Master of Science Topics 2021-2022, AI-MAS Laboratory

Machine Learning for Computer Vision and Image Analysis Tasks

Title: Understanding image classification based on deep neural networks (2 topics)

Coordinator: *Prof. Adina Magda Florea (adina.florea@upb.ro)*

Description:

While deep neural networks have had a large impact on a variety of different tasks, explaining their predictions is still challenging. The lack of tools to inspect the behavior of these models makes DNNs less trustable for those domains where interpretability and reliability are crucial, like autonomous driving, medical applications, and finance.

One of the main approaches to understand the behaviour of a DNN is assigning an attribution value, also called "relevance" or "contribution", to each input feature of a network. In this context, we propose 2 research topics:

1. Understanding image classification tasks using Layer-wise Relevance Propagation

Layer-wise Relevance Propagation (LRP) allows to visualize the contributions of pixels to predictions for multilayered neural networks. These pixel contributions can be visualized as heatmaps and are provided to a human expert who can intuitively not only verify the validity of the classification decision, but also focus further analysis on regions of potential interest. LRP is computed with a backward pass on the network. The method assigns a relevance to units of different layers of the network. It starts at the output layer L and assigns the relevance of the target neuron equal to the output of the neuron itself and the relevance of all other neurons to zero. Then it proceeds layer by layer, redistributing the prediction score until the input layer is reached.

Tasks:

- investigate the properties of LPR to understanding deep learning classification tasks;
- implement the approach;
- propose improvements;
- perform evaluations of the implementation, both technical and nontechnical.

S. Bach, A. Binder et.al. On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation, pp.1-46, PLOS One, 2015.

A. Binder, G. Montavon, S. Lapuschkin, K.R. Müller and W. Samek, "Layer-wise relevance propagation for neural networks with local renormalization layers". In: Proc. of Conf. on Artificial Neural Networks, pp.63-71, Springer, 2016.

2. Understanding image classification tasks by Class Activation Mapping

Class Activation Mapping (CAM) applies to CNN and uses global average pooling, which acts as a structural regularizer, preventing overfitting during training. The global average pooling also allows the network to retain its localization ability until the final layer. This allows identifying the discriminative image regions and creating activation maps that act as a detector for different patterns in the image, localized in space. By inspecting these activation maps, correct and incorrect classifications can be analyzed.

Tasks:

- investigate the properties of this approach to understanding deep learning classification tasks, including Gradient-weighted Class Activation Mapping (Grad- CAM), which extends CAM by producing discriminative heatmaps that achieve object localization without training;
- implement the approach;
- propose improvements;
- perform evaluations of the implementation.

B. Zhou, A. Khosla, A. Lapedriza, A. Oliva and A. Torralba, "Learning deep features for discriminative localization". In: Proc. of Conf. on Computer Vision and Pattern Recognition, pp.2921-2929, 2016.

R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization". In: Proc. of Conf. on Computer Vision, pp.618-626, 2017.

https://alexisbcook.github.io/2017/global-average-pooling-layers-for-object-localization/

Title: Semantic segmentation of images with a semi-supervised approach

Coordinator: Prof. Adina Magda Florea (adina.florea@upb.ro)

Description:

Deep learning supervised methods require large data sets that are labeled in order to achieve performance of image segmentation, which are not usually easy to obtain. Different self-supervised and semi-supervised approaches were proposed to alleviate this issue. The research topic will consist of investigating and implementing several methods of semi- supervised learning for image segmentation, e.g., [1].

[1] https://arxiv.org/pdf/2004.14960.pdf

Title: ML methods for space surveillance

Coordinator: Assist. Prof. Mihai Trascau (mihai.trascau@upb.ro)

Tasks:

- Develop ML algorithms able to process images from optical telescopes, to identify possible objects of interest, correlate the measurements in successive images according

to the observational scheme, and extract accurate information on the position and magnitude of detected objects.

- Detect and recognize interesting object in images taken by telescopes, such as stars, satellite or cosmic ray.
- Tracklet generation. Use ML algorithms to identify and predict the trend of the tracklet (tracklet pattern), in other words connecting consecutive images in an observational sequence to generate object tracklets.

Title: Character animation using deep learning

Coordinator: *Prof. dr. ing. Irina Mocanu (irina.mocanu@upb.ro)*

Description:

Interactively synthesizing novel combinations and variations of character movements from different motion skills is a key problem in computer animation. The scope of the project is to propose a deep learning framework for creating different movements, that imitates animation layering using neural networks with the aim to overcome typical challenges when mixing, blending and editing movements from unaligned motion sources.

Bibliography:

1. https://developer.nvidia.com/blog/virtual-character-animation-system-uses-ai-to-generate-more-human-like-movements/
2. https://github.com/sebastianstarke/AI4Animation
3. https://analyticsindiamag.com/how-disney-is-using-machine-learning-for-realistic-animation/.
4. https://medium.com/embarkstudios/transforming-animation-with-machine-learning-27ac694590c

Title: Fashion design assistant for creative support using GAN models

Coordinators: *dr. ing. Alexandru Sorici (alexandru.sorici@upb.ro)*
dr. ing. Mihai Trăscău (mihai.trascau@upb.ro),
drd. ing. Teodor Poncu (dan_teodor.poncu@upb.ro)

Description:

Topic designed in partnership with Infosys Consulting Romania

Generative Adversarial Networks (GANs) have been implemented with great success in many computer vision tasks, like image-to-image translation [1], conditional synthesis [2], data augmentation, style transfer [3], image super-resolution [4], etc.

In this project we aim to use GAN models to support creative efforts for the fashion design task. [5, 6, 7]. Starting from a given dataset of fashion articles, the goal is to generate new items conditioned by style considerations and desired attributes of the target.

The process of garment generation can be further improved/expanded to lead to the generation of 3D meshes of the clothing items, by further conditioning the GAN models on person specific physiological attributes, such as height, waist size, etc.

The physiological attributes are directly determined by a *person scan* procedure, carried out using the Kinect v2/v3 or Intel RealSense RGB-D cameras.

[1] ZHU, Jun-Yan, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Proceedings of the IEEE international conference on computer vision*. 2017. p. 2223-2232.
[2] PARK, Hyojin; YOO, YoungJoon; KWAK, Nojun. Mc-gan: Multi-conditional generative adversarial network for image synthesis. *arXiv preprint arXiv:1805.01123*, 2018.
[3] ISOLA, Phillip, et al. Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. p. 1125-1134.
[4] WANG, Xintao, et al. Esrgan: Enhanced super-resolution generative adversarial networks. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
[5] ZHU, Shizhan, et al. Be your own prada: Fashion synthesis with structural coherence. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2017. p. 1680-1688.
[6] SBAI, Othman, et al. Design: Design inspiration from generative networks. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
[7] KANG, Wang-Cheng, et al. Visually-aware fashion recommendation and design with generative image models. In: *2017 IEEE International Conference on Data Mining (ICDM)*. IEEE, 2017. p. 207-216.

Title: IntAli: Food recognition, food recommendation and autonomous detection of food in a refrigerator

Coordinators: *Prof. Adina Magda Florea (adina.florea@upb.ro),*

Assist. Prof. Mihai Trascau (mihai.trascau@upb.ro)

Tasks:

- Develop an intelligent application to recognize foods and cooks from images, recommend healthy food and recipes, automatically detect foods in a smart refrigerator and recommend necessary supplies.
- Allow the user to personalize the application according to his/her preferences, food dietary restriction, and available ingredients.

Title: Food recognition and recipe identification

Coordinators: *Prof. Adina Magda Florea (adina.florea@upb.ro)*
Assist. Prof. Mihai Trascau (mihai.trascau@upb.ro)

Description:

The research topic aims at recognizing basic food elements from pictures taken by the user with a mobile phone and then finding the recipes that can be prepared with these

items. A link with a smart refrigerator having the ability to capture in images the foods in the refrigerator is to be integrated in the topic. One of the main challenges is to also determine the amount of food items available for cooking in such a way as to scale the culinary recipe to the available quantity. Moreover, the system has to find the recipes with the least number of missing ingredients that match the current existing items, find on-lien stores that provide these missing items and propose them to the user.
https://foodai.org/#index