

A distributed deep learning approach for histopathology image retrieval

Auteur : Defraire, Stephan

Promoteur(s) : Maree, Raphael

Faculté : Faculté des Sciences appliquées

Diplôme : Master en ingénieur civil en informatique, à finalité spécialisée en "intelligent systems"

Année académique : 2020-2021

URI/URL : <http://hdl.handle.net/2268.2/13016>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.



UNIVERSITY OF LIÈGE
FACULTY OF APPLIED SCIENCE
ENGINEERING IN COMPUTER SCIENCE
ACADEMIC YEAR 2020-2021

Author
Stephan Defraire
Supervisor
Raphaël Marée

A distributed deep learning approach for histopathology image retrieval

Digital microscopy and radiology generate growing amounts of imagery data. To help practitioners find the information crucial to establish the most accurate possible diagnoses, Artificial Intelligence tools need to be developed.

This master thesis, based on the study of existing literature and open-source code, proposes a distributed deep learning architecture that allows a user, by using a fast approximate nearest neighbour search, to retrieve similar histopathology images to a query image.

The retained Deep Learning architecture, ResNet50 with some modifications, was distributed on different servers in order to allow the handling of up to million or billion images.

It was trained on a large-scale dataset of 67 classes of annotated medical images and the obtained results are quite promising, as well for the visual similarity of the retrieved images as for the search time. This research also analyses the generalisation to classes on which the system was not trained, and the impact of the approximated search on the accuracy and the retrieval time.

Nevertheless, even though the results are positive, this system might present some limitations as it was tested on only one dataset and was not reviewed by medical practitioners.