $x[n]$ , το οποίο είναι μία διακριτή ακολουθία με  $N$  δείγματα, μετασχηματίζεται μέσω του μετασχηματισμού Fourier, δίνοντας ένα νέο σήμα  $X[k]$ , με βάση τον παρακάτω τύπο:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi kn/N} \quad (5)$$

Το νέο σήμα  $X[k]$  δείχνει ποιες συχνότητες περιέχει το αρχικό σήμα και πόσο ισχυρές είναι (φασματική ανάλυση). Ακολουθεί ο αντίστροφος μετασχηματισμός Fourier (iFFT), που επαναφέρει το προσαρμοσμένο φάσμα στο πεδίο του χρόνου, δημιουργώντας ένα νέο σήμα με χαμηλότερη πυκνότητα δειγμάτων, αλλά με μορφολογικά παρόμοια χαρακτηριστικά. Η αντίστροφη μορφή του μετασχηματισμού Fourier (iFFT) υπολογίζει την αρχική διακριτή ακολουθία  $x[m]$  από το φασματικό σήμα  $X[k]$ , με βάση τον τύπο:

$$\tilde{x}[m] = \frac{1}{M} \sum_{k=0}^{M-1} X[k] \cdot e^{j2\pi km/M} \quad (6)$$

Ο τύπος αυτός επανασυνθέτει το αρχικό σήμα στο χωρικό (ή χρονικό) πεδίο από τις φασματικές του συνιστώσες. Το αποτέλεσμα είναι ένα νέο σήμα  $\tilde{x}[m]$  με  $M$  δείγματα, το οποίο αποτελεί μία συμπιεσμένη χρονική αναπαράσταση του αρχικού σήματος, χωρίς να αλλάζει τη διάρκεια του σήματος σε φυσικές μονάδες (π.χ. ms).

### 6.5.3 Zero-padding

Η τεχνική του zero padding αποτελεί μία τεχνική προεπεξεργασίας σημάτων, στην οποία προστίθενται τεχνητά μηδενικά στην αρχή ή/και στο τέλος ενός σήματος με σκοπό την επίτευξη ομοιόμορφου μήκους δεδομένων εισόδου για επεξεργασία ή εκπαίδευση σε μοντέλα μηχανικής μάθησης.

Έστω ένα αρχικό διακριτό σήμα  $x[n]$  μήκους  $L$ . Η διαδικασία του zero padding συνίσταται στην προσθήκη μηδενικών τιμών στο τέλος ή και στην αρχή του σήματος, ώστε να αποκτήσει νέο, προκαθορισμένο μήκος  $N$ , με  $N > L$ . Το νέο, επεκταμένο σήμα  $x_{\text{padded}}[n]$  ορίζεται ως:

$$x_{\text{padded}}[n] = \begin{cases} x[n], & 0 \leq n < L \\ 0, & L \leq n < N \end{cases}$$

ή, ισοδύναμα, σε διανυσματική μορφή:

$$x_{\text{padded}} = [0, 0, \dots, x_1, x_2, \dots, x_L, 0, \dots, 0]$$

όπου  $x_1, x_2, \dots, x_L$  είναι τα δείγματα του αρχικού σήματος. Ο συνολικός αριθμός των μηδενικών που προστίθενται είναι  $N - L$ .

Η κατανομή των μηδενικών μπορεί να είναι:

- Ασύμμετρη (π.χ. όλα τα μηδενικά στο τέλος),

- ή συμμετρική (ισομερώς στην αρχή και στο τέλος), ανάλογα με την εφαρμογή.

Στην περίπτωση ΗΚΓ χρησιμοποιείται η συμμετρική καθώς διαφορετικά θα υπήρχε κίνδυνος για χρονική μετατόπιση των χτύπων. Επιπλέον, προσθήκη μηδενικών τιμών έχει το πλεονέκτημα ότι διατηρεί αναλλοίωτο το αρχικό σήμα, χωρίς περικοπές, αποφεύγοντας έτσι απώλειες πληροφορίας.

Στην ταξινόμηση καρδιακού ρυθμού από σήματα ECG, κάθε παλμός (heartbeat) που εξάγεται με βάση τα R-peaks έχει διαφορετική διάρκεια, καθώς το διάστημα μεταξύ διαδοχικών καρδιακών παλμών (RR-διάστημα) εξαρτάται από τη συχνότητα και τη μορφολογία του καρδιακού κύκλου.

Αυτή η μεταβλητότητα στο μήκος των παλμών δημιουργεί πρόβλημα σε μοντέλα που απαιτούν ομοιόμορφο μήκος εισόδου. Τεχνικές όπως το zero-padding προσφέρει μια απλή και αποδοτική λύση σε αυτό το ζήτημα, καθώς επιτρέπει την ομογενοποίηση του μήκους των σημάτων, καθώς το αρχικό τμήμα συμπληρώνεται με κατάλληλο αριθμό δειγμάτων για να ταιριάζει με το μέγιστο μήκος περιόδου, συμπεριλαμβανόμενης της μηδενικής συμπλήρωσης [25].

## 6.6 Μετρικές Αξιολόγησης

### 6.6.1 Accuracy

Η ακρίβεια (accuracy) είναι ένα από τα βασικά μετρικά μεγέθη αξιολόγησης σε προβλήματα ταξινόμησης. Εκφράζει το ποσοστό των σωστών προβλέψεων του μοντέλου σε σχέση με το σύνολο των δειγμάτων. Δείχνει δηλαδή, πόσο συχνά το μοντέλο προβλέπει σωστά, χωρίς να λαμβάνει υπόψη το είδος των λαθών. Ένα υψηλό accuracy σημαίνει ότι το μοντέλο κάνει λίγα λάθη συνολικά και προβλέπει σωστά τα περισσότερα δείγματα, ενώ ένα χαμηλό σημαίνει ότι το μοντέλο αποτυγχάνει να ταξινομήσει σωστά μεγάλο μέρος των δεδομένων. Ωστόσο, η ακρίβεια (accuracy) δεν αποτελεί πάντα αξιόπιστο κριτήριο αξιολόγησης σε προβλήματα με ανισορροπία κατηγοριών (imbalanced datasets), καθώς μπορεί να δίνει την ψευδαίσθηση υψηλής απόδοσης. Το μοντέλο ενδέχεται να προβλέπει σωστά τις κατηγορίες που εμφανίζονται συχνότερα, ενώ αποτυγχάνει συστηματικά στις λιγότερο αντιπροσωπευμένες. Παρ' όλα αυτά, η συνολική ακρίβεια παραμένει υψηλή, επειδή κυριαρχείται από τις προβλέψεις της πλειοψηφικής κατηγορίας. Ο μαθηματικός τύπος δίνεται από τη σχέση:

**Μαθηματικός τύπος:**

$$\text{Accuracy} = \frac{\text{Αριθμός σωστών προβλέψεων}}{\text{Σύνολο προβλέψεων}} = \frac{TP + TN}{TP + TN + FP + FN}$$

Όπου:

- TP (True Positives): Τα δείγματα που είναι θετικά και προβλέφθηκαν σωστά ως θετικά.
- TN (True Negatives): Τα δείγματα που είναι αρνητικά και προβλέφθηκαν σωστά ως αρνητικά.
- FP (False Positives): Τα δείγματα που είναι αρνητικά αλλά προβλέφθηκαν λανθασμένα ως θετικά.
- FN (False Negatives): Τα δείγματα που είναι θετικά αλλά προβλέφθηκαν λανθασμένα ως αρνητικά.

### 6.6.2 Precision

Η εκλεκτικότητα (precision) είναι μια μετρική αξιολόγησης που εστιάζει στην εκτίμηση της ποιότητας των θετικών προβλέψεων ενός μοντέλου ταξινόμησης. Αντιπροσωπεύει το ποσοστό των σωστά προβλεπόμενων θετικών κλάσεων σε σχέση με όλες τις προβλέψεις που έγιναν ως θετικές και δίνεται από τον μαθηματικό τύπο:

**Μαθηματικός τύπος:**

$$\text{Precision} = \frac{TP}{TP + FP}$$

Όπου:

- TP (True Positives): Τα δείγματα που είναι θετικά και προβλέφθηκαν σωστά ως θετικά.
- FP (False Positives): Τα δείγματα που είναι αρνητικά αλλά προβλέφθηκαν λανθασμένα ως θετικά.

Το precision είναι χρήσιμο στις περιπτώσεις όπου είναι σημαντικό να μην υπάρχουν ψευδώς θετικές προβλέψεις. Υψηλό precision σημαίνει ότι το μοντέλο κάνει λίγα λάθη τύπου FP, δηλαδή δεν «υπερδιαγιγνώσκει» την κατηγορία ενώ χαμηλό υποδηλώνει ότι μεγάλο ποσοστό των θετικών προβλέψεων του μοντέλου είναι στην πραγματικότητα λανθασμένο, άρα το μοντέλο τείνει να κάνει πολλές προβλέψεις υπέρ της κατηγορίας ακόμη και όταν δεν πρέπει.

### 6.6.3 Recall

Η ανάκληση (recall), γνωστή και ως ευαισθησία (sensitivity) ή true positive rate, μετρά την ικανότητα του μοντέλου να εντοπίζει σωστά όλα τα θετικά δείγματα. Αντιπροσωπεύει το ποσοστό των πραγματικά θετικών περιπτώσεων που αναγνωρίστηκαν σωστά από το μοντέλο και δίνεται από τον τύπο:

---

Μαθηματικός τύπος:

$$\text{Recall} = \frac{TP}{TP + FN}$$

Όπου:

- TP (True Positives): Τα δείγματα που είναι θετικά και προβλέφθηκαν σωστά ως θετικά.
- FN (False Negatives): Τα δείγματα που είναι θετικά αλλά προβλέφθηκαν λανθασμένα ως αρνητικά.

#### 6.6.4 F1-score

Το F1-score εκπροσωπεί την ισορροπία μεταξύ recall και precision. Δίνεται από τον τύπο:

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Οι τιμές του κυμαίνονται από 0 έως 1. Εάν το F1 πάρει την τιμή 1, σημαίνει άριστη επίδοση (τέλεια precision και recall), ενώ αν πάρει τιμή κοντά στο 0 υποδηλώνει ότι είτε η precision, είτε η recall (ή και οι δύο) είναι χαμηλές. Σε προβλήματα με ανισορροπία κατηγοριών (imbalanced datasets), η μεγιστοποίηση μόνο της μίας από τις δύο μετρικές ενδέχεται να οδηγήσει σε παραπλανητικά συμπεράσματα. Το F1-score συνδυάζει και τις δύο σε έναν ενιαίο δείκτη απόδοσης, ελαχιστοποιώντας την επίδραση των ακραίων τιμών και παρέχοντας μια πιο αξιόπιστη εκτίμηση της συνολικής συμπεριφοράς του μοντέλου.

## 7 Μεθοδολογία

Η ροή εργασίας που ακολουθήθηκε και τα επιμέρους βήματά της περιγράφονται παρακάτω με λεπτομέρεια. Αρχικά, πραγματοποιείται εντοπισμός των κορυφών R-peaks στα σήματα ηλεκτροκαρδιογραφήματος και δημιουργούνται τα αντίστοιχα καρδιακά τμήματα (heartbeat segments) με τα αντίστοιχα διαγνωστικά τους σύμβολα. Στη συνέχεια, ακολουθεί επεξεργασία του μήκους κάθε segment, ώστε να διασφαλιστεί ότι όλα τα αποσπάσματα έχουν σταθερό μέγεθος 256 δειγμάτων, είτε μέσω μηδενικής συμπλήρωσης (zero padding) είτε με επαναδειγματοληψία (resampling). Αφού εξομαλυνθεί η χρονική διάρκεια, πραγματοποιείται κανονικοποίηση του κάθε χτύπου και ακολουθεί η μετατροπή των καρδιακών παλμών σε εικόνες με χρήση του μετασχηματισμού GAF (Gramian Angular Field). Οι εικόνες αυτές έχουν τελικό μέγεθος εικόνας 64×64 pixels.

Το σύνολο των παλμών που εξήχθησαν οργανώθηκε σε τρία υποσύνολα: εκπαίδευσης, επικύρωσης και δοκιμής (train, validation, test). Ο διαχωρισμός πραγματοποιήθηκε σε δύο στάδια: αρχικά, το αρχικό σύνολο χωρίστηκε σε εκπαίδευση και δοκιμή, και στη συνέχεια το υποσύνολο εκπαίδευσης διασπάστηκε εκ νέου για τη δημιουργία του συνόλου επικύρω-

---

σης. Για την τμηματική τροφοδότηση του μοντέλου, αξιοποιήθηκε ο μηχανισμός του PyTorch DataLoader, σε συνδυασμό με μια προσαρμοσμένη κλάση τύπου Dataset. Η κλάση αυτή αναλαμβάνει τη μετατροπή των δεδομένων στην κατάλληλη μορφή για να δοθούν ως είσοδος στο νευρωνικό δίκτυο. Συγκεκριμένα, διαμορφώνει κάθε δείγμα ως ζεύγος εισόδου-στόχου που περιλαμβάνει μία εικόνα ενός καναλιού, σταθερών διαστάσεων, η οποία έχει παραχθεί μέσω μετασχηματισμού Gramian Angular Field (GAF) από το κανάλι MLII, καθώς και την αντίστοιχη ετικέτα κατηγορίας σε ακέραιη μορφή. Το μέγεθος παρτίδας ορίστηκε σε 32, προκειμένου να επιτευχθεί ισορροπία μεταξύ κατανάλωσης μνήμης και ταχύτητας υπολογισμών. Το δίκτυο εκπαιδεύτηκε σε GPU για 30 εποχές με χρήση του αλγορίθμου Adam και τιμή learning rate που προέκυψε από διαδικασία βελτιστοποίησης. Για τη ρύθμιση υπερπαραμέτρων χρησιμοποιήθηκαν δύο τεχνικές: αναζήτηση μέσω πλέγματος (grid search) και αυτόματη βελτιστοποίηση με τη χρήση της βιβλιοθήκης Optuna, με στόχο την ενίσχυση της ακρίβειας του μοντέλου σε λιγότερο εκπροσωπούμενες περιπτώσεις.

## 7.1 Δεδομένα – MIT-BIH Arrhythmia Database

Το dataset που χρησιμοποιήθηκε ήταν το MIT-BIH Arrhythmia dataset [26] της βάσης Physionet, το οποίο είναι ένα ευρέως διαδεδομένο και καθιερωμένο σύνολο δεδομένων στον ερευνητικό χώρο. Η βάση δεδομένων περιέχει αρχεία καταγραφής καρδιογραφήματων διάρκειας 30 λεπτών σε συχνότητα καταγραφής 360Hz, από συνολικά 48 ασθενείς. Περίπου το 60% των καταγραφών που υπάρχουν στο dataset προέρχονται από νοσηλευόμενους ασθενείς, δηλαδή από ανθρώπους που βρίσκονταν ήδη υπό ιατρική παρακολούθηση και άρα έχουν αυξημένη την πιθανότητα να καταγραφούν στα καρδιογραφήματά τους παθολογικά ευρήματα. Συγκεκριμένα, η βάση δεδομένων περιλαμβάνει 23 καταγραφές (με αριθμούς από 100 έως 124, με μη συνεχόμενη αρίθμηση), οι οποίες επιλέχθηκαν τυχαία από ένα γενικό σύνολο, ώστε να είναι αντιπροσωπευτικές των μορφών κυματομορφών και των παρασίτων (artifacts) που ενδέχεται να συναντήσει ένας ανιχνευτής αρρυθμιών στην καθημερινή κλινική πράξη. Επιπλέον, περιλαμβάνονται 25 καταγραφές που επιλέχθηκαν στοχευμένα, προκειμένου να περιλαμβάνουν σπάνια αλλά κλινικά σημαντικά φαινόμενα. Η σκόπιμη αυτή επιλογή γίνεται επειδή οι συσκευές Holter, που χρησιμοποιούνται συνήθως, καταγράφουν τον καρδιακό ρυθμό για μεγάλα χρονικά διαστήματα. Ωστόσο, στο συγκεκριμένο dataset οι καταγραφές περιορίζονται σε διάρκειες περίπου 30 λεπτών. Για να μειωθεί επομένως η πιθανότητα απουσίας σπάνιων αλλά κλινικά σημαντικών φαινομένων, επιλέγονται στοχευμένα 25 εγγραφές που τα περιλαμβάνουν. Τα αρχεία αυτά εστιάζουν στις εξής σύνθετες παθολογικές καταστάσεις: κοιλιακές αρρυθμίες (ventricular), κολποκοιλιακές (junctional), υπερκοιλιακές (supraventricular) και διαταραχές αγωγιμότητας (conduction abnormalities). Στόχος της ένταξής τους είναι να δοκιμαστεί η ανθεκτικότητα και η αξιοπιστία των ανιχνευτών αρρυθμιών σε απαιτητικές ή προβληματικές περιπτώσεις. Συνολικά στο dataset υπάρχουν καρδιογραφήματα από 25 άν-

---

δρες ηλικίας 32 έως 89 ετών και από 22 γυναίκες ηλικίας 23 έως 89 ετών (Τα αρχεία 201 και 202 αντιστοιχούν στον ίδιο άνδρα συμμετέχοντα.).

Στην πλειονότητα των αρχείων για το πάνω σήμα (upper signal) χρησιμοποιήθηκε τροποποιημένη απαγωγή ΙΙ (MLII) και για το κάτω σήμα (lower signal) τροποποιημένη απαγωγή V1. Η οπτική της απαγωγής MLII είναι σχεδόν παράλληλη με τον ηλεκτρικό άξονα της καρδιάς, ενώ τη απαγωγή V1 έχει προσανατολισμό σχεδόν κάθετο προς τον άξονα αυτόν. Αυτό έχει ως αποτέλεσμα στο κάτω κανάλι να είναι πιο ευδιάκριτοι οι παθολογικοί παλμοί, ενώ στο πάνω κανάλι είναι ευκολότερος ο εντοπισμός των φυσιολογικών καρδιακών συμπλεγμάτων.

Ακολουθώντας τον τρόπο των περισσότερων δημοσιευμένων ερευνών, χρησιμοποιείται μόνο το πρώτο από τα δύο κανάλια, δηλαδή το MLII. Για αυτόν τον λόγο οι ασθενείς 102, 104, 114 δεν χρησιμοποιούνται καθώς δεν χρησιμοποιούν ως πρώτο κανάλι το MLII.

## 7.2 Προεπεξεργασία δεδομένων

### 7.2.1 Χρήση AAMI

Όπως αναφέρθηκε παραπάνω, η βάση δεδομένων MIT-BIH Arrhythmia Database συνοδεύει τα σήματα ηλεκτροκαρδιογραφήματος με μια εκτενή λίστα από σύμβολα annotations, όπου κάθε γράμμα (όπως 'A', 'V', 'L', 'R', 'j') αντιστοιχεί σε έναν συγκεκριμένο τύπο καρδιακού παλμού. Τα σύμβολα αυτά προκύπτουν από την κλινική αξιολόγηση των καταγραφών και έχουν σχεδιαστεί έτσι ώστε να περιγράφουν λεπτομερώς την ηλεκτροφυσιολογική προέλευση κάθε παλμού, όπως για παράδειγμα αν προέρχεται από τους κόλπους, τις κοιλίες ή την περιοχή του κολποκοιλιακού κόμβου. Αυτή η υψηλή διαγνωστική ακρίβεια είναι ιδιαίτερα χρήσιμη σε κλινικό περιβάλλον, ωστόσο οδηγεί στην ύπαρξη υπερβολικά πολλών κλάσεων για ένα πρόβλημα ταξινόμησης, ειδικά όταν πολλές από αυτές τις κλάσεις εμφανίζονται πολύ σπάνια καθιστώντας έτσι το πρόβλημα της ταξινόμησης υψηλά ανισόρροπο (imbalanced). Επιπλέον, αρκετές από αυτές τις κατηγορίες δεν παρουσιάζουν σημαντικές μορφολογικές διαφορές στο σήμα του ΗΚΓ, γεγονός που καθιστά δύσκολη τη διάκρισή τους από ένα σύστημα αυτόματης ανάλυσης, ακόμη και με τη χρήση νευρωνικών δικτύων. Για παράδειγμα, οι παλμοί με σύμβολα 'N' (Normal beat), 'L' (Left bundle branch block beat), 'R' (Right bundle branch block beat), 'e' (Atrial escape beat) και 'j' (Nodal escape beat), παρότι προέρχονται από από διαφορετικά σημεία της καρδιάς, παρουσιάζουν παρόμοια μορφολογία στο ECG, καθώς το σύμπλοκο QRS έχει φυσιολογικά ή σχεδόν φυσιολογικά χαρακτηριστικά. Για τον λόγο αυτό, όλοι οι παραπάνω τύποι ομαδοποιούνται στην κατηγορία N, ώστε να θεωρούνται ως «φυσιολογικοί» για σκοπούς αυτόματης ταξινόμησης. Για την αντιμετώπιση αυτών των προβλημάτων, γίνεται χρήση του προτύπου AAMI [27], το οποίο προτείνει τη συγχώνευση των λεπτομερών annotations σε πέντε ευρύτερες κατηγορίες αρρυθμιών: N (Normal), S (Supraventricular), V (Ventricular), F (Fusion), Q (Unknown). Η ομαδοποίηση

---

αυτή βασίζεται τόσο στη μορφολογία των παλμών όσο και στη διαγνωστική τους σημασία. Στη συγκεκριμένη εργασία για την ενοποίηση των διαφορετικών τύπων καρδιακών παλμών σε πέντε βασικές κατηγορίες σύμφωνα με το πρότυπο AAMI, χρησιμοποιείται ένας πίνακας αντιστοίχισης (annotation\_map). Το λεξικό αυτό αντιστοιχίζει ή χαρτογραφεί τα σύμβολα annotations του MIT-BIH, τα οποία περιγράφουν τον τύπο κάθε παλμού (όπως 'N', 'L', 'V', 'A' κ.ά.), στις αντίστοιχες υποκατηγορίες του AAMI:

- Τα σύμβολα 'N', 'L', 'R', 'e', 'j' αντιστοιχίζονται στην κατηγορία N (Normal).
- Τα σύμβολα 'A', 'a', 'J', 'S' στην κατηγορία S (Supraventricular).
- Τα σύμβολα 'V', 'E' στην κατηγορία V (Ventricular).
- Το σύμβολο 'F' στην κατηγορία F (Fusion).
- Τα σύμβολα '/', 'Q' στην κατηγορία Q (Unknown).

## 7.2.2 Ανίχνευση R-peaks και RR-based Segmentation

Προτού χρησιμοποιηθεί η ταξινόμηση των χτύπων βάση AAMI, είναι απαραίτητη η κατάλληλη τμηματοποίηση του ΗΚΓ σήματος στους χτύπους αυτούς. Στη συνέχεια παρουσιάζεται η διαδικασία ανίχνευσης RR διαστημάτων και δημιουργίας των αντίστοιχων segments. Η επεξεργασία, εφαρμόζεται διαδοχικά σε κάθε αρχείο καταγραφής (record) και έχει ως στόχο την καθαρή εξαγωγή των R-peaks και την κατασκευή των επιμέρους heartbeat segments. Σε ορισμένα πειράματα εφαρμόζεται πρώτα στα δεδομένα bandpass φίλτρο με συχνότητες αποκοπής 0.5 με 40 Hz για την αποθρομβοποίηση του σήματος. Αρχικά, από κάθε αρχείο καταγραφής ανακτώνται τα σήματα ECG, καθώς και οι αντίστοιχες χρονικές σημάνσεις που υποδεικνύουν τις θέσεις των καρδιακών παλμών (R-peaks) και η διαγνωστική τους ετικέτα. Από τις ανακτημένες σημάνσεις (annotations), διατηρούνται μόνο εκείνοι οι παλμοί των οποίων τα σύμβολα αντιστοιχούν σε κάποια από τις πέντε κατηγορίες του προτύπου AAMI. Το φιλτράρισμα αυτό, επιτρέπει την απομάκρυνση ανεπιθύμητων ή μη διαγνώσιμων παλμών, διασφαλίζοντας έτσι ότι στο τελικό σύνολο δεδομένων περιλαμβάνονται μόνο οι παλμοί που είναι μορφολογικά και διαγνωστικά κατάλληλοι για ταξινόμηση. Αφού ολοκληρωθεί ο καθαρισμός των σημάνσεων, ακολουθεί η κατασκευή των επιμέρους heartbeat segments, με βάση τα διαδοχικά RR-διαστήματα (RR-intervals). Συγκεκριμένα, κάθε έγκυρος παλμός (heartbeat segment) ορίζεται ως το χρονικό διάστημα που ξεκινά από ένα εντοπισμένο R-peak και καταλήγει στο αμέσως επόμενο. Τα κανάλια που υπάρχουν σε κάθε αρχείο είναι δύο, αλλά στην συγκεκριμένη εργασία χρησιμοποιείται μόνο το πρώτο κανάλι, το οποίο περιέχει το MLI. Κάθε heartbeat συνοδεύεται από το όνομα της καταγραφής από την οποία προήλθε, τη θέση του στην ακολουθία (heartbeat index), το αντίστοιχο απόσπασμα σήματος

από το κανάλι, καθώς και το σύμβολο annotation που το χαρακτηρίζει, το οποίο έχει αντιστοιχηθεί σε μία από τις πέντε κατηγορίες του προτύπου AAMI. Στον Πίνακα 1 φαίνεται η τελική κατανομή των κλάσεων στο σύνολο των δεδομένων.

Πίνακας 1: Κατανομή παλμών ανά κατηγορία AAMI στο τελικό σύνολο δεδομένων

| Κατηγορία AAMI | Πλήθος παλμών |
|----------------|---------------|
| N              | 90589         |
| V              | 7236          |
| Q              | 7057          |
| S              | 2779          |
| F              | 803           |

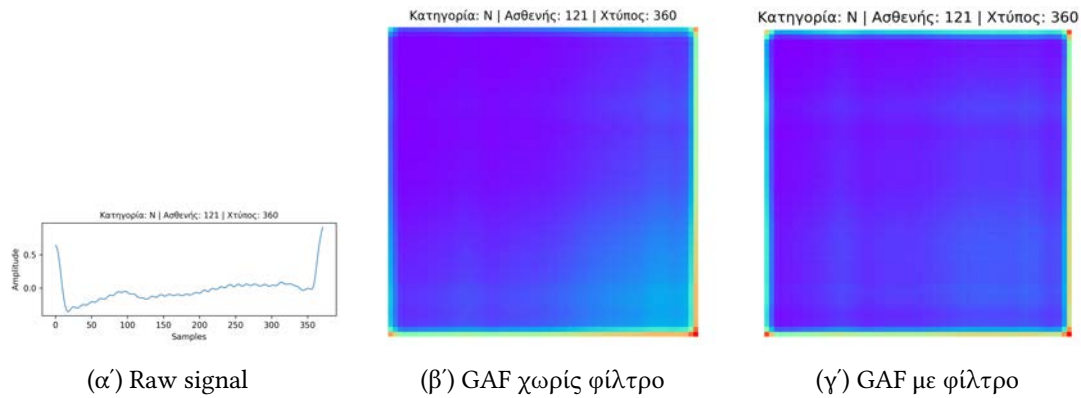
### 7.2.3 Μετασχηματισμός GAF

Τα segments που έχουν δημιουργηθεί, δεν έχουν σταθερό μήκος, καθώς οι χτύποι και οι θέσεις των R-peaks διαφέρουν ακόμα και στον ίδιο άνθρωπο, υπό φυσιολογικές συνθήκες. Η συνάρτηση gaf (Gramian Angular Field ) που θα χρησιμοποιηθεί απαιτεί κομμάτια σταθερού μήκους για είσοδο, επομένως για την τυποποίηση του μήκους όλων των RR-based καρδιοχτύπων υλοποιήθηκε και εφαρμόστηκε επεξεργασία του μήκους, με στόχο την εξασφάλιση σταθερού αριθμού δειγμάτων ανά segment με target length=256. Συγκεκριμένα, όταν το μήκος του segment είναι μικρότερο από το προκαθορισμένο εφαρμόζεται zero-padding και όταν το segment είναι μεγαλύτερο από το επιθυμητό, χρησιμοποιείται επαναδειγματοληψία (resampling). Σε περίπτωση που το μήκος του σήματος συνέπιπτε ήδη με τον επιθυμητό αριθμό δειγμάτων, το σήμα παρέμενε αμετάβλητο. Με τον τρόπο αυτό, κάθε heartbeat segment απέκτησε ενιαίο μήκος εισόδου, καθιστώντας δυνατή τη χρήση του ως input σε απευθείας μορφή για μετασχηματισμό εικόνας αλλά και για την μετέπειτα είσοδό του σε νευρωνικό δίκτυο. Στη συνέχεια, πραγματοποιείται κανονικοποίηση των segments στην περιοχή [0,1] με τη χρήση της MinMaxScaler, πριν μετασχηματιστούν με GAF (Gramian Angular Field). Οι επεξεργασίες (κανονικοποίηση μήκους, κανονικοποίηση) των τμημάτων (segments) και η εφαρμογή GAF, πραγματοποιούνται ξεχωριστά για κάθε χτύπο, και επομένως, η κανονικοποίηση πραγματοποιείται βάση των τιμών στο εκάστοτε τμήμα προς επεξεργασία και όχι σε όλο το εύρος των segments, καθώς αυτό θα μπορούσε να δημιουργήσει προβλήματα όπως data leak.

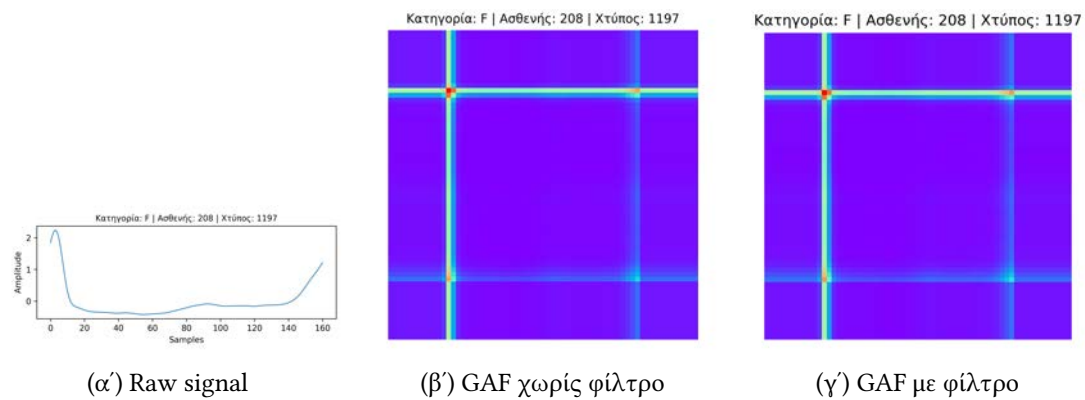
Στις εικόνες 1 , 2, 3, 4 και 5 για κάθε μία από τις πέντε κατηγορίες AAMI έχει δημιουργηθεί μία ενδεικτική εικόνα που περιλαμβάνει τρεις απεικονίσεις του ίδιου καρδιακού παλμού:

1. το αρχικό σήμα (raw signal),
2. την εικόνα GAF που προκύπτει από το ακατέργαστο (μη φιλτραρισμένο) σήμα και

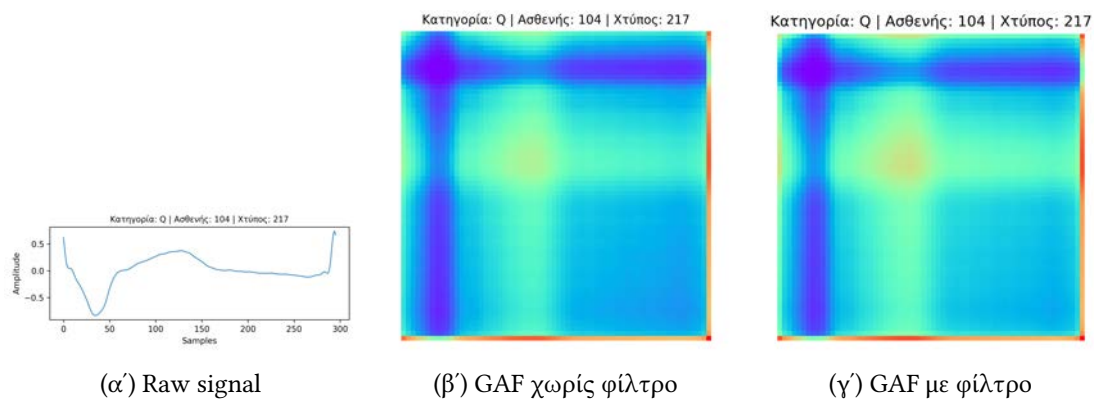
3. την αντίστοιχη εικόνα GAF που προκύπτει από το σήμα μετά την εφαρμογή φίλτρου αφαίρεσης θορύβου



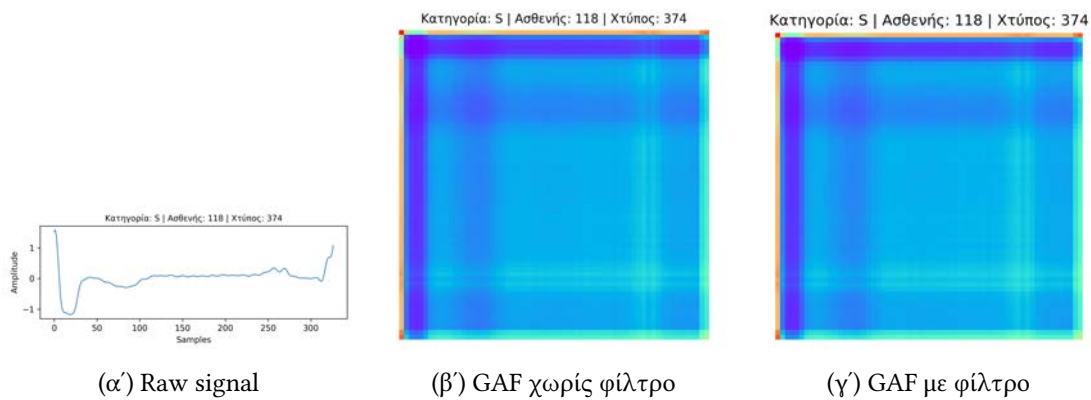
Σχήμα 1: Σύγκριση τριών απεικονίσεων για έναν χτύπο (κατηγορία N).



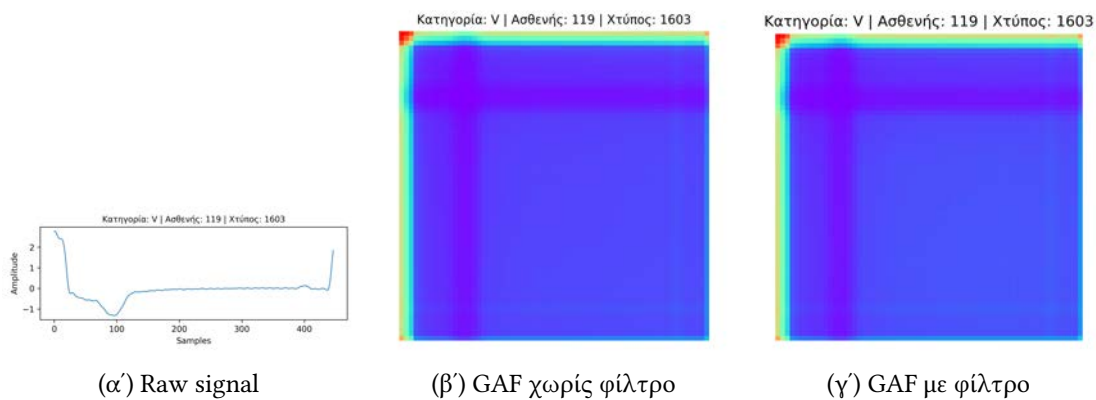
Σχήμα 2: Σύγκριση τριών απεικονίσεων για έναν χτύπο (κατηγορία F).



Σχήμα 3: Σύγκριση τριών απεικονίσεων για έναν χτύπο (κατηγορία Q).



Σχήμα 4: Σύγκριση τριών απεικονίσεων για έναν χτύπο (κατηγορία S).



Σχήμα 5: Σύγκριση τριών απεικονίσεων για έναν χτύπο (κατηγορία V).

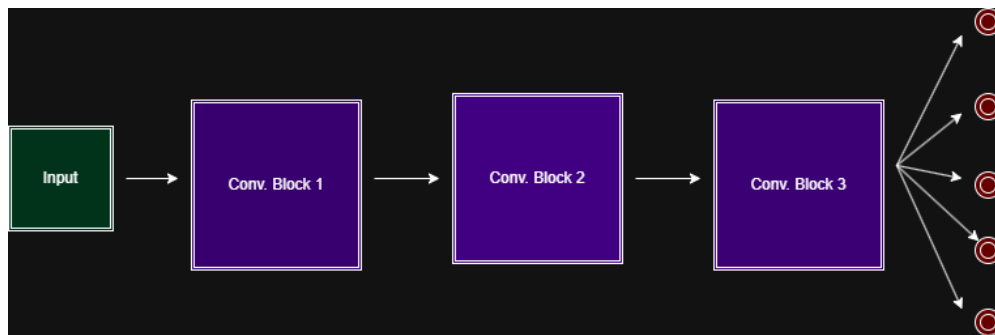
#### 7.2.4 Δημιουργία και επεξεργασία εισόδου για το νευρωνικό

Έχοντας τις εικόνες GAF από κάθε segment, στο επόμενο στάδιο χρειάζεται η κατάλληλη επεξεργασία τους για να δοθούν ως είσοδοι στο νευρωνικό δίκτυο. Συγκεκριμένα, πραγματοποιείται χωρισμός των δεδομένων αρχικά σε train και test datasets και στη συνέχεια σε train και validation. Αφού έχουν δημιουργηθεί τα train, test, validation datasets χρησιμοποιούνται οι συναρτήσεις ECGDataset και DataLoader. Η κλάση ECGDataset χρησιμοποιείται για να προετοιμάσει τα δεδομένα στην κατάλληλη μορφή πριν αυτά δοθούν στον DataLoader. Κάθε φορά που καλείται μέσω ενός index από τον DataLoader, πραγματοποιεί τις απαραίτητες μετατροπές των δεδομένων σε μορφή numpy και torch.Tensor, και στη συνέχεια δημιουργεί μία εικόνα διαστάσεων (1, 64, 64). Η εικόνα αυτή δηλαδή θα αποτελείται από ένα κανάλι(channel), το οποίο αντιστοιχεί στην GAF εικόνα που προέρχεται από το αρχικό κανάλι MLI του ECG σήματος. Για την τμηματική τροφοδότηση του μοντέλου με δεδομένα κατά την εκπαίδευση και την επικύρωση, χρησιμοποιήθηκε ο μηχανισμός DataLoader του PyTorch, με μέγεθος παρτίδας (batch size) ίσο με 32. Ο μηχανισμός αυτός επιτρέπει τη σταδιακή φόρ-

τωση των δεδομένων από custom Dataset κλάση, διευκολύνοντας την αποδοτική χρήση της μνήμης.

### 7.3 Αρχιτεκτονική του Νευρωνικού Δικτύου (CNN)

Για την ταξινόμηση των heartbeat segments σε μία από τις 5 κατηγορίες χρησιμοποιείται συνελκτικό δίκτυο (CNN), το οποίο αποτελείται από τρία διαδοχικά συνελκτικά blocks κι έναν classifier όπως φαίνεται στην Εικόνα 6.



Σχήμα 6: Αρχιτεκτονική του συνελκτικού νευρωνικού δικτύου (CNN) που χρησιμοποιείται για την ταξινόμηση των καρδιακών παλμών.

Κάθε block αποτελείται από ένα συνελκτικό επίπεδο με kernel  $3 \times 3$  και padding 1, ακολουθούμενο από κανονικοποίηση παρτίδας (Batch Normalization), ενεργοποίηση ReLU, υποδειγματοληψία μέσω max pooling ( $2 \times 2$ ) και dropout με πιθανότητα  $p$  (εδώ  $p = 0.3$ ). Από το χώρισμα των δεδομένων σε train, test, validation datasets υπολογίστηκαν ποσοστά των κλάσεων και επιβεβαιώθηκε αυτό που είχε παρατηρηθεί από την αρχή, η ύπαρξη ανισορροπίας μεταξύ κλάσεων. Οπότε ήδη από την αρχή προσθέτονται dropout layers και L2 regularization. Παρόλο που οι μέθοδοι αυτές δεν βοηθάνε απευθείας στη λύση του προβλήματος της ανισορροπίας, λειτουργούν συμπληρωματικά ενισχύοντας τη γενίκευση του μοντέλου και περιορίζοντας την ευαισθησία του σε υπερπροσαρμοσμένα πρότυπα. Τα φίλτρα (δηλαδή οι διαφορετικοί πυρήνες που εξάγουν χαρακτηριστικά) αυξάνονται σε κάθε επίπεδο, από 16 σε 32 και 64, ενώ αντίστροφα το μέγεθος του χάρτη χαρακτηριστικών μειώνεται από 64 σε 8 και 8. Ο τελικός αυτός χάρτης έχοντας υποστεί flattening δίνεται σε classifier, που αποτελείται από 128 πλήρως συνδεδεμένους νευρώνες και συνάρτηση ενεργοποίησης ReLU, ακολουθούμενο από dropout. Η έξοδος του δικτύου είναι ένα διάνυσμα 5 διαστάσεων, όπου κάθε τιμή αντιπροσωπεύει την τιμή συσχέτισης της εικόνας εισόδου με την κάθε κλάση.

---

## 7.4 Παραλλαγές Μοντέλων και Βελτιστοποιήσεις

Πέρα από τον ορισμό αρχιτεκτονικής μοντέλου, εφαρμόστηκαν και ορισμένες τεχνικές βελτιστοποίησης, προκειμένου να αντιμετωπιστούν οι επιπτώσεις της ανισορροπίας των κατηγοριών και να ενισχυθεί η συνολική απόδοση. Οι μέθοδοι αυτοί παρουσιάζονται παρακάτω.

### 7.4.1 Υλοποίηση προσθήκης Optuna για βάρη

Στα δεδομένα υπάρχει σημαντική ανισορροπία κλάσεων, με τις κλάσεις S και F να σημειώνουν πολύ μικρότερη εμφάνιση στα δεδομένα. Οπότε αρχικά γίνεται προσθήκη βαρών για τις 5 κατηγορίες. Προκειμένου να αντιμετωπιστεί περαιτέρω η ανισορροπία στις κλάσεις χρησιμοποιείται η βιβλιοθήκη Optuna για προσδιορισμό βέλτιστων βαρών για τις δύο αυτές κατηγορίες στα βάρη που έχει ήδη υπολογίσει. Για την ενίσχυση του recall στο validation set των κατηγοριών S και F με βάρη, σχεδιάστηκε η συνάρτηση objective, η οποία χρησιμοποιείται από το Optuna για την αξιολόγηση της απόδοσης κάθε συνδυασμού παραμέτρων. Η συνάρτηση αυτή δέχεται τις παραμέτρους που προτείνονται σε κάθε trial, εκπαιδεύει το μοντέλο με αυτές και επιστρέφει τη σχετική μετρική απόδοσης.

Για κάθε δοκιμή συνδυασμών βαρών, γίνεται εκπαίδευση του μοντέλου για 3 εποχές ώστε να αξιολογηθεί γρήγορα η επίδραση των επιλεγμένων βαρών. Το μοντέλο που εκπαιδεύτηκε αξιολογείται στο validation set και υπολογίζεται το macro recall. Η τιμή αυτή επιστρέφεται από τη συνάρτηση objective στο Optuna, το οποίο χρησιμοποιεί αλγόριθμο Bayesian-based για να διαμορφωθεί σύγκριση με άλλες τιμές βαρών που θα προκύψουν από επόμενα trials και τις αποδόσεις τους. Πριν από κάθε δοκιμή (trial), το Optuna καθοδηγεί τη διαδικασία επιλογής, εστιάζοντας σε ένα στοχευμένο υποσύνολο του συνολικού διαστήματος τιμών. Μετά το πέρας των 50 δοκιμών, το Optuna επιστρέφει τις τιμές των μεταβλητών που οδήγησαν στο μέγιστο macro recall στο validation set. Βάση αυτών των τιμών θα εκπαιδευτεί πλήρως το συνελικτικό δίκτυο και στη συνέχεια θα αξιολογηθεί στο test dataset.

### 7.4.2 Υλοποίηση grid search για κατώφλια

Η χρήση βαρών βελτίωσε σημαντικά τα recalls των κλάσεων που ήταν σε μειοψηφία, μειώνοντας όμως το precision τους, τοποθετώντας και προβλέψεις με χαμηλή τιμή στις κλάσεις αυτές (αυξημένα false positives). Σε προσπάθεια να ενισχυθούν τα precision, χωρίς σημαντική απώλεια στα recalls, τοποθετήθηκαν κατώφλια για τις κλάσεις S, F έτσι ώστε αν το ποσοστό προβλεψής στην κλάση αυτή είναι πολύ μικρό να μην τοποθετείται στην κλάση αυτή, αλλά by default στην κλάση N. Στην πράξη, αφού το μοντέλο εξάγει τις πιθανότητες για κάθε κατηγορία, εντοπίζεται η κλάση με τη μέγιστη πιθανότητα (η προβλεπόμενη) και συγκρίνεται η τιμή της με το αντίστοιχο κατώφλι. Εάν η πιθανότητα υπερβαίνει το threshold,

η πρόβλεψη διατηρείται· διαφορετικά, η κατηγορία αντικαθίσταται με την N (Normal). Με αυτόν τον τρόπο απορρίπτονται προβλέψεις χαμηλής εμπιστοσύνης για τις S και F, επιτρέποντας τον έλεγχο του trade-off μεταξύ precision και recall. Τέλος, για κάθε δοκιμή υπολογίζεται το macro F1-score, και διατηρείται ο συνδυασμός κατωφλίων που οδηγεί στην καλύτερη τιμή της μετρικής. Με βάση αυτό, στο τέλος του αλγορίθμου προτείνεται ως βέλτιστο το ζεύγος κατωφλίων που απέδωσε τη μέγιστη τιμή F1.

## 7.5 Overfitting

Εκτός της ανάγκης βελτιστοποίησης παραμέτρων, ένα άλλο πρόβλημα που μπορεί να παρουσιαστεί κατά την εκπαίδευση νευρωνικών δικτύων είναι η υπερπροσαρμογή (overfitting), κατά την οποία το μοντέλο αποστηθίζει τα δεδομένα εκπαίδευσης, συμπεριλαμβανομένων των θορύβων και των τυχαίων διαφορών. Αυτό οδηγεί σε πολύ καλή απόδοση του μοντέλου στα δεδομένα εκπαίδευσης, αλλά σε χαμηλή απόδοση σε καινούρια και άγνωστα δεδομένα, κάτι που φαίνεται στα σύνολα επικύρωσης ή δοκιμών (validation and test dataset) [28]. Δύο σημαντικές τεχνικές αντιμετώπισης της υπεροπροσαρμογής είναι η κανονικοποίηση (regularization) και η πρόωπη διακοπή (early stopping). Στην κανονικοποίηση ανήκουν τεχνικές όπως το dropout, η L1 και η L2 κανονικοποίηση, οι οποίες περιορίζουν την πολυπλοκότητα του μοντέλου. Η πρόωπη διακοπή σταματά την εκπαίδευση όταν η απόδοση στο σύνολο επικύρωσης (validation dataset) αρχίσει να επιδεινώνεται.

### 7.5.1 L1 Regularization

Η L1 regularization είναι μια μέθοδος κανονικοποίησης που προσθέτει έναν όρο στη συνάρτηση απώλειας, βασισμένο στο άθροισμα των απόλυτων τιμών των βαρών. Η L1 δημιουργεί αραιές λύσεις, όπου η πλειονότητα των τιμών των βαρών είναι 0, εξαιτίας του ότι λόγω της απόλυτης τιμής των βαρών αγνοούνται τα περισσότερα υπολογιστικά σφάλματα. Είσοδοι δηλαδή των οποίων τα βάρη παίζουν μικρό ή και καθόλου ρόλο, αγνοούνται και με αυτόν τρόπο η L1 πραγματοποιεί επιλογή προβλέψιμων χαρακτηριστικών. Επιπλέον το γεγονός ότι τα βάρη κάποιων συνδέσεων έχουν απορριφθεί, οδηγεί σε αραιώση του νευρωνικού δικτύου [29]. Παρακάτω φαίνεται ο τύπος που χρησιμοποιεί η L1:

$$\mathcal{L}_{L1} = \mathcal{L}_{\text{original}} + \lambda \sum_{i=1}^n |w_i| \quad (7)$$

Όπου:

- $w = w_i$ , με  $w_i$  τα βάρη των χαρακτηριστικών
- $\mathcal{L}_{L1}$  είναι η νέα συνάρτηση απώλειας με κανονικοποίηση L1

- $\mathcal{L}_{\text{original}}$  είναι η αρχική συνάρτηση απώλειας χωρίς κανονικοποίηση
- $\lambda$  είναι η παράμετρος κανονικοποίησης
- $n$  είναι ο αριθμός των χαρακτηριστικών

Δύο σημαντικές τεχνικές αντιμετώπισης της υπεροπροσαρμογής είναι η κανονικοποίηση (regularization) και η πρόωπη διακοπή (early stopping). Στην κανονικοποίηση ανήκουν τεχνικές όπως το dropout, η L1 και η L2 κανονικοποίηση, οι οποίες περιορίζουν την πολυπλοκότητα του μοντέλου. Η πρόωπη διακοπή σταματά την εκπαίδευση όταν η απόδοση στο σύνολο επικύρωσης (validation dataset) αρχίσει να επιδεινώνεται.

### 7.5.2 L2 Regularization

Η L2 λειτουργεί παροτρύνει το μοντέλο να καταναίμει με ίσο τρόπο τα βάρη στα διάφορα χαρακτηριστικά, αποτρέποντας έτσι το δίκτυο να εστιάσει και άρα και να εξαρτηθεί σε μεγάλο βαθμό από ορισμένα μόνο χαρακτηριστικά. Με αυτόν τον τρόπο, επιτυγχάνεται η μείωση της υπεροπροσαρμογής. Παρακάτω φαίνεται ο τύπος που χρησιμοποιεί η L2:

$$\mathcal{L}_{L2} = \mathcal{L}_{\text{original}} + \lambda \sum_{i=1}^n w_i^2 \quad (8)$$

Όπου:

- $w = w_i$ , με  $w_i$  τα βάρη των χαρακτηριστικών
- $\mathcal{L}_{L2}$  είναι η νέα συνάρτηση απώλειας με κανονικοποίηση L2
- $\mathcal{L}_{\text{original}}$  είναι η αρχική συνάρτηση απώλειας χωρίς κανονικοποίηση
- $\lambda$  είναι η παράμετρος κανονικοποίησης
- $n$  είναι ο αριθμός των χαρακτηριστικών

Η L2 κατηγοριοποίηση, λόγω του τετραγώνου στον τύπο, είναι ευαίσθητη σε ακραίες τιμές (outliers), ενώ διατηρεί όλα τα χαρακτηριστικά των δεδομένων, οδηγώντας σε λιγότερο ακραία μοντέλα από ό,τι η L1 κατηγοριοποίηση.

## 8 Αποτελέσματα

### 8.1 Πειραματικές ρυθμίσεις

Για την πειραματική μελέτη αξιολογήθηκε η ικανότητα ενός συνελκτικού νευρωνικού δικτύου να ταξινομεί παλμούς ηλεκτροκαρδιογραφήματος (ECG) σε πέντε διαγνωστικές κατηγορίες, όπως ορίζονται από το πρότυπο AAMI. Τα δεδομένα προήλθαν από τη βάση MIT-BIH

---

Arrhythmia Database, από την οποία επιλέχθηκαν μόνο οι καταγραφές που περιλαμβάνουν ως πρώτο κανάλι το MLII, το οποίο είναι και το πιο ευρέως χρησιμοποιούμενο σε σχετικές μελέτες. Τα αρχεία των ασθενών 102, 104 και 114 είναι περιπτώσεις όπου το πρώτο κανάλι δεν αντιστοιχούσε στο πρότυπο MLII και για αυτόν το λόγο, εξαιρέθηκαν από τη μελέτη, ώστε να διασφαλιστεί η ομοιογένεια της εισόδου προς το μοντέλο. Εκτελούνται διαδοχικές δοκιμές πάνω στην ίδια αρχιτεκτονική, εφαρμόζοντας κάθε φορά διαφορετικές ρυθμίσεις και παραμέτρους με σκοπό τη βελτίωση της τελικής απόδοσης. Η αποτίμηση της απόδοσης του μοντέλου πραγματοποιήθηκε με βάση τέσσερις βασικές μετρικές: ακρίβεια (accuracy), θετική προγνωστική τιμή (precision), ευαισθησία (recall) και F1-score. Οι μετρικές αυτές εφαρμόστηκαν τόσο στο συνολικό σύνολο όσο και ανά κατηγορία, ώστε να είναι εμφανής η συμπεριφορά του μοντέλου για όλες τις κατηγορίες και να διακρίνεται η απόδοσή του, τόσο στις σπάνιες και όσο και στις συχνές κατηγορίες. Το F1-score αποτέλεσε βασικό κριτήριο αξιολόγησης, καθώς χρησιμοποιήθηκε για την εκτίμηση της ισορροπίας μεταξύ recall και precision, ιδιαίτερα στο πλαίσιο ενίσχυσης της απόδοσης στις υποεκπροσωπούμενες κατηγορίες. Συγκεκριμένα και δεδομένου ότι η αύξηση του recall συχνά συνεπάγεται μείωση στο precision, το F1-score παρείχε μια πιο αντιπροσωπευτική εικόνα της συνολικής επίδοσης του μοντέλου, ιδίως όταν επιδιώκεται ταυτόχρονη ικανοποιητική συμπεριφορά και στους δύο δείκτες.

## 8.2 CNN μοντέλο χωρίς βάρη και κατώφλι με φίλτρο

Στο πλαίσιο αυτής της προσέγγισης, το μοντέλο εκπαιδεύεται και αξιολογείται πάνω σε δεδομένα που έχουν υποστεί προεπεξεργασία αφαίρεσης θορύβου. Στο μοντέλο εφαρμόζονται γενικές τεχνικές regularization, όπως dropout και L2, αλλά δεν χρησιμοποιούνται ειδικά σχεδιασμένες τεχνικές για την αντιμετώπιση της ανισορροπίας των κατηγοριών, όπως η τροποποίηση των βαρών απώλειας ή η ρύθμιση κατωφλίων πρόβλεψης. Κατ' επέκταση, η ταξινόμηση βασίζεται αποκλειστικά στη βασική ικανότητα γενίκευσης της αρχιτεκτονικής του μοντέλου.

Το μοντέλο εμφανίζει γενικά καλά metrics. Συγκεκριμένα, οι κλάσεις N και Q φαίνεται να αναγνωρίζονται με καθολική ακρίβεια από το μοντέλο, καθώς όπως φαίνεται και στο πίνακα 2 εμφανίζουν εξαιρετικά υψηλές επιδόσεις σε όλα τα metrics (Precision, Recall, F1-score). Η κατηγορία V παρουσιάζει ελαφρώς χαμηλότερες τιμές, αλλά διατηρεί επίσης υψηλή επίδοση άνω του 92.00% σε όλα τα metrics.

Παρά τις υψηλές επιδόσεις που καταγράφονται στις πλειοψηφικές κατηγορίες, το μοντέλο παρουσιάζει χαμηλότερη απόδοση στις κλάσεις S και F, οι οποίες αποτελούν και τις λιγότερο αντιπροσωπευμένες στο σύνολο των δεδομένων. Συγκεκριμένα, το recall για την κατηγορία S περιορίζεται στο 65.34% και για την κατηγορία F στο 58.33%, κάτι το οποίο επηρεάζει και φαίνεται και στις αντίστοιχες τιμές των F1-score, οι οποίες παραμένουν εμφανώς χαμηλότερες συγκριτικά με τις υπόλοιπες.

Πίνακας 2: Ανάλυση ανά κατηγορία για το μοντέλο: 1. Simple CNN με φίλτρο

| Κλάση | Precision | Recall | F1-score |
|-------|-----------|--------|----------|
| N     | 0.9810    | 0.9977 | 0.9893   |
| S     | 0.9524    | 0.6534 | 0.7750   |
| V     | 0.9679    | 0.9251 | 0.9460   |
| F     | 0.8522    | 0.5833 | 0.6926   |
| Q     | 0.9973    | 0.9855 | 0.9914   |

### 8.3 CNN μοντέλο χωρίς βάρη και κατώφλι και χωρίς φίλτρο

Στην παρούσα παραλλαγή του μοντέλου, το συνελκτικό νευρωνικό δίκτυο εκπαιδεύτηκε και αξιολογήθηκε στις εικόνες GAF που προέρχονται από το ακατέργαστο ΗΚΓ σήμα, χωρίς να έχει εφαρμοσθεί προεπεξεργασία αφαίρεσης θορύβου στα δεδομένα. Επιπλέον, δεν εφαρμόστηκαν τεχνικές εξισορρόπησης των δεδομένων, όπως η χρήση βαρών απώλειας ή η ρύθμιση κατωφλίων πρόβλεψης.

Το μοντέλο παρουσιάζει πολύ υψηλά επίπεδα precision, σε όλες τις κατηγορίες, συμπεριλαμβανομένων και των μειοψηφικών S και F. Συγκεκριμένα όλες οι τιμές precision υπερβαίνουν το 94.00%.

Ωστόσο, η απόδοση στο recall για τις κατηγορίες S και F υπολείπεται σημαντικά, γεγονός που αντικατοπτρίζεται και στις αντίστοιχες τιμές του F1-score. Ειδικά για την κατηγορία F, το recall της είναι 42.26% και το F1-score έχει την τιμή 59.17%, ενώ η κατηγορία S έχει recall 63.88% και F1-score 76.19%, τιμές που απέχουν αισθητά από εκείνες των υπόλοιπων κλάσεων, όπως φαίνεται στον πίνακα 3.

Πίνακας 3: Ανάλυση ανά κατηγορία για το μοντέλο: 2. Simple CNN χωρίς φίλτρο

| Κλάση | Precision | Recall | F1-score |
|-------|-----------|--------|----------|
| N     | 0.9806    | 0.9979 | 0.9892   |
| S     | 0.9437    | 0.6388 | 0.7619   |
| V     | 0.9576    | 0.9367 | 0.9470   |
| F     | 0.9861    | 0.4226 | 0.5917   |
| Q     | 0.9960    | 0.9894 | 0.9927   |

### 8.4 CNN μοντέλο με βάρη και φίλτρο

Σε αυτό το πείραμα έχουν εφαρμοσθεί προσαρμοσμένα βάρη στο loss function με σκοπό την αντιστάθμιση της ανισορροπίας στις κλάσεις του συνόλου εκπαίδευσης.

Η επίδραση των βαρών αποτυπώνεται στο recall των μειοψηφικών κατηγοριών, καθώς

το recall για την κατηγορία S είναι 84.21% και για την κατηγορία F 84.52%.

Ωστόσο, παρατηρούνται για τις δύο αυτές κλάσεις, χαμηλά precision. Συγκεκριμένα, για την κατηγορία S το precision είναι 52.67%, ενώ για την κατηγορία F μόλις 23.87%. Η χαμηλή αυτή απόδοση αποτυπώνεται και στις αντίστοιχες τιμές F1-score, με την κατηγορία F να εμφανίζει 37.22%, παρά το υψηλό recall της.

Οι κλάσεις N, Q έχουν πολύ καλές αποδόσεις, με όλα τους τα metrics να έχουν τιμές πάνω από 94.00%, ενώ η κλάση V έχει επίσης ισχυρό accuracy και F1-score με τιμές 94.49% και 90,54% αντίστοιχα και λίγο χαμηλότερο αλλά εξίσου καλό precision 86,91%. Οι τιμές των metrics για τις κλάσεις σε αυτό το πείραμα φαίνονται στον πίνακα 4.

Πίνακας 4: Ανάλυση ανά κατηγορία για το μοντέλο: 3. Προσθήκη βαρών με φίλτρο

| Κλάση | Precision | Recall | F1-score |
|-------|-----------|--------|----------|
| N     | 0.9933    | 0.9433 | 0.9676   |
| S     | 0.5267    | 0.8421 | 0.6480   |
| V     | 0.8691    | 0.9449 | 0.9054   |
| F     | 0.2387    | 0.8452 | 0.3722   |
| Q     | 0.9908    | 0.9934 | 0.9921   |

## 8.5 CNN μοντέλο με βάρη χωρίς φίλτρο

Σε αυτό το πείραμα, δεν χρησιμοποιείται στα δεδομένα φίλτρο αφαίρεσης του θορύβου ενώ γίνεται προσπάθεια εξισορρόπησης των κλάσεων χρησιμοποιώντας βάρη.

Για τις μειοψηφικές κατηγορίες S και F, καταγράφονται υψηλές τιμές recall 82.21% και 83.33% αντίστοιχα, υποδεικνύοντας αυξημένη ικανότητα εντοπισμού θετικών περιπτώσεων. Ωστόσο, το precision για τις δύο αυτές κλάσεις είναι χαμηλό, με την S να φτάνει το 40.12% και την F το 24.22%. Οι τιμές αυτές επηρεάζουν άμεσα και τα αντίστοιχα F1-scores, τα οποία διαμορφώνονται σε 53.93% για την S και 37.53% για την F.

Για τις κλάσεις N, Q παρατηρούνται υψηλά metrics, με τιμές μεγαλύτερες του 93.00%. Η κλάση V έχει επίσης υψηλές τιμές precision 84.42%, recall 95.51% και F1-score 89.62%. Από τις αποδόσεις αυτές, παρατηρείται ότι τα βάρη οδηγούν σε υψηλό recall των λιγότερα εμφανιζόμενων κλάσεων, ρίχνοντας όμως σημαντικά το precision τους, όπως φαίνεται στον πίνακα 5.

Πίνακας 5: Ανάλυση ανά κατηγορία για το μοντέλο: 4. Προσθήκη βαρών χωρίς φίλτρο

| Κλάση | Precision | Recall | F1-score |
|-------|-----------|--------|----------|
| N     | 0.9926    | 0.9262 | 0.9583   |
| S     | 0.4012    | 0.8221 | 0.5393   |
| V     | 0.8442    | 0.9551 | 0.8962   |
| F     | 0.2422    | 0.8333 | 0.3753   |
| Q     | 0.9881    | 0.9855 | 0.9868   |

## 8.6 CNN μοντέλο με fine tuning βαρών και φίλτρο

Παρακάτω παρουσιάζονται τα αποτελέσματα για CNN μοντέλο στο οποίο έχει χρησιμοποιηθεί *ortuna* για εύρεση των βέλτιστων τιμών βαρών των κλάσεων που υποεκπροσωπούνται, δηλαδή των κλάσεων S και F. Στις υπόλοιπες κλάσεις παραμένουν τα βάρη που έχουν βρεθεί από την κατάλληλη συνάρτηση εύρεσης προσαρμοσμένων βαρών. Όπως φαίνεται και στον πίνακα 6 παρατηρούνται υψηλά *metrics* στις τρεις υπερέχουσες κλάσεις N, V, Q, με την κλάση Q να έχει τιμές μεγαλύτερες του 98.00% σε όλα τα *metrics*. Μόνο το *precision* της κλάσης V έχει λίγο χαμηλότερη τιμή 79.74%. Οι κλάσεις S και F οι οποίες βρίσκονται σε μικρότερο ποσοστό στα δεδομένα, εμφανίζουν και οι δύο υψηλό *recall* της τιμής 86.39% και 82.74% αντίστοιχα, αλλά έχουν χαμηλό *precision*, 36.12% για την κλάση S και 28.60% για την κλάση F, γεγονός που σημαίνει ότι λιγότερες από τις μισές προβλέψεις του μοντέλου για τις κατηγορίες αυτές είναι ορθές.

Πίνακας 6: Ανάλυση ανά κατηγορία για το μοντέλο: 5. Fine tuning βαρών με φίλτρο

| Κλάση | Precision | Recall | F1-score |
|-------|-----------|--------|----------|
| N     | 0.9942    | 0.9169 | 0.9539   |
| S     | 0.3612    | 0.8639 | 0.5094   |
| V     | 0.7974    | 0.9483 | 0.8663   |
| F     | 0.2860    | 0.8274 | 0.4251   |
| Q     | 0.9843    | 0.9947 | 0.9895   |

## 8.7 CNN μοντέλο με fine tuning βαρών χωρίς φίλτρο

Σε αυτήν την περίπτωση χρησιμοποιείται μέθοδος βελτιστοποίησης των βαρών για τις κατηγορίες S και F, ενώ δεν χρησιμοποιείται φίλτρο αποθρομβοποίησης. Διαπιστώνεται πως η προσθήκη βαρών έχει επίδραση στις τιμές των *recalls* των κλάσεων S και F, οι οποίες είναι πολύ κοντά μεταξύ τους. Συγκεκριμένα, η κλάση S έχει *recall* 84.75% και η κλάση F 84.52%. Γενικά, όπως φαίνεται και στον πίνακα 7, τα *recalls* όλων των κλάσεων είναι υψηλά, πάνω

από 84.00% για τις μειοψηφικές κλάσεις και πάνω από 93.00% για τις υπόλοιπες. Ωστόσο, παρατηρείται διαφορά στις αποδόσεις ανά κλάση των precisions και των αντίστοιχων F1-scores, όπου οι κλάσεις S και F έχουν πολύ χαμηλότερες τιμές από τις υπόλοιπες κλάσεις, με την S να έχει 39.58% και την F 27.36%.

Πίνακας 7: Ανάλυση ανά κατηγορία για το μοντέλο: 6. Fine tuning βαρών χωρίς φίλτρο

| Κλάση | Precision | Recall | F1-score |
|-------|-----------|--------|----------|
| N     | 0.9929    | 0.9313 | 0.9611   |
| S     | 0.3958    | 0.8475 | 0.5396   |
| V     | 0.8707    | 0.9353 | 0.9019   |
| F     | 0.2736    | 0.8452 | 0.4134   |
| Q     | 0.9882    | 0.9960 | 0.9921   |

## 8.8 CNN μοντέλο με βάρη, κατώφλι και φίλτρο

Το παρόν μοντέλο αξιοποιεί ρύθμιση κατωφλίων πρόβλεψης για τις κατηγορίες S και F, με στόχο την ενίσχυση της απόδοσης στις μειοψηφικές κλάσεις S και F, σε συνδυασμό με φίλτρο αφαίρεσης θορύβου στα δεδομένα.

Όπως παρουσιάζεται στον Πίνακα 8, οι πλειοψηφικές κατηγορίες N και Q εμφανίζουν υψηλές επιδόσεις σε όλα τα metrics, με F1-score 97.77% και 98.95% αντίστοιχα, υποδηλώνοντας ακρίβεια και αξιοπιστία στην ταξινόμησή τους.

Η κατηγορία V εμφανίζει ισχυρό recall 94.83% και ικανοποιητικό precision 79.74%, με F1-score 86.63%, γεγονός που δείχνει αποτελεσματική ανίχνευση παθολογικών περιστατικών αυτής της κατηγορίας, παρά την ελαφρώς μειωμένη ακρίβεια.

Η κλάση S έχει ισορροπία μεταξύ precision και recall με τιμές 74.22% και 73.14% αντίστοιχα κάτι που οδηγεί και στην τιμή του F1-score να είναι κοντινή στο 73.76%. Τέλος, η κλάση F, έχει παρόμοια τιμή recall 75.00% αλλά χαμηλότερο precision 53.62%.

Πίνακας 8: Ανάλυση ανά κατηγορία για το μοντέλο: 7. κατώφλι με φίλτρο

| Κλάση | Precision | Recall | F1-score |
|-------|-----------|--------|----------|
| N     | 0.9874    | 0.9681 | 0.9777   |
| S     | 0.7422    | 0.7314 | 0.7367   |
| V     | 0.7974    | 0.9483 | 0.8663   |
| F     | 0.5362    | 0.7500 | 0.6253   |
| Q     | 0.9843    | 0.9947 | 0.9895   |

## 8.9 CNN μοντέλο με βάρη, κατώφλι χωρίς φίλτρο

Σε αυτήν την εκδοχή πραγματοποιείται βελτιστοποίηση βαρών των κλάσεων S και F και γίνεται χρήση εύρεσης βέλτιστης τιμής κατωφλίων προβλέψεων για τις δύο αυτές κλάσεις, ενώ δεν έχει πραγματοποιηθεί στα δεδομένα προεπεξεργασία σήματος με φίλτρο αφαίρεσης θορύβου. Όπως φαίνεται και από τον πίνακα 9 οι κλάσεις N, Q που έχουν πολύ υψηλά metrics άνω του 97.00%, ενώ η κλάση V έχει επίσης υψηλές τιμές με recall 93.53%, F1-score 90.019% και με χαμηλότερο precision 87,07%.

Όσον αφορά τις κατηγορίες S και F, η εφαρμογή βαρών και κατωφλίων οδηγεί σε ικανοποιητικές τιμές recall 72.96% για την S και 73.81% για την F, γεγονός που δείχνει την ικανότητα εντοπισμού των θετικών περιπτώσεων. Το precision είναι χαμηλότερο σε σύγκριση με τις υπόλοιπες κατηγορίες (71.53% για την S και 55.86% για την F), επηρεάζοντας αντίστοιχα και τις τιμές F1-score, οι οποίες διαμορφώνονται σε 72.24% και 63.59% αντίστοιχα.

Πίνακας 9: Ανάλυση ανά κατηγορία για το μοντέλο: 8. κατώφλι χωρίς φίλτρο

| Κλάση | Precision | Recall | F1-score |
|-------|-----------|--------|----------|
| N     | 0.9862    | 0.9762 | 0.9812   |
| S     | 0.7153    | 0.7296 | 0.7224   |
| V     | 0.8707    | 0.9353 | 0.9019   |
| F     | 0.5586    | 0.7381 | 0.6359   |
| Q     | 0.9882    | 0.9960 | 0.9921   |

## 8.10 Σύγκριση αποτελεσμάτων με και χωρίς φίλτρο

Πραγματοποιείται σύγκριση της αρχιτεκτονικής του νευρωνικού συνελικτικού δικτύου που έχει περιγραφεί για τις 8 διαφορετικές περιπτώσεις στην περίπτωση χρήσης ή όχι φίλτρου απομάκρυνσης θορύβου.

Ως προς τη σύγκριση με ή χωρίς φίλτρο αφαίρεσης θορύβου στα αρχικά simple CNN, παρατηρείται ότι δεν υπάρχουν σημαντικές διαφορές στα νούμερα των metrics και δεν υπάρχει ξεκάθαρη βελτίωση με την προσθήκη ή αφαίρεση του θορύβου, όπως φαίνεται στους πίνακες 2 και 3. Μόνη διαφορά υπάρχει στην κλάση F, όπου χωρίς φίλτρο το precision είναι 98.61% και με φίλτρο 85.22%, χωρίς φίλτρο το recall είναι 42.26% και με φίλτρο 58.33%.

Στην περίπτωση με την προσθήκη βαρών καλύτερα αποτελέσματα δίνει το μοντέλο στο οποίο έχει γίνει αφαίρεση του θορύβου, χωρίς όμως μεγάλες αριθμητικές διαφορές στα metrics, όπως φαίνεται στους πίνακες 4 και 5. Η μεγαλύτερη διαφορά στις τιμές των metrics με και χωρίς φίλτρο είναι για την κλάση S, όπου έχει λίγο υψηλότερο precision και F1- score (precision 52.67% με αφαίρεση θορύβου και 40.12% χωρίς αφαίρεση θορύβου και F1- score 64.80% με αφαίρεση θορύβου και 53.93% χωρίς αφαίρεση θορύβου).

---

Στις εκδοχές όπου πραγματοποιείται εύρεση των βέλτιστων βαρών γίνεται εστίαση στη βελτίωση του χαμηλού precision των κλάσεων S,F οι οποίες αποτελούν τη μειοψηφία. Καλύτερο φαίνεται να είναι το μοντέλο στο οποίο δεν υπάρχει φίλτρο αφαίρεσης θορύβου, αλλά και πάλι οι διαφορές είναι πολύ μικρές. Οι μεγαλύτερες αποκλίσεις τιμών metrics παρατηρούνται στην κλάση V, όπου υπάρχει το precision χωρίς φίλτρο είναι 87.07% και με φίλτρο 79.74% και F1-score 90.19% χωρίς φίλτρο και 86.63% με φίλτρο, όπως φαίνεται στους πίνακες 6 και 7.

Τέλος, στα μοντέλα όπου έχει γίνει εύρεση βέλτιστου κατωφλίου πρόβλεψης, δεν υπάρχουν σημαντικές διαφορές με και χωρίς φίλτρο και παρατηρείται ότι οι κλάσεις V,Q έχουν ίδιες τιμές metrics με τις αντίστοιχες περιπτώσεις χωρίς κατώφλι, όπως φαίνεται στους πίνακες 8 και 9.

Γενικά σε όλες τις περιπτώσεις παρατηρήθηκε πως η προσθήκη φίλτρων επέφερε μικρές διαφορές στις επιδόσεις του νευρωνικού. Βελτώνει την απόδοση ορισμένων κατηγοριών, αλλά όχι σημαντικά, ενώ υπήρχαν και περιπτώσεις όπου το μοντέλο χωρίς την εφαρμογή φίλτρου είχε καλύτερα αποτελέσματα σε ορισμένα metrics κατηγοριών, χωρίς όμως μεγάλες διαφορές.

## 8.11 Σύγκριση αποτελεσμάτων των 8 πειραμάτων

Γενικά, από την ανάλυση των οκτώ διαφορετικών διαμορφώσεων του μοντέλου, το υψηλότερο AUC (94.61%) καταγράφεται στο μοντέλο όπου έχουν προστεθεί προσαρμοσμένα βάρη χωρίς περαιτέρω βελτιστοποίησή τους. Η απόδοση αυτή υποδηλώνει ότι η χρήση βαρών βελτιώνει την διαχωριστική ικανότητα του μοντέλου, ιδιαίτερα σε επίπεδο συνολικής καμπύλης ROC. Ωστόσο, στην εν λόγω εκδοχή το precision των μειοψηφικών κατηγοριών S και F παραμένει εξαιρετικά χαμηλό, γεγονός που φανερώνει αυξημένο αριθμό ψευδώς θετικών προβλέψεων. Ειδικά για την κατηγορία F, το precision ανέρχεται μόλις σε 23.87% με εφαρμογή φίλτρου θορύβου και 24.22% χωρίς φιλτράρισμα, υποδηλώνοντας περιορισμένη αξιοπιστία στις θετικές προβλέψεις της συγκεκριμένης κατηγορίας.

Το υψηλότερο accuracy επιτυγχάνεται από το απλό CNN χωρίς χρήση βαρών ή άλλων τεχνικών ενίσχυσης, φτάνοντας το 97.94%. Παρά τη συνολική του ακρίβεια, το μοντέλο αυτό καταγράφει τις χαμηλότερες τιμές AUC, με 89.31 με φίλτρο θορύβου και 87.32 χωρίς φίλτρο.

Σε όλα τα πειράματα, οι κατηγορίες N και Q παραμένουν σταθερά υψηλές σε όλες τις μετρικές, παρουσιάζοντας ελάχιστες διακυμάνσεις μεταξύ precision και recall. Το γεγονός αυτό δείχνει ότι το μοντέλο διατηρεί σταθερά διακριτική ικανότητα σε αυτές τις κατηγορίες, ανεξαρτήτως του αν εφαρμόζονται τεχνικές ενίσχυσης ή όχι.

Γενικά, στόχος ήταν η αύξηση του recall χωρίς όμως να προκληθεί σημαντική πτώση στο precision. Επίσης στα μοντέλα με βέλτιστα κατώφλια υπάρχουν ήδη και χρησιμοποιούνται τα βέλτιστα βάρη που βρέθηκαν και με τα οποία αξιολογήθηκαν οι περιπτώσεις καθώς στό-

---

χος ήταν η σταδιακή βελτίωση και ισορροπία στις αποδόσεις των προβλέψεων ανά κλάση αλλά και συνολικά.

## 8.12 Σύγκριση αποτελεσμάτων με άλλες έρευνες

Σε αυτό το κομμάτι πραγματοποιείται σύγκριση των αποτελεσμάτων με άλλες έρευνες.

Οι έρευνες που χρησιμοποιούνται είναι των Ahmad et al. [6] όπου χρησιμοποιούνται δύο σχήματα συνδυασμού πληροφοριών (fusion frameworks): το Multimodal Image Fusion (MIF) και το Multimodal Feature Fusion (MFF), των Zhou et al. [7] εφαρμόζει πολυτροπική απεικόνιση (GAF, RP, MTF) σε συνδυασμό με Residual CNN ενισχυμένο με μηχανισμό Frequency Channel Attention (FCA) και των Zhang et al. [8] που προτείνει μια αυτοεπιβλεπόμενη προσέγγιση (SMC), συνδυάζοντας τροποποιημένο MAE με Gaussian Mixture Models για ομαδοποίηση.

Στον πίνακα 10 υπάρχουν τα αποτελέσματα των metrics των ερευνών αυτών ενώ από τα συνολικά οκτώ πειράματα που πραγματοποιήθηκαν στο πλαίσιο της παρούσας εργασίας, στον συγκριτικό πίνακα περιλαμβάνονται τέσσερα. Συγκεκριμένα, για κάθε ζεύγος παραλλαγών (με ή χωρίς φίλτρο θορύβου), επιλέχθηκε να παρουσιαστεί μόνο εκείνη η εκδοχή που εμφάνισε καλύτερη συνολική απόδοση ως προς τις βασικές μετρικές (accuracy, precision, recall), οι οποίες χρησιμοποιούνται από τις άλλες έρευνες. Δηλαδή, τα μοντέλα αυτής της εργασίας που παρουσιάζονται δεν είναι τα μοντέλα που αξιολογήθηκαν και χαρακτηρίζονται γενικώς ως καλύτερα αλλά αυτά τα οποία είχαν μεγαλύτερα γενικά metrics. Το μοντέλο CNN χωρίς φίλτρο για παράδειγμα, ενώ έχει υψηλή τιμή recall, η τιμή του εξειδικευμένα για τις κλάσεις S, F είναι χαμηλότερη του 70.00%.

Στον πίνακα 10, παρατηρείται αρχικά, ότι τα μοντέλα MFF και CNN + FCA εμφανίζουν τις υψηλότερες επιδόσεις σε accuracy 99.7% και 99.6% αντίστοιχα, ενώ το μοντέλο CNN με φίλτρο θορύβου παρουσιάζει επίσης υψηλό accuracy 97.94% και υψηλό precision 95.02%. Γενικά το accuracy όλων των μοντέλων είναι ικανοποιητικό καθώς είναι μεγαλύτερο ή ίσο με 93.0%. Ως προς το precision το μοντέλο MFF παρουσιάζει το μεγαλύτερο με τιμή 98.00, αμέσως μετά είναι το μοντέλο CNN με φίλτρο θορύβου με precision 95.00% και μετά το μοντέλο MIF με 93.00%. Τα υπόλοιπα μοντέλα έχουν precision μικρότερο του 90.00% με χαμηλότερο το μοντέλο CNN με optuna χωρίς φίλτρο θορύβου με 70.42% και το μοντέλο CNN: Με χρήση GAF, SMOTE, denoising με τιμή 74.2%. Μεγαλύτερο recall και ίσο με 98.00% έχουν τα μοντέλα MFF και CNN: Με χρήση GAF, SMOTE, denoising. Αμέσως μετά είναι τα MIF (92.00%), CNN με βάρη και φίλτρο θορύβου (91.38%), CNN με optuna χωρίς φίλτρο θορύβου (91.11%) και AlexNet: Με χρήση GAF, SMOTE, denoising(91.00%) με κοντινές τιμές.

Από τα 8 μοντέλα των πειραμάτων της συγκεκριμένης εργασίας, παρόλο που ορισμένα μοντέλα, όπως το απλό CNN, παρουσιάζουν υψηλή τιμή σε μεμονωμένες μετρικές (π.χ. recall), η απόδοσή τους δεν είναι το ίδιο ικανοποιητική όταν εξετάζονται αναλυτικά οι κατηγορίες.

Συγκεκριμένα, τόσο το F1-score όσο και το precision για τις λιγότερο εκπροσωπημένες κατηγορίες εμφανίζονται αισθητά χαμηλότερες σε σχέση με πιο ενισχυμένα μοντέλα, όπως αυτά που εκπαιδεύτηκαν με χρήση βαρών loss function, ρυθμισμένων υπερπαραμέτρων μέσω Optuna ή με κατώφλια εμπιστοσύνης στις προβλέψεις. Τέλος, να αναφερθεί ότι τα τέσσερα πρώτα μοντέλα του συγκριτικού πίνακα (AlexNet, MIF, MFF, CNN) προέρχονται από την ίδια εργασία, όπου παρουσιάζονται διαφορετικές παραλλαγές της αρχιτεκτονικής που υλοποίησαν οι συγγραφείς. Τα αποτελέσματα αυτά παρατίθενται εδώ για σκοπούς σύγκρισης, καθώς βασίζονται στις ίδιες τεχνικές προεπεξεργασίας (GAF, SMOTE, denoising) που χρησιμοποιούνται και στο παρόν πείραμα.

Πίνακας 10: Αποτελέσματα ταξινόμησης ως προς τις μετρικές accuracy, precision και recall

| Μοντέλο / Πείραμα                                         | Accuracy | Precision | Recall |
|-----------------------------------------------------------|----------|-----------|--------|
| AlexNet: Με χρήση GAF, SMOTE, denoising [6]               | 97.3     | 85.00     | 91.00  |
| MIF: Με χρήση 3-καναλων εικόνων, SMOTE, denoising [6]     | 98.6     | 93.00     | 92.00  |
| MFF: Με χρήση 3-καναλων εικόνων, SMOTE, denoising [6]     | 99.7     | 98.00     | 98.00  |
| CNN: Με χρήση GAF, SMOTE, denoising [6]                   | 94.2     | 74.20     | 98.00  |
| CNN + FCA: Με χρήση φίλτρων, SMOTE, 3-καναλων εικόνων [7] | 99.6     | –         | –      |
| MAE + Clustering: Με χρήση GAF [8]                        | 97.84    | 84.25     | 65.51  |
| CNN με φίλτρο θορύβου (εργασίας)                          | 97.94    | 95.02     | 82.90  |
| CNN με βάρη και φίλτρο θορύβου (παρούσας εργασίας)        | 94.17    | 72.37     | 91.38  |
| CNN με optuna χωρίς φίλτρο θορύβου (παρούσας εργασίας)    | 93.10    | 70.42     | 91.11  |
| CNN με κατώφλι χωρίς φίλτρο θορύβου (παρούσας εργασίας)   | 96.55    | 82.38     | 87.50  |

## 9 Συζήτηση

Όπως φαίνεται από τα αποτελέσματα, οι αποδόσεις των μοντέλων CNN, χωρίς βάρη και χρήση κατωφλίων, αναδεικνύουν την επίδραση που έχει η ανισορροπία κλάσεων στα δεδομένα, κυρίως στα recalls των μειοψηφικών κλάσεων και την ανάγκη αντιμετώπισης της ανισορροπίας στα δεδομένα.

Στα μοντέλα που εφαρμόζονται προσαρμοσμένα βάρη, αυξάνεται το recall των κατηγοριών S και F, αλλά ταυτόχρονα μειώνεται το precision καθώς το μοντέλο ωθείται να αποδώσει μεγαλύτερη σημασία στην ορθή ταξινόμηση των κλάσεων αυτών, πραγματοποιώντας έτσι περισσότερες προβλέψεις για τις κλάσεις αυτές. Αυτό οδηγεί σε πτώση του precision, αφού αυτή η στρατηγική αυξάνει ταυτόχρονα και τον αριθμό των false positives.

Ομοίως, στην περίπτωση χρήσης optuna φαίνεται ότι η προσθήκη βέλτιστων βαρών, οδήγησε στην εστίαση του μοντέλου στο recall των μειοψηφικών κλάσεων με αποτέλεσμα το μοντέλο να πραγματοποιεί περισσότερες προβλέψεις για τις κλάσεις αυτές, αυξάνοντας το

---

recall και μειώνοντας το precision. Επιπλέον, το F1-score αποτελεί τον αρμονικό μέσο του precision και του recall, οπότε είναι αναμενόμενο οι χαμηλές τιμές στο precision να επηρεάζουν και τα F1-scores των αντίστοιχων κλάσεων. Συνεπώς, εδώ φαίνεται πως ακόμη και όταν το recall είναι υψηλό, μια σημαντικά μειωμένη τιμή precision οδηγεί σε μέτρια ή χαμηλή τελική τιμή F1-score, όπως παρατηρείται χαρακτηριστικά στις κατηγορίες S και F.

Όσο αφορά τη χρήση ρύθμισης κατωφλίων πρόβλεψης (threshold tuning), φαίνεται να επιτρέπει τον έλεγχο της ευαισθησίας του μοντέλου στις προβλέψεις των επιμέρους κατηγοριών, επιδιώκοντας καλύτερη ισορροπία μεταξύ precision και recall. Η χαμηλή τιμή precision της κλάσης F μπορεί να οφείλεται σε περισσότερες προβλέψεις αυτής της κατηγορίας με στόχο την ενίσχυση της ευαισθησίας της, εις βάρος της ακρίβειας ή σε μορφολογική ομοιότητα παλμών F με άλλες κατηγορίες, είτε σε μειωμένη ποιότητα του αρχικού σήματος σε χτύπους αυτής της κατηγορίας, αφού δεν εφαρμόστηκε φίλτρο αφαίρεσης θορύβου για βελτίωση της ποιότητας του αρχικού σήματος. Η συνολική εικόνα υποδεικνύει ότι ο συνδυασμός κατάλληλων κατωφλίων πρόβλεψης και βελτιστοποιημένων βαρών ακόμη και χωρίς καθαρισμό σήματος, μπορεί να επηρεάσει σημαντικά τις επιδόσεις των προβλέψεων του μοντέλου, ειδικά για τις λιγότερο αντιπροσωπευμένες κλάσεις.

Συνολικά, τα αποτελέσματα δείχνουν ότι το μοντέλο είναι ιδιαίτερα σταθερό στην ταξινόμηση των πλειοψηφικών κατηγοριών, ενώ η απόδοσή του στις μειοψηφικές εξαρτάται από τις τεχνικές ενίσχυσης. Οι στρατηγικές όπως η χρήση βαρών, τα κατώφλια και ο καθαρισμός του σήματος από θόρυβο βελτιώνουν την απόδοσή του.

Επίσης, προκύπτει ότι η εφαρμογή φίλτρου αφαίρεσης θορύβου στα δεδομένα δεν φαίνεται να επιφέρει μεγάλες διαφορές στις επιδόσεις του μοντέλου. Ιδιαίτερα οι πλειοψηφικές κατηγορίες N και Q, παρουσιάζουν σταθερά υψηλά αποτελέσματα ανεξαρτήτως αφαίρεσης θορύβου. Το γεγονός αυτό υποδηλώνει ότι οι συγκεκριμένες κλάσεις διαθέτουν επαρκή και διακριτά μορφολογικά χαρακτηριστικά, που επιτρέπουν στο μοντέλο να τις αναγνωρίζει.

Ως προς τη χρήση GAF, όπως φαίνεται και στις ενδεικτικές εικόνες 1, 2, 3, 4 και 5, η εφαρμογή φίλτρου αποθορυβοποίησης δεν επιφέρει σημαντικές οπτικές διαφοροποιήσεις. Οι βασικές διαφορές εντοπίζονται κυρίως στην ένταση των χρωμάτων και όχι στη μορφή των μοτίβων, πιθανώς γιατί ο μετασχηματισμός GAF διατηρεί τη συνολική μορφολογία του σήματος και το φίλτρο δεν επηρεάζει τη μορφολογία αλλά ορισμένες ακραίες τιμές.

Τέλος, σημειώνεται ότι υπάρχουν ορισμένοι περιορισμοί στην εργασία. Παρόλο που η απόδοση του μοντέλου είναι υψηλή, δεν υπερτερεί έναντι πιο σύνθετων αρχιτεκτονικών ή πολυσταδιακών προσεγγίσεων που ενσωματώνουν μηχανισμούς ενίσχυσης ή συνδυασμούς χαρακτηριστικών. Ωστόσο, η απλότητα του μοντέλου προσφέρει το πλεονέκτημα της χαμηλότερης υπολογιστικής πολυπλοκότητας και της ευκολότερης αναπαραγωγής των αποτελεσμάτων.

---

## 10 Συμπεράσματα και μελλοντικές έρευνες

Στην παρούσα εργασία μελετήθηκε το πρόβλημα της ταξινόμησης καρδιακών παλμών σε πέντε κατηγορίες, σύμφωνα με το πρότυπο AAMI, με τη χρήση συνελκτικού νευρωνικού δικτύου. Για τον σκοπό αυτό, τα σήματα ΗΚΓ μετασχηματίστηκαν σε εικόνες μέσω του Gramian Angular Field (GAF), ώστε να καταστούν κατάλληλα για είσοδο σε μοντέλα βασισμένα σε εικόνες. Τα αποτελέσματα της μελέτης κατέδειξαν ότι, ακόμη και χωρίς την εφαρμογή τεχνικών ενίσχυσης των υποεκπροσωπούμενων κατηγοριών, όπως το data augmentation ή το SMOTE, το μοντέλο παρουσίασε ικανοποιητική απόδοση. Συγκεκριμένα, επιτεύχθηκε μέγιστη ακρίβεια (accuracy) 0.9789 και μέγιστο F1-score 0.9111, ενώ ιδιαίτερη έμφαση δόθηκε στη βελτίωση των τιμών recall για τις κατηγορίες με αρχικά χαμηλή επίδοση. Σε σύγκριση με προηγούμενες μελέτες, η παρούσα εργασία διαφοροποιείται καθώς βασίζεται αποκλειστικά στον GAF μετασχηματισμό και σε ένα σχετικά απλό συνελκτικό νευρωνικό δίκτυο, αποφεύγοντας την παραγωγή τεχνητών δεδομένων. Παρά την απλότητα της αρχιτεκτονικής και την απουσία τεχνικών εξισορρόπησης, τα αποτελέσματα είναι ανταγωνιστικά, υπογραμμίζοντας τη χρηστικότητα της μεθόδου. Για μελλοντικές εργασίες, προτείνεται η διερεύνηση εναλλακτικών μετασχηματισμών εικόνας, όπως το Recurrence Plot (RP) και το Markov Transition Field (MTF), καθώς και η εφαρμογή τεχνικών συνδυασμού (fusion) διαφορετικών χαρακτηριστικών ή μετασχηματισμών, με στόχο την περαιτέρω ενίσχυση της ακρίβειας και της γενίκευσης του μοντέλου, καθώς και ο προσδιορισμός κατάλληλου μήκους segment.

---

## Βιβλιογραφία

### Αναφορές

- [1] M. Zoni-Berisso, F. Lercari, T. Carazza, and S. Domenicucci, “Cardiac arrhythmias: diagnosis, symptoms, and management,” *Clinical interventions in aging*, vol. 11, p. 655, 2016.
- [2] American Heart Association, “Types of arrhythmias,” 2023. Accessed May 2025.
- [3] World Health Organization, “Cardiovascular diseases (cvds),” [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)), 2021. Accessed May 2025.
- [4] J. E. Hall, *Guyton and Hall Textbook of Medical Physiology*. Philadelphia, PA: Elsevier, 14 ed., 2020.
- [5] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, “Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network,” *Nature medicine*, vol. 25, no. 1, pp. 65–69, 2019.
- [6] Z. Ahmad, A. Tabassum, L. Guan, and N. M. Khan, “Ecg heartbeat classification using multimodal fusion,” *IEEE Access*, vol. 9, pp. 100615–100626, 2021.
- [7] F. Zhou and D. Fang, “Multimodal ECG heartbeat classification method based on a convolutional neural network embedded with FCA,” *Scientific Reports*, vol. 14, p. 8804, 2024.
- [8] Y. Zhang, X. Li, L. Zhang, J. Wang, S. Jiang, Y. Ma, and D. Li, “A sequential mae-clustering self-supervised learning method for arrhythmia detection,” *Expert Systems with Applications*, vol. 269, p. 126379, 2025.
- [9] E. Izci, M. A. Ozdemir, M. Degirmenci, and A. Akan, “Cardiac arrhythmia detection from 2d ecg images by using deep learning technique,” in *2019 Medical Technologies Congress (TIPTEKNO)*, pp. 1–4, 2019.
- [10] M. Zhang, H. Jin, B. Zheng, and W. Luo, “Deep learning modeling of cardiac arrhythmia classification on information feature fusion image with attention mechanism,” *Entropy*, vol. 25, no. 9, p. 1264, 2023.
- [11] J. Qin, F. Gao, Z. Wang, L. Liu, and C. Ji, “Arrhythmia detection based on wgan-gp and se-resnet1d,” *Electronics*, vol. 11, no. 21, p. 3427, 2022.

- 
- [12] Cleveland Clinic, “Heart.” <https://my.clevelandclinic.org/health/body/21704-heart>, n.d. Accessed: 2025-05-25.
- [13] “Characteristics of the normal ecg (p-wave, qrs complex, st segment, t-wave, j-point).” <https://ecgwaves.com/topic/ecg-normal-p-wave-qrs-complex-st-segment-t-wave-j-point/>. Accessed: 2025-05-25.
- [14] E. A. Ashley and J. Niebauer, *Cardiology Explained*. London: Remedica, 2004. Conquering the ECG. Available at <https://www.ncbi.nlm.nih.gov/books/NBK2214/>.
- [15] G. Lannoy, D. François, J. Delbeke, and M. Verleysen, “Feature relevance assessment in automatic inter-patient heart beat classification,” in *Proceedings of the 3rd International Workshop on Biomedical Data Engineering (BMDE)*, pp. 13–20, 2010.
- [16] M. K. Das and S. Ari, “Ecg beats classification using mixture of features,” *International Scholarly Research Notices*, vol. 2014, p. 178436, 2014.
- [17] E. Burns and R. Buttner, “Premature atrial complex (pac) – ecg library diagnosis.” <https://litfl.com/premature-atrial-complex-pac/>, 2024. Life in the Fast Lane (LITFL), ανακτήθηκε τον Μάιο 2025.
- [18] E. Burns and R. Buttner, “Premature ventricular complex (pvc) – ecg library diagnosis.” <https://litfl.com/premature-ventricular-complex-pvc-ecg-library/>, 2024. Life in the Fast Lane (LITFL), accessed May 2025.
- [19] E. Burns and R. Buttner, “Fusion beats – ecg library diagnosis.” <https://litfl.com/fusion-beat-dressler-beat-ecg-library/>, 2024. Life in the Fast Lane (LITFL), accessed May 2025.
- [20] Z. Wang, T. Oates, *et al.*, “Encoding time series as images for visual inspection and classification using tiled convolutional neural networks,” in *Workshops at the twenty-ninth AAAI conference on artificial intelligence*, vol. 1, pp. 1–7, Austin, TX, 2015.
- [21] C.-L. Yang, Z.-X. Chen, and C.-Y. Yang, “Sensor classification using convolutional neural network by encoding multivariate time series as two-dimensional colored images,” *Sensors*, vol. 20, no. 1, p. 168, 2019.
- [22] A. H. Ribeiro *et al.*, “Automatic diagnosis of the 12-lead ecg using a deep neural network,” *Nature Communications*, vol. 11, no. 1, p. 1760, 2020.
- [23] P. Nejedly *et al.*, “Classification of ecg using ensemble of residual cnns with attention mechanism,” in *Computing in Cardiology*, 2021.

- 
- [24] M. Pesl *et al.*, “Cardiac arrhythmias classification in kardiovize population study,” in *Computing in Cardiology*, 2021.
  - [25] S. Llamedo and J. P. Martínez, “Heartbeat classification using feature selection driven by database generalization criteria,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 616–625, 2011.
  - [26] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
  - [27] Association for the Advancement of Medical Instrumentation, *Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms*. Arlington, VA: AAMI, 1998. ANSI/AAMI EC57:1998.
  - [28] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.
  - [29] A. Y. Ng, “Feature selection, l1 vs. l2 regularization, and rotational invariance,” in *Proceedings of the Twenty-First International Conference on Machine Learning (ICML)*, p. 78, ACM, 2004.
  - [30] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
  - [31] S. Min *et al.*, “Bag of tricks for electrocardiogram classification with deep neural networks,” in *Computing in Cardiology*, 2020.
  - [32] Y. Ou, X. Li, Z. Guo, and Y. Wang, “Anobeat: Anomaly detection for electrocardiography beat signals,” in *2020 IEEE Fifth International Conference on Data Science in Cyberspace (DSC)*, pp. 142–149, 2020.
  - [33] C. fang Zhao, W. yun Yao, M. juan Yi, C. Wan, and Y. le Tian, “Arrhythmia classification algorithm based on a two-dimensional image and modified efficientnet,” *Computational Intelligence and Neuroscience*, vol. 2022, p. 8683855, 2022. Special Issue: Advanced Deep Learning and Neuro-Evolution Metaheuristic Techniques in Medical Applications.
  - [34] F. Lüer, D. Mautz, and C. Böhm, “Anomaly detection in time series using generative adversarial networks,” in *2019 International Conference on Data Mining Workshops (ICDMW)*, pp. 1047–1048, 2019.

- 
- [35] A. Rajkumar, M. Ganesan, and R. Lavanya, "Arrhythmia classification on ecg using deep learning," in *2019 5th International Conference on Advanced Computing Communication Systems (ICACCS)*, pp. 365–369, 2019.
- [36] E. Izci, M. A. Ozdemir, M. Degirmenci, and A. Akan, "Cardiac arrhythmia detection from 2d ecg images by using deep learning technique," in *2019 Medical Technologies Congress (TIPTEKNO)*, pp. 1–4, 2019.
- [37] G. B. Moody and R. G. Mark, "The mit-bih arrhythmia database." <https://www.physionet.org/content/mitdb/1.0.0/>, 2001. PhysioNet: PhysioBank Archive, Accessed: May 25, 2025.