

FID3018 - Opposition Report

Opposition to presentation of Abubakreledik Karali

Tianze Wang *EECS, SCS*
KTH Royal Institute of Technology
Stockholm, Sweden
tianzew@kth.se

I. SUMMARY

The seminar presents topics on self-supervised learning methods for learning through large unlabeled datasets. The three papers selected touch on topics of self-supervised learning on computer vision tasks. The seminar is very interesting and has very good slides.

II. CHOICE OF PAPERS

- 1) Gidaris, Spyros, Praveer Singh, and Nikos Komodakis. "Unsupervised Representation Learning by Predicting Image Rotations." *International Conference on Learning Representations*. 2018. [1]
- 2) Mundhenk, T. Nathan, Daniel Ho, and Barry Y. Chen. "Improvements to context based self-supervised learning." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018. [2]
- 3) Kolesnikov, Alexander, Xiaohua Zhai, and Lucas Beyer. "Revisiting self-supervised visual representation learning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019. [3]

The three papers selected all study the topic of self-supervised visual representation learning which is concerned with learning visual representations on large unlabeled datasets by some predefined tasks that do not need manual labeling. The first paper aims to learn visual representations by predicting the rotations of images. The second paper aims to improve the qualities of learning by preventing the model to learn from common artifacts in images such as chromatic aberration. The third paper studies the choice of convolutional neural networks (CNNs) in self-supervised learning and uncovers that standard designs CNN models that have a good performance on supervised tasks do not necessarily translate to good performance in self-supervised learning settings.

Overall, the choice of papers focuses on the recent progress in self-supervised visual representation learning and can be viewed as one way of mining through massive datasets without explicit supervision.

III. PRESENTATION

The presentation is clear and coherent with a nice choice of papers. The presentation of each paper follows the same strategy with motivation, contributions, methods, strong points, and weak points. The transition between the three papers is smooth. The presenter also gives the necessary background needed for understanding each paper which is always appreciable. The inclusion of personal summaries on each paper is also welcomed.

The slides are also nicely made with clear and concise visualizations that capture the key points that the presenter tries to convey. It would be interesting if more self-supervised learning tasks can be introduced and compared with the ones in the selected papers. Although it is outside the scope of visual representation learning of the seminar, it would be also interesting to see how well self-supervised learning methods proposed for visual representation learning can transfer to learning tasks of other domains, e.g., natural language processing.

Despite some network issues during the seminar, the presentation went smoothly. The presenter has a decent understanding of the topic and the discussion session is also inspiring and insightful.

REFERENCES

- [1] S. Gidaris, P. Singh, and N. Komodakis, "Unsupervised representation learning by predicting image rotations," in *International Conference on Learning Representations*, 2018.
- [2] T. N. Mundhenk, D. Ho, and B. Y. Chen, "Improvements to context based self-supervised learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9339–9348.
- [3] A. Kolesnikov, X. Zhai, and L. Beyer, "Revisiting self-supervised visual representation learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 1920–1929.