EURA NOVA

FACULTIES OF SCIENCE AND ENGINEERING

Master Thesis Offers 2021-2022

For the latest information, please visit https://research.euranova.eu.
20th November 2020

CONTENT

Eura Nova

Introduction

Eura Nova is a data-driven Belgian company founded in September 2008 and located in Brussels, Marseille, and Tunis. Our mission is simple: bring life to our customers' great ideas, by offering best-in-class services in data science, software engineering, and data architecture. To do so, we invest significantly in in-house expertise and state-of-the-art knowledge. In line with this course of action, we offer academic programs in collaboration with universities. These offers include bootcamps, master theses topics, research internships, and PhDs topics. See below for details.

Our Master Theses Offers

This document presents master theses topics supervised by our research & development department. Each project is an opportunity to be actively involved in the development of solutions to address tomorrow's challenges in ICTs and to implement them today. The students will work in a dedicated international team of engineers with diverse expertise in machine learning, graph theory, artificial intelligence, high performance computing, etc. They will keep Eura Nova informed of the project advancement and share their ideas and challenges using the in-house knowledge management tool. We value continuous learning and teamwork. We love to have a good time together. For more information on our R&D activities, please visit our website at https://research.euranova.eu.

How To Apply

When you have gone through our master thesis offers, pick your favourite. Draft a short text, stating why you find it interesting and what you would do about it. Send us this statement, along with your CV at career@euranova.eu. If you are interested in working on a topic that is not in our range of offers, we would be delighted to hear your proposition and invite you get in touch as well.

GridNet with Attention for Semantic Segmentation

Context

Image semantic segmentation is a popular field of research which consists in assigning a semantic label to each pixel of an image. Many methods have been proposed to tackle this problem, including GridNet [1], a little-known architecture. GridNet uses a grid-based approach where each row of grid processes features maps at different resolutions, and each column allows communication between lines. Due to its specific architecture, GridNet cannot use a pre-trained network and has to be trained from scratch. Despite this, it shows promising results.

Recently, a new technique, named Attention [2, 3], was developed by the deep learning community. Attention initially gave remarkable results in Natural Language Processing [4] and was soon adapted to image processing [5, 6]. The attention mechanism is based on a system of queries, keys and values, which allows the network to pay more attention to different input parts that may be far apart.

The objective of this thesis is to explore how the attention mechanism can be integrated into a GridNet. Especially, can the attention mechanism allow better communication between the resolutions (lines) of a GridNet? Can the attention mechanism capture better contextual information? And finally, can it facilitate the training? We expect that the use of the Attention mechanism will allow a GridNet to better capture contextual information, which will improve the long-range dependency and may also reduce the complexity of the model by reducing the training time and inference time. The study will be done on public datasets such as Cityscapes [7] and results will be compared with state-of-the-art methods [8].

Business opportunities

The exploration of the Attention mechanism carried out as part of this project will provide a good understanding of a very popular and cutting-edge technique. It is important for Eura Nova to know and master state-of-the-art techniques. In addition, it will set up a codebase useful for tasks related to image segmentation such as semantic segmentation, instances segmentation and panoptic segmentation.

Contribution

The objective is to implement and improve a GridNet using an attention mechanism. It will take a strong understanding of GridNet and Attention mechanisms as well as good creativity to combine them.

- Explore the state of the art of Attention for image processing.
- Read and understand GridNet architecture.
- Propose different approaches to use attention in a GridNet and implement them.
- Analyze the different results and explain the pros and cons.

Technologies/expertise to develop

- Keras and Tensorflow
- Semantic segmentation
- Attention mechanism

References

- [1] Fourure, D., Emonet, R., Fromont, E., Muselet, D., Tremeau, A., & Wolf, C. (2017). Residual conv-deconv grid network for semantic segmentation. arXiv preprint arXiv:1707.07958.
- [2] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.
- [3] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).
- [4] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Agarwal, S. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.
- [5] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. arXiv preprint arXiv:2005.12872.
- [6] Under double-blind review. LambdaNetworks: Modeling Long-Range Interactions Without Attention
- [7] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [8] Tao, A., Sapra, K., & Catanzaro, B. (2020). Hierarchical Multi-Scale Attention for Semantic Segmentation. arXiv preprint arXiv:2005.10821.

Combining Generative Adversarial Networks and Attention Mechanism for Multi-View Representation Learning

Context

In the real world, multi-view or multi-modal data are quite common. Indeed, a sample can have different representations depending on its source, on the sensor that captured it, or even on the applied method generating its features. For instance, if we consider social media content or e-commerce websites, most of the items are represented by an image and by some description text. A second example can be related to healthcare applications where patients can be described by their symptoms, the results from tests, and data coming from sensors monitoring their vitals.

Several techniques have been used to handle multi-modal or multi-view datasets. The objective is to learn the relationships between the different views that refer to the same item and embed them in a common latent space. The learnt common latent manifold can then be used to solve several machine learning tasks such as clustering, classification, etc.

Recently, generative adversarial networks (GANs) [1] have gained popularity in the field of multi-modal embedding thanks to their limited need for input data, their ability to generate fake data that are very similar to the real ones, as well as their ability to handle missing data or partial views. For instance, in [2], high-level features extracted from each view are embedded and fused to give a unique representation of the given sample. The proposed framework called GP-MVC is even able to handle missing views by the means of the cycle consistency between the generators.

The main challenge with the fusion approach in the context of multi-view representation learning is to soft-select relevant features of the views by assigning learnable weights. If we consider the context of action recognition using RGB and Depth images, some actions can be more detectable within the first sensor while others may appear clearer with the second type of images [4].

The objective of this thesis is to explore how the attention mechanism [4] can lead to discovering meaningful weights/scores and provide complementary information and thus enhance multi-view representation as shown in [5] [6] [7]. The main idea is to integrate one or more attention networks within a GAN-based architecture to better represent multimodal/view data.

Business opportunities

The exploration of Attention mechanism carried out as part of this project will provide a good understanding of a very popular and cutting-edge technology. This thesis might lead to a base implementation that would be used for several tasks related to multimodal/view data handling for machine learning applications.

Contribution

The objective is to integrate the attention mechanism within a GAN-based architecture for multiview representation learning.

- Explore the state of the art of multimodal/view representation learning.
- Explore the state of the art of Attention and self-attention mechanisms.
- Read and understand the GP-MVC architecture and code.
- Propose different approaches to integrate attention networks with GP-MVC and implement them.
- Conduct a comparative analysis of the results.

Technologies/expertise to develop

- Encoders/decoders, GANs.
- Tensorflow, Keras and Pytorch.
- Multiview representation learning.
- Attention mechanism.

References

- [1] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. C., and Bengio, Y. Generative adversarial nets. In NIPS, 2014
- [2] Wang, Q., Ding, Z., Tao, Z., Gao, Q., & Fu, Y. (2020). Generative Partial Multi-View Clustering. arXiv preprint arXiv:2003.13088
- [3] Wang, L., Ding, Z., Tao, Z., Liu, Y., & Fu, Y. (2019). Generative multi-view human action recognition. In Proceedings of the IEEE International Conference on Computer Vision (pp. 6212-6221)
- [4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).
- [5] Yuan, Y., Xun, G., Ma, F., Wang, Y., Du, N., Jia, K., ... & Zhang, A. (2018, November). Muvan: A multi-view attention network for multivariate temporal data. In 2018 IEEE International Conference on Data Mining (ICDM) (pp. 717-726). IEEE.
- [6] Liu, J. W., Ding, X. H., Lu, R. K., & Xionglin, L. U. O. (2020, August). Self-attention Multi-view Representation Learning with Diversity-promoting Complementarity. In 2020 Chinese Control And Decision Conference (CCDC) (pp. 3972-3978). IEEE
- [7] Li, X., Wang, C., Tan, J., Zeng, X., Ou, D., Ou, D., & Zheng, B. (2020, April). Adversarial Multimodal Representation Learning for Click-Through Rate Prediction. In Proceedings of The Web Conference 2020 (pp. 827-836).