MSc Data Science Project: Can a Convolutional Neural Network judge a book by its cover?

Ryan Hill MSci

Supervisor: Dr Hubie Chen

Birkbeck, University Of London

rhill06@mail.bbk.ac.uk

July 18, 2020

Abstract

Coming soon...

Contents

1 Introduction

Introduce the project, describe the goal and the idea behind why this is important. Explain the structure of this work,

1.1 Related Works

Discuss the work of Iwana as the main source, reference their methods and results but keep much for later where relevant for comparison with my work.

2 Theory

2.1 Components of Convolutional Neural Networks

Describe the maths and the theory behind the various components of CNNs to the best of my ability, referring to the source papers for more complex combined units where sensible.

2.2 Transfer Learning

Describe the theory and maths behind transfer learning to the reader.

2.3 Activation Maximisation

Describe the theory and/or maths of activation maximisation and the methods used within the package we use the regularisations applied.

3 Data collection and pre-processing

3.1 Image Download and Manual Review

Detail the process of collecting the data and manually reviewing those chosen for use within the train/test set, as well as the validation split.

Rule of thumb: single front facing book cover, book cover did not contain pictures of multiple books, and was not just an arrangement of book covers. Covers designed specifically for a boxset were allowed.

3.2 Image Pre-processing Technique

Detail the approaches taken when considering how to pre-process the images and their results, including what the final choice was.

4 Model Training

4.1 Configuration

Detail the models used, the hyperparameters set as well as the optimiser and the loss functions chosen. Denote the pre-trained sources and the reason for choosing these models.

4.2 Training Performance

Detail and visualise the performance of the models through training; compare length of training to that of Iwana.

5 Evaluation

5.1 Results

Present and discuss the basic results for each model, including a comparison with Iwana with the caveat of a slightly different dataset and different pre-processing.

5.2 Further Analysis

Consider further, in depth, analysis on the best performing model, including some of the insight provided by Iwana.

Idea: compare class average colour by predicted and actual and compare to accuracy?

5.3 Feature Visualisation

Detail the approach taken to try and visualise the ideal input to trigger the most accurate classes in an attempt to understand what feature the model is picking up on. Show output and discuss impact.

5.4 Discussion

Discussion around the reasons we have seen some of these results if not already covered in the previous subsections.

6 Conclusion

Discuss the project as a whole, the impact of the results when compared to the literature, and what future work there could be off the back of this.

References

A CNN Architectures

Detail the exact configurations of these architectures

A.1 MobileNetV2

A.2 Inception-ResnetV2

A.3 ResNeXt50

B Technology

List of software and packages/libraries including their versions.