Contribution ID: **14** Type: **not specified**

Picture Captioning

The research focused on classic image captioning based on a coder-decoder structure, where the coder encodes the image features. At the same time, the decoder produces a caption –a phrase describing the image content. We investigated the decoder part by testing multiple convolutional-neural-network-based backbones –feature extractors. This investigation aimed to find the optimal encoder, i.e., one that maximizes text generation metrics BLEU_1-Bleu_4, CIDEr, SPICE, and METEOR. Moreover, we worked on optimizing beam-search parameters used by the decoder to generate alternative phrases. Our research proves that an optimal choice of model's hyperparameters increases caption generation efficiency.

Authors: Prof. IWANOWSKI, Marcin; Mr BARTOSIEWICZ, Mateusz; Mr SZCZEPAŃSKI, Piotr; Mr ZIELIŃSKI, Karol; Mr ZIÓŁKIEWICZ, Albert

Presenters: Mr SZCZEPAŃSKI, Piotr; Mr ZIELIŃSKI, Karol; Mr ZIÓŁKIEWICZ, Albert

Session Classification: Session C (Poster)