Self-Supervised Learning for Robust Image Classification

Ladyna Wittscher
Jan Diers, Christian Pigorsch

Self-supervised learning

- Training without big, high-quality or labeled datasets (Kolesnikov et al., 2019; Zhang et al., 2016; Zhang et al. 2017)

- Training in two steps: pretraining and downstream task (Doersch and Zisserman, 2017; Gidaris et al., 2019; Jing and Tian, 2020; Doersch et al., 2015)
  - Supervisory signals generated from the training data itself (such as position, color or self-motion)
  - Self-generated pseudo-labels
  - Pretraining requires a high-level of semantic understanding of the data
  - Network learns meaningful visual representations
  - Initialization of weights for more complex downstream tasks

Research question

- Does pretraining increase robustness concerning:
  - Small datasets?
  - Data imbalance?
  - Pixel loss?
  - Combinations?

Conclusion

- Self-supervised learning is significantly better than the standard approach given data manipulations
- More difficult conditons → increased difference between the models
- Relative classification improvement due to pretraining given highest levels of data manipulation

Scenario	MNIST	Fashion-MNIST
Data reduction	11.2%	15.2%
Pixel loss	7.3%	2.2%
Imbalance	0.9%	0.8%
Combination data reduction & pixel loss	42.9%	

- Combinations of challenges pose a greater difficulty than the sum of the individual ones

References

Carl Doersch, Abhinav Gupta, and Alexei A. Efros. Unsupervised visual representation learning by context prediction. In *Proceedings of the IEEE International Conference on Computer Vision,* pages 1422–1430, 2015.

Carl Doersch and Andrew Zisserman. Multi-task self-supervised visual learning. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2051–2060, 2017.

Spyros Gidaris, Andrei Bursuc, Nikos Komodakis, Patrick Pérez, and Matthieu Cord. Boosting few-shot visual learning with self-supervision. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 8059–8068, 2019.

Longlong Jing and Yingli Tian. Self-supervised visual feature learning with deep neural networks: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020.

Alexander Kolesnikov, Xiaohua Zhai, and Lucas Beyer. Revisiting self-supervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1920–1929, 2019.

Richard Zhang, Phillip Isola, and Alexei A. Efros. Colorful image colorization. In *European conference on computer vision*, pages 649–666. Springer, 2016

Richard Zhang, Phillip Isola, and Alexei A. Efros. Split-brain autoencoders: Unsupervised learning by cross-channel prediction. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 645–654, 2017.

Xin Zheng, Yong Wang, Guoyou Wang, and Jianguo Liu. Fast and robust segmentation of white blood cell images by self-supervised learning. *Micron*, 107:55–71, 2018.

Thank you for your attention!