

M2 internship - 2021

Expert-guided Deep-Attention Saliency Models for Computer Vision

Feb. 2022 - Aug. 2022 (6 months)

Supervisor(s) : Christophe Montagne and Dominique Fourer
Team / Laboratory : IRA2-SIAM / IBISC (EA 4526) - Univ. Évry/Paris-Sacay
Contact: christophe.montagne@univ-evry.fr, dominique.fourer@univ-evry.fr

Abstract : Attention is a promising deep learning paradigm to develop efficient and almost interpretable computer vision methods. This is a hot topic as shown by an increasing number of related work in the literature. The present work aims at further investigating attention model under the assumption that a Human expert can convey relevant information to enhance the accuracy of trained machine learning model. Hence, we aim at developing novel deep neural architectures including an deep-attention mechanism that is directly trained using visual information generated by human experts involved in visual segmentation task in real-world scenarios (eg. eye-tracking coordinates and/or ocular measurements). The considered applications in this study will possibly include autonomous driving and precision medicine tasks for which we aim at showing how the resulting deep-saliency function guided by expert human can enhance the efficiency of an existing deep-learning-based prediction method (eg. MRI segmentation, image recognition, etc.). To reach this goal, the recruited internship candidate will first implement a state-of-the-art attention-based deep learning based method applied to a given segmentation task. Second, we will modify the preliminary proposed method to separately train the attention block using a distinct training dataset. Finally, we will assess the different proposed deep learning methods in different scenarios to validate our assumption that a saliency function trained by experts is more efficient.

keywords : computer vision, image segmentation, attention-based model, MRI, autonomous driving, deep learning

Goals

- Bibliographical study for identifying the best state-of-the-art methods for attention-based deep neural network architectures
- Implementation of one or several state-of-the-art methods for the sake of reproducible research
- Construction of a new expert-based dataset made of eye-tracking and ocular recordings in collaboration with the partners of the project
- Development and evaluation of one or several new attention-based deep neural network architectures considering several perception-based toy problems in autonomous driving and/or medical image segmentation

Methodology

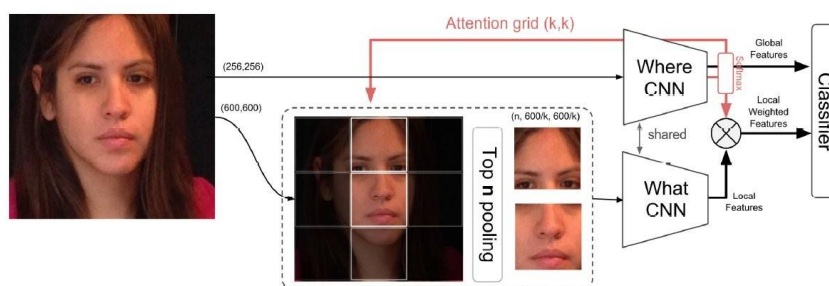


Figure 1 – Example of a state-of-the-art attention-based architecture involving two distinct convolutional neural networks as proposed by Rodriguez et al. [5].

The starting point of this research is our previous work focusing on deep learning-based methods applied to MRI segmentation in which we compared several neural architectures [1]. In this research work, we showed that the use of a bounding

box as a preprocessing step can significantly improve our MRI segmentation results. We also showed that the use of cascade architecture where the combination of 3 distinct cascaded deep neural networks can obtain promising results despite a very high computational cost. Hence, the recruited internship candidate will also investigate and compare new promising attention-based deep neural architecture such as U-net transformer with self- and cross-attention [4].

Following this idea which was recently proposed in reinforcement learning driving scenarios [3], we now propose to develop new attention-based deep neural architectures where the attention block is trained using a distinct dataset constructed by measurements collected from experts. The future proposed method will involve a deep neural architecture including two distinct convolutional neural networks (CNNs) as illustrated in Fig. 1 where the “What” CNN provides a prediction using a preprocessed version of the input preprocessed by the “Where” CNN computing the prominent regions in an image.

The difference of the new proposed method is that, we will now transfer information from a distinct training dataset dedicated to the deep-attention “Where” CNN which is designed to provide a saliency map guided by experts involved in the construction of this dual saliency dataset. The goal is to obtain a generalized saliency function targeting the regions of interest in an image to enhance the efficiency of a given machine-learning-based prediction model.

As a preliminary step, we expect to conduct a bibliographical study to discuss advantages and trade-off of the most promising state-of-the-art attention-based deep neural network architectures designed for salient object detection and image segmentation tasks [6, 2]. A second part of this research work is dedicated to the design of computer vision experiments and the collection of eye-tracking data from Human experts involved in image recognition tasks in different scenarios (driving or biomedicine). This novel saliency dataset will be used to train the new proposed deep-attention mechanism combined with the chosen baseline computer vision method.

Finally, we expect to objectively evaluate several proposed methods including or not the new proposed expert-guided deep-attention model by considering several application scenarios designed for autonomous driving and/or biomedical images segmentation.

Required profile

- good machine learning and signal processing knowledge
- mathematical understanding of the formal background
- excellent programming skills (Python, Matlab, C/C++, keras, tensorflow, pytorch, etc.)
- good motivation, high productivity and methodical works

Salary an perspectives

According to background and experience (a minimum of 577.50 euros/month). Possibility to pursue with a 3-year-funded PhD contract¹.

Références

- [1] Ikram Brahim, Dominique Fourer, Vincent Vigneron, and Hichem Maaref. Deep learning methods for mri brain tumor segmentation : a comparative study. In *Proc. IEEE IPTA 2019*, Istanbul, Turkey, November 2019.
- [2] Shervin Minaee, Yuri Y Boykov, Fatih Porikli, Antonio J Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning : A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.
- [3] Zhenghao Peng, Quanyi Li, Chunxiao Liu, and Bolei Zhou. Safe driving via expert guided policy optimization. *arXiv preprint arXiv :2110.06831*, 2021.
- [4] Olivier Petit, Nicolas Thome, Clement Rambour, Loic Themyr, Toby Collins, and Luc Soler. U-net transformer : Self and cross attention for medical image segmentation. In *International Workshop on Machine Learning in Medical Imaging*, pages 267–276. Springer, 2021.
- [5] Pau Rodríguez, Guillem Cucurull, Josep M Gonfau, F Xavier Roca, and Jordi Gonzalez. Age and gender recognition in the wild with deep attention. *Pattern Recognition*, 72 :563–571, 2017.
- [6] Wenguan Wang, Qixia Lai, Huazhu Fu, Jianbing Shen, Haibin Ling, and Ruigang Yang. Salient object detection in the deep learning era : An in-depth survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

¹ . https://www.adum.fr/as/ed/voirproposition.pl?matricule_prop=38723&site=adumR

