Assignment: Some Key Ideas From Deep Convolutional Neural Network Architectures

ENGN8536, 2020

August 21, 2020

Until around 2018, some major progress was made in Deep Convolutional Neural Networks by augmentations that were introduced in Architectures. It is important to understand these innovations, not just from a sense of (recent) history. Firstly, many of these networks (e.g., ResNet, VGG, DenseNet) are frequently used as backbone networks for new problems with transfer learning. Secondly, and perhaps more importantly, these augmentations represent key ways in which research has dealt with key problems in CNNs. These innovations make a significant difference to the performance of networks.

This assignment asks you to investigate these innovations in the theory and practice of CNNs in a non-trivial way: what they were, how did they work, what have these led to in subsequent research. It also builds research skills, getting you to investigate and understand, and track innovations and understanding through the literature. For Deep CNNs in particular, we would like you to be able to understand how key ideas propagate into other papers, and how they are used to solve related problems in varying contexts.

1 Part 1: GoogLeNet (3 marks)

- What were the key innovations of GoogLeNet with respect to VGG?
- One key mechanism reduces the number of parameters in a block. Explain how it does this.
- Explain how this/these mechanisms help improve accuracy.

2 Part 2: ResNet (4 marks)

ResNet created much deeper networks than its predecessors (particularly VGG).

- Describe the key innovation(s) of ResNet.
- Explain what was the problem that this feature addressed and how ResNet addressed the problem. (Hint: you need to talk about vanishing gradients.)
- Compare and contrast the design of ResNet vs that of VGG.
- What is the importance of Batch Normalization in this network?

3 Part 3: Squeeze and Excitation Networks (3 marks)

- What is the key innovation of Squeeze and Excitation networks?
- How does this feature improve performance?
- Squeeze and excitation networks introduce a new module. Explain how many parameters are used by one of these modules (use and define variables for this).

4 Part 4: What has this lead to? (10 marks)

Each of these networks has had influence on subsequent research. For example, ResNet and DenseNet use bottlenecks. Part 4 is an open-ended research question. It requires you to:

- Choose one of the papers below - your **source paper** (you may chose more than one if you wish).
- Search through more recent papers, and find a paper - your **follow-up paper** - that uses a key innovation from your source paper in some way (again, you may use more than one if you wish).
- Then your answer should at least:
  - Write a high level summary of the paper including the problem, the proposed solution, the results and any weaknesses. (2 marks)
  - Name and briefly describe the key innovation(s) from the source paper. (2 mark)
  - Then for the follow-up paper describe how this innovation was used (6 marks overall for this part):
    - Explain how the follow-up paper used the key innovation(s).
    - Describe what the follow-up paper did in terms of adaptations of the innovation(s) in order to use in the context of the follow-up paper's task.
    - Explain why they used the new innovation(s).
    - Explain what was the key contribution of the follow-up paper. (It may have nothing to do with the source paper at all).
  - The key requirement is show insight and initiative in your study for the follow-up paper, and the key content could come in any of these.
- Note that your report must reference all the papers you discuss in proper bibliographic format (any format is OK, you can use the citation style of any of the papers as a style).
- Latex can be a helpful tool for writing such documents.

Note that your answer can go into issues other than these questions as you consider important. The key idea is to show a deep understanding of a key innovation and how it progresses through subsequent research.

For your choice of papers, you must choose at least one paper for your follow-up paper from the papers that have been published in IEEE-Computer Vision and Pattern Recognition, IEEE-International Conference on Computer Vision, or the European Conference on Computer Vision. Feel free to use other publications if you look at more than one follow-up paper.

5 Resources

- The week 6 lecture notes and lecture cover some highlights.
- The source papers (note these are all available on arXiv):
  - Gradient-Based Learning Applied to Document Recognition, LaCun et al, Proc IEEE, Nov 1998, Cited by 28998
  - ImageNet Classification with Deep Convolutional Neural Networks, Alex Krishevsky, Ilya Sutskever, and Geoffrey Hinton, Proc NIPS, 2012, Cited by 67748
  - Very Deep Convolutional Networks for Large Scale Image Recognition, ICLR 2015, Simonyan and Zisserman, ARXIV 2014, Cited by 42645
  - Going deeper with convolutions, Szegedy et al, IEEE-CVPR, June 2015, Cited by 23430
  - Deep Residual Learning for Image Recognition, He et al, IEEE-CVPR, June 2016, Cited by 52996

  - G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243. (Cited by 10519)
  - J. Hu, L. Shen and G. Sun, "Squeeze-and-Excitation Networks," 2018 IEEE-CVPR, Salt Lake City, UT, 2018, pp. 7132-7141, doi: 10.1109/CVPR.2018.00745. (Cited by 3667)

- Many lectures, such as Stanford CS231 also cover architectures and may give other leads.
- Here are some sample papers that you could look at:
  - DenseNet as above.
  - Image Super-Resolution Using Very Deep Residual Channel Attention Networks (available at: https://arxiv.org/abs/1807.02758).
  - SCA-CNN: Spatial and Channel-wise Attention in Convolutional Networksfor Image Captioning (https://openaccess.thecvf.com/content_cvpr_2017/papers/Chen_SCA-CNN_Spatial_and_C

6 Notes

1. The assignment submission will be due on Sunday, 27 September 2020, 6:00pm. You are required to complete all tasks and answer all questions. Please submit a single .pdf report file. There is no requirement for any coding for any other documents, but if you want to submit additional files you should create a single zip file with the .pdf report and any additional files. Do not submit any datasets.
2. Name your zip file as Assignment_u0000000.zip, replacing u0000000 with your uni-ID.
3. The full grade of this lab is 20 marks. This assignment makes up 20% of the total course assessment.