



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τεχνολογία Ήχου Και Εικόνας Κατηγοριοποίηση Καρδιακών Ήχων

9^ο Εξάμηνο

Μουστάκας Γεώργιος 9365
Σαρρής Αναστάσιος Λουκάς 9451
Στεφανίδης Ιωάννης 9587
Σφυράκης Εμμανουήλ 9507

22 Φεβρουαρίου 2022

Περιεχόμενα

1	Εισαγωγή	2
2	Εξαγωγή χαρακτηριστικών	2
2.1	Mel Frequency Cepstral Coefficients	2
2.2	Εξαγωγή στην Python	2

1 Εισαγωγή

Εισαγωγή

2 Εξαγωγή χαρακτηριστικών

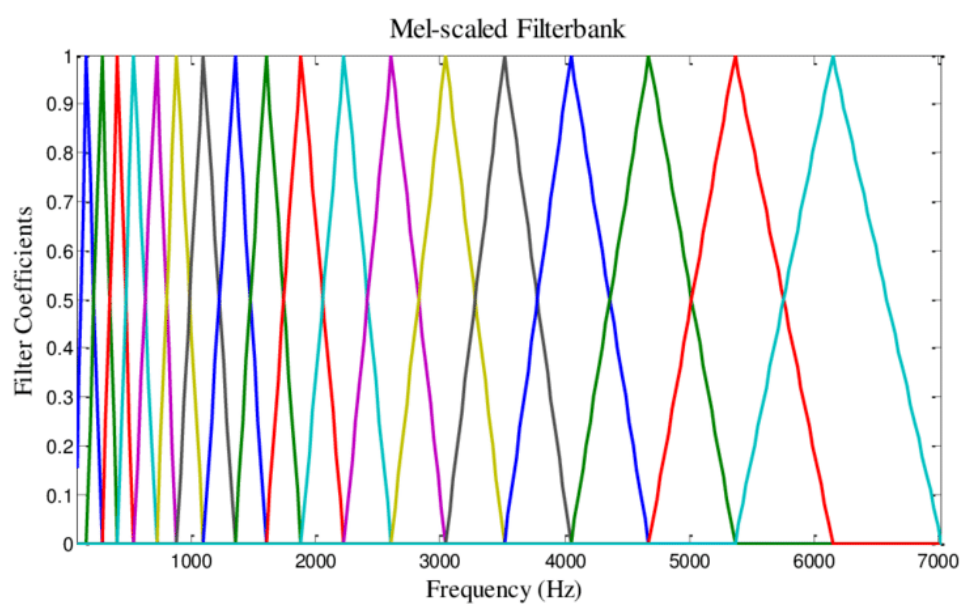
2.1 Mel Frequency Cepstral Coefficients

Ένα από τα χαρακτηριστικά που εξάγαμε από τα σήματα καρδιακών ήχων που είχαμε στη διάθεσή μας ώστε να εκπαιδεύσουμε το νευρωνικό δίκτυο είναι τα Mel Frequency Cepstral Coefficients. Ο υπολογισμός τους έγινε στα ήδη προεπεξεργασμένα δεδομένα και συγκεκριμένα στο κάθε ένα σήμα που δημιουργήθηκε από την έξοδο των παραθύρων. Το σύνολο των συντελεστών που παρήγαγε αυτή η διαδικασία είναι 13 από τους οποίους οι 12 αναπαριστούν την περισσότερη πληροφορία της φασματικής περιβάλλουσας και ο 13^{ος} αναπαριστά τη συνολική ενέργεια του σήματος. Δεν επιλέχθηκαν περισσότεροι από 13 συντελεστές καθώς η αύξηση του αριθμού τους πάνω από αυτό το όριο έχει ως αποτέλεσμα την ταχεία μεταβολή των συντελεστών γεγονός που δυσχαιρένει την εκπαίδευση του νευρωνικού δικτύου. Ο κύριος λόγος που επιλέχθηκαν τα mfcc's είναι ότι αποτελούν την καλύτερη προσέγγιση της λειτουργίας του κοχλίου του ανθρώπινου αυτιού που είναι επιλεκτικός στις συχνότητες και στο πως αντιδρά σε αυτές. Επιγραμματικά η διαδικασία για τον υπολογισμό των mfcc είναι

- Εφαρμογή παραθύρων στο σήμα
- Υπολογισμός του φάσματος ενέργειας
- Εφαρμογή του φίλτρου Mel και άθροισμα της ενέργειας του κάθε φίλτρου
- Λογαρίθμηση του αποτελέσματος του προηγούμενου βήματος
- Εφαρμογή μετασχηματισμού συνημιτόνων

2.2 Εξαγωγή στην Python

Για τις ανάγκες υλοποίησης του νευρωνικού δικτύου ήταν απαραίτητη η εξαγωγή των mfcc's από τα φωνοκαρδιογραφήματα που είχαμε στη διάθεσή μας. Μετά την προεπεξεργασία στην οποία υποβλήθηκαν όταν πλέον είχαμε τα δείγματα μας χωρισμένα σε μικρότερα από επικαλυπτόμενα παράθυρα τότε σε κάθε ένα καινούριο ηχητικό σήμα το οποίο είχε δημιουργηθεί εφαρμόστηκε η συνάρτηση mfcc της βιβλιοθήκης `python_speech_features` η έξοδος της οποίας, τα 13 mfcc's, αποθηκεύτηκαν σε μορφή εικόνας που είναι και τα χαρακτηριστικά εκπαίδευσης του συνελκτικού νευρωνικού δικτύου.



Σχήμα 2.1: Φίλτρο Mel