

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

**Ψαρουλάκης Κωνσταντίνος
Μεταπτυχιακός Φοιτητής**

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Επόπτης Μεταπτυχιακής Εργασίας: Καθηγητής, Π. Τσακαλίδης

Τετάρτη, 7 Οκτωβρίου 2020 ,ώρα 11:00 π.μ.

**Τηλεδιάσκεψη (μέσω του συστήματος e:Presence), Τμήμα Επιστήμης Υπολογιστών,
Πανεπιστήμιο Κρήτης**

Διεύθυνση μετάδοσης (url): <http://video.ucnet.uoc.gr/live/show/316>

Κανάλι YouTube του Τμήματος

https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB_Gnt6Q/live

“Χρήση τεχνικών μηχανικής μάθησης για τον εντοπισμό ανθρωπογενούς δραστηριότητας από ηχογραφήσεις σε προστατευόμενες περιοχές”

Περίληψη

Η ανθρώπινη δραστηριότητα αποτελεί σήμερα τη σημαντικότερη αιτία εξαφάνισης βιοτόπων μεγάλου αριθμού φυτών και ζώων στον πλανήτη. Αυτή η δραστηριότητα οδηγεί αφενός στην ολική εξαφάνιση πολλών ειδών κι αφετέρου στη αποδυνάμωση των οικοσυστημάτων, γεγονός που διαταράσσει τις ισορροπίες στον πλανήτη και την ποιότητα ζωής του ανθρώπου. Ένα μέτρο προστασίας είναι η θέσπιση προστατευόμενων περιοχών όπου η ανθρώπινη δραστηριότητα είναι περιορισμένη. Σε αυτές τις περιοχές, συστήματα ασφαλείας αποτελούμενα από πολλές κάμερες και μικρόφωνα μπορούν να

συμβάλλουν σημαντικά στην παρακολούθηση της ισορροπίας του εκάστοτε οικοσυστήματος καθώς και στην αποτροπή ανθρώπινης δραστηριότητας που αποτελεί απειλή για το περιβάλλον.

Αυτή η εργασία επικεντρώνεται στην εφαρμογή τεχνικών μηχανικής μάθησης και επεξεργασίας σήματος σε δεδομένα ήχου από προστατευόμενες περιοχές ανά την Ελλάδα, με σκοπό τον αυτόματο εντοπισμό ηχητικών γεγονότων που σηματοδοτούν παράνομη ανθρώπινη δραστηριότητα, όπως παράνομη υλοτομία, βόσκηση, κυνήγι κ.α.

Για τη συλλογή και κατηγοριοποίηση των δεδομένων που απαιτούνται για την εκπαίδευση των μοντέλων μηχανικής μάθησης, παρουσιάζουμε τη χρησιμότητα μιας μεθόδου εντοπισμού φωνής που ενεργοποιείται με την ύπαρξη αρμονικής δομής σε ένα σήμα ήχου. Η μέθοδος αξιοποιείται για την αυτόματη κατάτμηση εκατοντάδων ωρών ηχογράφησης σε χιλιάδες μικρής διάρκειας ηχητικά αποσπάσματα που εν δυνάμει φέρουν το υποκείμενο μοτίβο ενδιαφέροντος.

Στη συνέχεια, εκτελούμε πολλαπλά πειράματα με στόχο την εύρεση της βέλτιστης προσέγγισης για (i) ένα πρόβλημα δυαδικής ταξινόμησης που επικεντρώνεται την περίπτωση του ήχου του αλυσοπρίονου και (ii) ένα πρόβλημα έξι κλάσεων που περιλαμβάνει πρόσθετα ηχητικά μοτίβα σχετικά με την παράνομη ανθρώπινη δραστηριότητα. Τα αποτελέσματα που παρουσιάζουμε αναδεικνύουν την υπεροχή των Βαθιών Νευρωνικών Δικτύων έναντι γνωστών συμβατικών προσεγγίσεων και αναδεικνύουν βέλτιστες επιλογές όσον αφορά τις αρχιτεκτονικές των δικτύων και τον τύπο των ακουστικών χαρακτηριστικών ανάλογα με το πρόβλημα ταξινόμησης που εξετάζεται.

University of Crete

Computer Science Department

M.Sc. Thesis presentation / examination

Psaroulakis Konstantinos

Master's Thesis Supervisor: Professor P. Tsakalides

Wednesday, 7 October 2020, 11:00 a.m

**Teleconference (will use the e: Presence system), Computer Science Department,
University of Crete**

(url) : <http://video.ucnet.uoc.gr/live/show/316>

YouTube channel :

https://www.youtube.com/channel/UC7uE3QiMTQjkrpByB_Gnt6Q/live

“Machine learning techniques for the detection of illegal human activity in audio recordings from protected areas”

Abstract

Human activity is considered today as the primary reason for habitat loss for a large number of Earth's plant and animal species. This activity results to the permanent loss of species and to the weakening of the ecosystems that are of significant importance for the overall health of the planet and as a consequence, to the quality of the human life. One key measure to protect habitats is the establishment of protected areas where human activity is restricted. In these areas, systems employing multiple cameras and microphones may offer a significant assistance in monitoring the health of the ecosystem but also as the means to prevent human intervention that is harmful to the environment.

This Thesis concerns the application of signal processing and machine learning techniques to audio recordings acquired in protected areas in Greece, with the aim to automatically detect sound events that are indicative of illegal human activity such as illegal logging, grazing, hunting, etc. To collect and annotate the data that is required for training such a scheme, we illustrate the usefulness of a Voice Activity Detector (VAD) that is activated on the presence of harmonic structure in the audio content. The VAD is used in order to automatically segment hundreds of hours of audio recording into thousands of short duration audio clips that potentially carry the underlying pattern of interest.

Continuing, we perform numerous experiments with the goal to find the optimal approach for (i) a binary classification problem that focuses on the case of chainsaw sound and (ii) a six class problem that includes additional patterns relating to illegal human activity. Experimental results illustrate the superiority of Deep Neural Networks (DNN) against other well known conventional classifiers and furthermore, highlight choices that are advantageous for the intended task in terms of the DNN architecture and the type of acoustic features.