

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

**Λυδάκης Γεώργιος
Μεταπτυχιακός Φοιτητής**

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Επόπτης Μεταπτυχιακής Εργασίας: Καθηγητής, Α. Αργυρός

Πέμπτη, 8 Ιουλίου 2021, ώρα 15:00 μ.μ.

Join Zoom Meeting

<https://zoom.us/j/95221129501>

“ Εκμετάλλευση αυτόματης επισημείωσης δεδομένων για εκπαίδευση Συνελικτικών Νευρωνικών Δικτύων και εφαρμογή στην κατηγοριοποίηση χειρομορφών”

Περίληψη

Τα τελευταία χρόνια, η πρόοδος σε τεχνικές βαθιάς μάθησης έχει επιφέρει μια επανάσταση μεγάλης κλίμακας στο πεδίο της Τεχνητής Νοημοσύνης. Μέθοδοι βαθιάς μάθησης έχουν εφαρμοστεί επιτυχώς σε μια πληθώρα ερευνητικών τομέων, από επεξεργασία φυσικών γλωσσών και βιοπληροφορική μέχρι αναγνώριση ομιλίας και υπολογιστική όραση, με κοινό στόχο την αυτόματη εκτίμηση μιας συνάρτησης που απεικονίζει ένα πεδίο εισόδου σε ένα πεδίο επιθυμητού αποτελέσματος. Ωστόσο, η επιτυχία των μεθόδων αυτών στο να εξάγουν μια τέτοια συνάρτηση συνήθως εξαρτάται από την ύπαρξη ενός μεγάλου όγκου επισημειωμένων δεδομένων εκπαίδευσης, δηλαδή δειγμάτων εισόδου για τα οποία η έξοδος είναι καθορισμένη. Λόγω της απαίτησης για μεγάλα σύνολα τέτοιων δεδομένων, σημαντική έρευνα διεξάγεται πάνω σε μεθόδους

που μειώνουν το κόστος σε ανθρώπινη προσπάθεια που απαιτείται για την επισημείωση αυτών των δεδομένων.

Προσεγγίσεις ημιεπίβλεψης, μέθοδοι δημιουργίας συνθετικών δεδομένων και τεχνικές για δημιουργία και χειρισμό αυτόματης επισημείωσης συγκεντρώνουν αυξανόμενο ενδιαφέρον. Σε αυτή την εργασία, διερευνούμε μια τεχνική για αξιοποίηση αυτόματα επισημειωμένων δεδομένων σε προβλήματα κατηγοριοποίησης. Χρησιμοποιώντας έναν μικρό αριθμό δεδομένων που έχουν επισημειωθεί από κάποιο ειδικό, και ένα μεγάλο σύνολο δεδομένων που χαρακτηρίζονται από αυτόματα εκτιμημένες, θορυβώδεις ετικέτες, η προσέγγισή μας εκπαιδεύει ένα Συνελικτικό Νευρωνικό Δίκτυο (CNN) με επαναληπτικό τρόπο. Οι αυτόματες επισημειώσεις συνδυάζονται με τις προβλέψεις του δικτύου ώστε να επεκταθεί σταδιακά το σύνολο δεδομένων εκπαίδευσης. Αυτή η επέκταση επιχειρεί να επιλέξει αυτόματα επισημειωμένα δείγματα των οποίων η ετικέτα κρίνεται σωστή.

Η προτεινόμενη προσέγγιση είναι γενική και μπορεί να εφαρμοστεί σε οποιοδήποτε πρόβλημα κατηγοριοποίησης. Προκειμένου να αποτιμήσουμε την απόδοσή της, την εφαρμόζουμε στο πρόβλημα της αναγνώρισης χειρομορφών από έγχρωμες (RGB) εικόνες εισόδου. Γενικά, η παρατήρηση και η ερμηνεία του ανθρώπινου χεριού είναι πολύ χρήσιμη σε ποικίλες εφαρμογές, οπότε και έχει αναπτυχθεί σημαντική έρευνα στα θέματα της τρισδιάστατης παρακολούθησης του χεριού, της παρατήρησης των αλληλεπιδράσεων του χεριού με αντικείμενα, καθώς και της αναγνώρισης χειρομορφών και χειρονομιών. Η Αναγνώριση Νοηματικής Γλώσσας είναι μια περιοχή όπου η αναγνώριση χειρομορφών είναι ιδιαίτερα χρήσιμη, επειδή οι μορφές των χεριών ενός νοηματιστή είναι κρίσιμα χαρακτηριστικά για την μετάφραση μιας νοηματικής γλώσσας.

Παρακινημένοι από την χρησιμότητα και τον αντίκτυπο της Αναγνώρισης Νοηματικής Γλώσσας, αναπτύσσουμε μια μέθοδο για αυτόματη επισημείωση εικόνων ή βίντεο χειρομορφών, και την εφαρμόζουμε στο πρόβλημα της κατηγοριοποίησης 19 χειρομορφών κοινών στην Ελληνική Νοηματική Γλώσσα. Η χειροκίνητη επισημείωση τέτοιων δεδομένων είναι μια χρονοβόρα διαδικασία. Η αυτόματη επισημείωσή τους βασίζεται στο συσχετισμό τρισδιάστατων αναπαραστάσεων των χεριών με τις κλάσεις (ετικέτες χειρομορφών), και δημιουργεί θορυβώδεις ετικέτες που μπορούν να χρησιμοποιηθούν για την εκπαίδευση Συνελικτικών Νευρωνικών Δικτύων. Αυτά τα χαρακτηριστικά του προβλήματος το καθιστούν υποψήφιο για την εφαρμογή της μεθόδου που προτείνουμε για την αυτόματη επέκταση συνόλων δεδομένων εκπαίδευσης. Συγκρίνουμε τα αποτελέσματα εκπαίδευσης ενός CNN με και χωρίς τη χρήση της τεχνικής μας. Η μεθοδός μας επιφέρει σημαντική αύξηση στην μέση ακρίβεια κατηγοριοποίησης, και επιπλέον μειώνει την απόκλιση της ακρίβειας ανά κλάση,

καταδεικνύοντας έτσι την εγκυρότητα και τη χρησιμότητα της προτεινόμενης προσέγγισης.

University of Crete

Computer Science Department

M.Sc. Thesis presentation / examination

Lydakis Georgios

Master's Thesis Supervisor: Professor, A. Argyros

Thursday, 8 July 2021, 15:00 p.m.

Join Zoom Meeting

<https://zoom.us/j/95221129501>

“Exploitation of noisy automatic data annotation for CNN training and its application to hand posture classification”

Abstract

In recent years, advances in deep learning have resulted in a large-scale revolution in the field of Artificial Intelligence. Deep learning methods have been successfully applied to a variety of research topics, ranging from natural language processing and bioinformatics to speech recognition and computer vision, with the common goal of inferring a function which maps an input domain to a target one. However, the success of deep models at inferring such a function usually relies on the existence of a large amount of annotated training data, that is, input samples for which the output is specified. Due to the requirement for large such training datasets, significant research is being conducted on methods for reducing the human effort that is necessary for their annotation.

Semi-supervised approaches, methods for synthetic data generation, and techniques for generating and handling automatic annotation are receiving increasing attention. In this work, we investigate a technique for utilizing automatically annotated data in classification problems. Using a small number of manually annotated samples, and a large set of data that feature automatically created, noisy labels, our approach trains a

Convolutional Neural Network (CNN) in an iterative manner. The automatic annotations are combined with the predictions of the network in order to gradually expand the training set. This expansion attempts to select automatically annotated samples for which the label is deemed to be correct.

The proposed approach is generic and can be applied to any classification problem. In order to evaluate its performance, we apply it to the problem of hand posture recognition from RGB images. In general, the observation and interpretation of the human hand is highly useful in several application areas, so significant research has been carried out in the topics of 3-D hand tracking, observation of interaction of hands with objects as well as hand posture and gesture recognition. Sign Language Recognition is one area where hand posture recognition is especially useful, as the postures of a signer's hands are essential features in the translation of a sign language.

Motivated by the usefulness and impact of Sign Language Recognition, we develop a method for automatically annotating images or videos of hand postures, and apply it to the problem of classifying 19 postures that are common in the Greek Sign Language. The manual annotation of such data is a time-consuming process. Their automatic annotation is based on associating 3D joint configurations of hands with classes (hand posture labels), and yields noisy such labels which can be used to train a Convolutional Neural Network. These characteristics of the problem make it a suitable candidate for applying our proposed method for automatic expansion of training datasets. We compare the results of training a CNN classifier with and without the use of our technique. Our method yields a significant increase in average classification accuracy, and also decreases the deviation in class accuracies, thus indicating the validity and the usefulness of the proposed approach.