

Θέμα πτυχιακής: Μελέτη τεχνικών διαχωρισμού ηχητικών πηγών με χρήση τεχνικών βαθιάς μάθησης

(Utilization of deep learning techniques for the task of sound source)

Επιβλέπων: Παναγιώτης Ζέρβας

Email: p.zervas@uop.gr

Μέλη: -

Ακαδημαϊκό έτος: 2022-23

Στόχοι: Στο πλαίσιο της προτεινόμενης πτυχιακής θα λάβουν χώρα τα παρακάτω

1. Θα μελετηθεί η βιβλιογραφία αναφορικά με το πρόβλημα του διαχωρισμού ηχητικών πηγών
2. Θα μελετηθεί η αρχιτεκτονική U-Net και θα γίνει εμβάθυνση στη παραλλαγή της Wave-U-Net για ηχητικά δεδομένα της οποία τελικά θα γίνει υλοποίηση σε Python
3. Θα μελετηθούν και υλοποιηθούν διάφορες τεχνικές αναπαράστασης σημάτων που μπορούν να τροφοδοτήσουν την είσοδο ενός Wave-U-Net
4. Θα γίνουν πειράματα υποκειμενικής και αντικειμενικής αξιολόγησης των πειραμάτων διαχωρισμού φωνητικών και άλλων μουσικών πηγών.

Αντικείμενο: Ο διαχωρισμός πηγής ήχου, γνωστός και ως το πρόβλημα του κοκτέιλ πάρτι, είναι ένα από τα μεγαλύτερα προβλήματα στον ήχο λόγω της πρακτικής του χρήσης. Με τον όρο διαχωρισμός πηγών ήχου εννοούμε τη διαδικασία διαχωρισμού ενός μείγματος (π.χ. ηχογράφηση ενός ποπ συγκροτήματος) σε απομονωμένους ήχους από μεμονωμένες πηγές (π.χ. μόνο τα φωνητικά). Στο πλαίσιο της προτεινόμενης πτυχιακής θα μελετηθεί η χρήση της αρχιτεκτονικής Wave-U-Net, μια προσαρμογή του U-Net στο μονοδιάστατο πεδίο του χρόνου, για το διαχωρισμό πηγών σε μια ηχογράφηση.

Η εργασία περιλαμβάνει:

- θεωρητική μελέτη, ανάπτυξη και ανάλυση αλγορίθμων
- σχεδιασμό και ανάπτυξη δεδομένων ομιλίας και θορύβου
- σχεδιασμό και ανάπτυξη συστήματος
- αξιολόγηση απόδοσης και σύγκριση με άλλες τεχνικές

Σχετιζόμενα μαθήματα:

Σήματα και συστήματα, Ψηφιακή επεξεργασία σημάτων, Στατιστική επεξεργασία σημάτων, Τεχνικές Προγραμματισμού

Απαιτήσεις παρουσίας φοιτητή: Ναι