

Spotkanie 06

Zagadnienia obowiązujące na quizokartkówkę

- Na czym polega architektura Vision-Transformer?
- Jak wykorzystujemy transformery do przetwarzania obrazów?
- Jak działa multihead-attention?
- Jak przygotowujemy obraz do przetwarzania transformerem?
- Co to są embeddingi?
- Jak działa augmentacja, dlaczego ją stosujemy?

Pomocne materiały

- <https://arxiv.org/pdf/1706.03762.pdf>
- <https://arxiv.org/pdf/2010.11929.pdf>
- https://www.youtube.com/watch?v=OyFJWRnt_AY
- <https://www.youtube.com/watch?v=g2BRluln4uc>
- <http://proceedings.mlr.press/v119/xiong20b/xiong20b.pdf>
- <https://datascience.stackexchange.com/questions/90649/class-token-in-vit-and-bert>
- <https://storrs.io/attention-rollout/>