ΠΑΡΟΥΣΙΑΣΗ

ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

ΗΜΕΡΟΜΗΝΙΑ: 24 Νοεμβρίου 2021

ΩΡΑ: 10:00 – 13:00

ΑΙΘΟΥΣΑ: Σεμιναρίων ΤΜΗΥΠ

ΟΜΙΛΗΤΗΣ: **Άγγελος Γιώτης**

Θ έ μ α

«Εντοπισμός λέξεων σε εικόνες χειρόγραφων κειμένων με επιβλεπόμενες και μη επιβλεπόμενες αναπαραστάσεις τους»

Επταμελής Εξεταστική Επιτροπή:

1. **Χριστόφορος Νίκου** (Επιβλέπων), Καθηγητής πρώτης βαθμίδας, Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
2. **Κωνσταντίνος Μπλέκας**, Καθηγητής πρώτης βαθμίδας, Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
3. **Βασίλειος Κατσούρος**, Διευθυντής Ερευνών, Ινστιτούτο Επεξεργασίας του Λόγου (ΙΕΛ), Ερευνητικό Κέντρο Καινοτομίας στις Τεχνολογίες της Πληροφορίας, των Επικοινωνιών & της Γνώσης “ΑΘΗΝΑ”
4. **Βασίλειος Γάτος**, Διευθυντής Ερευνών, Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, Εθνικό Κέντρο Έρευνας Φυσικών Επιστημών “ΔΗΜΟΚΡΙΤΟΣ”
5. **Μιχαήλ Βρίγκας**, Επίκουρος Καθηγητής, Τμήμα Επικοινωνίας και Ψηφιακών Μέσων, Πανεπιστήμιο Δυτικής Μακεδονίας
6. **Αριστείδης Λύκας**, Καθηγητής πρώτης βαθμίδας, Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
7. **Λυσίμαχος-Παύλος Κόντης**, Καθηγητής πρώτης βαθμίδας, Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Ιωαννίνων

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
Τ.Θ. 1186, ΙΩΑΝΝΙΝΑ 45110
Τ: 265100 8817 - 8813 - 7196
http:/www.cse.uoi.gr/

DEPT. OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF IOANNINA
P.O. BOX 1186, IOANNINA
GR - 45110 GREECE
T: +30 265100 8817 - 8813 - 7196
http:/www.cse.uoi.gr/

Π ε ρ ί λ η ψ η:

Vast collections of documents available in image format need to be efficiently digitized for information retrieval purposes. Many approaches from the document analysis and recognition research community have been proposed to alleviate the search process. However, the automatic recognition of degraded manuscripts using traditional Optical Character Recognition (OCR) is impractical, due to inherent challenges of these documents such as unknown layouts and fonts, the variability of handwriting and the open vocabulary. For this reason, the recent attraction for large-scale document indexing is based on a recognition-free image retrieval technique, known as keyword spotting (KWS).

The main focus of this PhD thesis lies on the systematical study and development of handwritten KWS methods as a practical solution, contrary to a costly and error-prone full text transcription. KWS methods aim to retrieve all instances of a user query in a set of document images. In an attempt to denote which parts of a KWS system require most attention to achieve high accuracy, we present a comprehensive survey of KWS techniques. To this end, each fundamental step of the respective pipeline, including layout analysis and preprocessing, feature selection and extraction, representation learning, alignment and matching is thoroughly explored. Several aspects that need to be taken into account such as robustness to writing style variabilities, the availability of training data, the evaluation protocols and measures as well as enhancement techniques which further boost the performance are highlighted and composed to a structured methodology. By these means, we suggest a theoretical foundation to be adopted by future works for unbiased evaluation and comparison.

Of most importance is the appropriate selection of features to form discriminative word image representations which can yield accurate and fast retrieval. In this thesis, we developed two template-based methods using translation and scale-invariant handcrafted features for KWS on modern and historical manuscripts. In the first work, supervised local contour features are used to train a representative shape of a word-class to address intra-class writing style variations. Its only limitation is related to out of vocabulary queries. The second method comprises an adaptation of the initial system into an unsupervised scheme for efficient and accurate script independent KWS.

Nevertheless, both former approaches are based on variable length image representations which are not fast to compare. Hence, a methodology which adapts a family of supervised, fixed-length representations that encode attribute-like features of the word image transcription is proposed for fast word retrieval. Attributes are properties that reflect the occurrence or absence of

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

Τ.Θ. 1186, ΙΩΑΝΝΙΝΑ 45110
Τ: 265100 8817 - 8813 - 7196
http:/www.cse.uoi.gr/

DEPT. OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF IOANNINA

P.O. BOX 1186, IOANNINA
GR - 45110 GREECE
T: +30 265100 8817 - 8813 - 7196
http:/www.cse.uoi.gr/

textual components (e.g. characters) at specific positions of the word. The proposed method extends this binary word image representation to include language-dependent features present in polytonic Greek text.

Following recent trends with respect to the deep learning era, in order to improve the representational power of word images, we propose a deep learning-based framework as an extraction model of deep features which are used to adapt KWS on weakly supervised diverse manuscripts with high distribution shift between source and target datasets. To this end, spatial transformations of the convolutional feature space aim to deter the KWS model so as to adversarially improve its robustness to unknown writing styles and word-classes. Finally, a technique to spot text regions in challenging historical natural images is proposed relying on adversarial learning of quaternion image descriptors which are far less resource demanding than vanilla neural network representations.

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

Τ.Θ. 1186, ΙΩΑΝΝΙΝΑ 45110
Τ: 265100 8817 - 8813 - 7196
http:/www.cse.uoi.gr/

DEPT. OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF IOANNINA

P.O. BOX 1186, IOANNINA
GR - 45110 GREECE
T: +30 265100 8817 - 8813 - 7196
http:/www.cse.uoi.gr/