Review: Group 46 - Image Classification of real estate images with transfer learning

Authors: Oskar Råhlén & Sacharias Sjöqvist

Reviewer: Mattias Larsson

May 23, 2019

1 Summary of the paper

In this paper classification of real estate photos using different types of deep neural network architectures was investigated. Models pre-trained on famous image classification data sets were used to further train on classifying real estate photos. This was both done using feature extraction as well as fine-tuning, two transfer learning approaches. The goal of the trained models was to be able to correctly classify the photos into "fireplace" or not, "balcony" or not and type of room (kitchen, bathroom or bedroom). The results of the research showed that the "balcony" and type of room models both achieved really good accuracy when tested. The "fireplace" model was slightly worse, but still very accurate. The results furthermore did not indicate on any type of architecture being clearly superior.

2 Title and abstract

- The title clearly describes the report and its content. You could maybe sneak in "deep learning" as well, if you want to make it really clear exactly what is being used.

- The abstract is well written and covers the whole report.

- The English abstract is missing, which you probably already are aware of.

3 Introduction - problem statement

- The research question is easily identified, but might be a bit vague on its own. The scope section however makes it clear what type of attribute in the images you are going to look at, but it is not understood in this section what type of machine learning methods that are going to be researched. It becomes clear later on in the method section, but it could be said already in this section, either woven into the research question or under the scope section.

- The introduction is interesting and really helps building up the case for the purpose of the report. It makes it obvious to the reader why this is something interesting to research.

- It is a bit unclear how the question will be answered, since it is only asking if it is possible to classify the real estate images. How do you define possible in this case? The models performing better than pure chance or better than some other already existing methods? This is something that you might want to clarify to the reader.

4 Background and related work

- The background section is extensive, but still very relevant to the report.

- The way the subsections follow upon each others is logical and makes it easy for the reader to understand how the concepts are related.

- The usage of illustrations and references are done really well.

- Really relevant sections describing the different architectures that are used later on, as well as the ImageNet data base.

- A relevant or related work section is something that could be added. Even if you looked into if something similar had been done before (research around classifying real estate images) and you did not find anything, there could be something said about that.

5 Methods

- The background makes it easy for the reader to fully understand this section.

- It is very clear how the experiments were conducted. If the code still exists, maybe you could include it in an appendix or link to a repository if that was used?

- The motivation and description of the data gathering is good.

- The motivation for why you chose the architectures could maybe be moved to the beginning of that section ("Implementation"), instead of coming at the very end. This would make it much more clear to the reader.

- In the evaluation section you describe how you actually evaluate the experiments. You could define more precisely what you mean with "noggrannhet". My assumption is that you simply measure how many images are classified as their correct classes, but do you for example at all care about e.g. false/true positives/negatives, and the ratios?

6 Results

- You might want to clarify in the beginning of this section that all the figures can be found in the appendix.

- Could make it clear that the accuracy "Max. noggrannhet" in the tables are percentages of correctly classified images, which is my assumption.

- The result section is straight to the point and only relevant information is provided, which is great. It is easy to find specific results by looking at the subsection titles and tables.

- It is a bit annoying that all the figures are in the appendix, but it is somewhat understandable since there are so many of them. It could be an idea to at least include the most interesting and relevant graphs in the result section. After all the result is the most interesting part of the report and you might want to highlight it with the help of a graph.

- Maybe include examples of the images that are classified?

- Is it the cut-off in the y-axis on the graphs in the appendix that makes it look like the accuracy is already very high (clearly larger than 50 %) at 0 epochs or are the models actually that good already without training? My assumption is that they are just visually misleading, which is a bit of a problem, but if it is the other case that would be really weird and something you should look into?

7 Discussion and conclusions

- The most important results and possible explanations are discussed, which is appropriate.

- The reasoning around the limitations and what could be done in the future is very interesting.

- As you describe, looking into which images that were classified incorrectly to find a common cause would be very interesting for future research.

- Nice discussion about the ethical aspects and sustainability, which I did not expect.

- The conclusion definitely answers the research question and is based on the results in a proper fashion.

8 Overall characteristics

- The report is well written and well structured.

- The text is adjusted to the reader.

- The tables in the result section are clear and graphs in the appendix have clear lables etc.

- The front page is missing the names of the supervisor and examiner.

- Overall a good report. Well done.