

COMP 576 Final Project Proposal

Deep CNN Book Cover First Impression Analysis

Josh Dunning & Danny Andreini

November 2018



1. Introduction

Despite the age old adage, readers often *do* judge books by their covers, and a poor design can separate a flop from a bestseller. By designing a deep convolutional network that can determine which genre a given book's cover conveys, we can analyze the effectiveness of cover designs in promoting book sales.

2. Goals

We would like to create a model that can predict a book's genre based off of its cover with accuracy such that the model's top three choices for genre contain the actual genre over 60% of the time. This would allow our model to give an author or publisher a basic understanding of what sorts of genres a potential book cover conveys, even if the model does not properly identify the books true genre.

3. Feasibility

Some studies, as referenced at the end of this document, have done research in this area, and have produced promising results. However, we plan on exploring a variety of different CNNs not included in those studies, and as we continue our research and obtain more results, we will consider utilizing a variety of different datasets and ways of prioritizing genres within those datasets.

4. Data

We will be using the *Goodbooks-10k* dataset, which contains large amounts of user data collected from *goodreads.com*. Of the five provided files in this dataset, we will be utilizing *books.csv*, *book-tags.csv*, and *tags.csv*. We will take advantage of the following rows in these datasets:

books.csv

`goodreads_book_id`: A unique ID number assigned to this book

`title`: The english title of this book

`small_image_url`: A link to a png file of the book's cover

book-tags.csv

goodreads_book_id: *A unique ID number assigned to this book*
tag_id: *A unique ID number representing a user defined book genre tag*
count: *The number of users that assigned this tag to this book*

tags.csv

tag_id: *A unique ID number representing a user defined book genre tag*
tag_name: *The string of user input corresponding to a unique tag ID*

For each book, we will use SQL to create an object of the following form:

```
{  
  goodreads_book_id: 2657,  
  title: "To Kill a Mockingbird",  
  small_image_url: "https://images.gr-assets.com/books/1361975680s/2657.jpg",  
  top_tags: [ "fantasy", "adventure", "romance" ]  
}
```

These objects will be used to generate the inputs to our network.

5. Model and Method

We will be using a convolutional neural network for its ability to divide covers into distinct sections for analysis. To train this model, we will classify all books in our dataset into one or more of ten overarching categories, then determine loss by the model's ability to correctly identify any of the genres in this set. We will be comparing several popular modern CNN structures for image recognition, such as ResNet and Inception, and reporting our results for each.

6. Reference

<http://fastml.com/goodbooks-10k-a-new-dataset-for-book-recommendations>

<https://www.technologyreview.com/s/602807/deep-neural-network-learns-to-judge-books-by-their-covers>

http://cs229.stanford.edu/proj2015/127_poster.pdf

<http://cs231n.stanford.edu/reports/2017/pdfs/814.pdf>

<https://hackernoon.com/a-comprehensive-design-guide-for-image-classification-cnns-46091260fb92>