

# Judging A Book By Its Cover

Ishika Tarin

ishikaime6561@gmail.com

Bangladesh University of Engineering and Technology  
Dhaka, Bangladesh

Afroza Parvin Disa

1805114@ugrad.cse.buet.ac.bd

Bangladesh University of Engineering and Technology  
Dhaka, Bangladesh

## 1 INTRODUCTION

While convolutional neural networks (CNNs) have proven adept at extracting features from visual data, their application in automatically classifying book genres based solely on cover images presents a unique challenge.

This research delves into the potential of CNNs to uncover the intricate relationship between the visual design elements employed on book covers and the genre they represent. However, our investigations suggest that a single CNN model might not be sufficient for achieving high accuracy in genre prediction. The inherent ambiguity associated with book cover design, where elements can hold multiple meanings or cater to a broader genre category, can pose a significant obstacle for a single model to learn definitive patterns.

That's why, recognizing this challenge, we shifted our approach towards a more nuanced signal processing strategy. This new direction aims to extract not just high-level features but also potentially underlying patterns and relationships within the cover image with a much better performance than CNN.

## 2 RELATED WORKS

A study was suggested to explore the learnability of connections between books and their covers. The aim was to employ a deep Convolutional Neural Network (CNN) to forecast a book's genre by analyzing the visual cues presented on its cover.[4]

Works have been done using new sets of color features in the field of computer vision and image processing which are inspired by the work of artists, the aim was to classify different subjective properties of paintings, including aesthetic quality, beauty, and liking of color. We then investigate if observers have individual tastes and opinions when evaluating different properties of artworks. [1] Gatys, et al. [3] used deep CNNs to learn and copy the artistic style of paintings.

In the field of genre classification, there have been several attempts to classify music [6], [7], text [5], painting [2] etc.

## 3 METHODOLOGY

### 3.1 Dataset

Our research leveraged a substantial dataset encompassing nearly 200,000 books across 29 distinct genres. However, to optimize the training process for our model, we curated a customized subset.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

Conference'17, July 2017, Washington, DC, USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM

<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

This refined dataset comprised 2,819 entries, ensuring a balanced representation of each genre. Specifically, we incorporated approximately 100 books for each genre. Each entry within the dataset corresponded to a unique identifier associated with its respective genre. By meticulously constructing this tailored dataset, we aimed to provide our model with a focused and informative training foundation that would enhance its ability to accurately predict book genres based solely on cover image analysis. This train dataset is available in public github .

### 3.2 Our First Approach Using CNN

We first used CNN approach with 23 layers for the proposed project. Here is the breakdown of the model layers:

- **6 Convolutional Layers:**
  - 1st Convolutional Layer
  - 2nd Convolutional Layer
  - 3rd Convolutional Layer
  - 4th Convolutional Layer
  - 5th Convolutional Layer
- **5 MaxPooling Layers:** Following each of the first 5 Convolutional Layers (one per layer)
- **5 BatchNormalization Layers:** Following each of the first 5 Convolutional Layers (one per layer)
- **3 Activation Layers (ReLU):** Following each BatchNormalization layer after the Convolutional Layers
- **2 Dropout Layers:** Within the Fully Connected layers
- **6 Fully Connected Layers:**
  - 1st Fully Connected Layer
  - 2nd Fully Connected Layer
  - 3rd Fully Connected Layer
  - Output Layer (Dense with 26 units and softmax activation)

Adding all these layers together gives us a total of 23 layers in the CNN architecture, that can be shown in the figure 1.

### 3.3 Our Second Approach using CBIR

CBIR stands for Content-Based Image Retrieval. It's a technique used in computer vision and information retrieval that analyzes the visual features of images such as color, texture, shape, and spatial layout to perform tasks like image search, similarity retrieval, and object recognition.

Here using CBIR, the images of book covers are checked if they match with other book covers and the matching probabilities are calculated in percentage. It uses 3-Level matching system to retrieve the closest matches of the book-cover and based on the retrieved book-cover it displays the possible genre using book cover design principles.

**3.3.1 Book Cover Design Principles:** There are certain features for each type of genres. Some of them can be described:

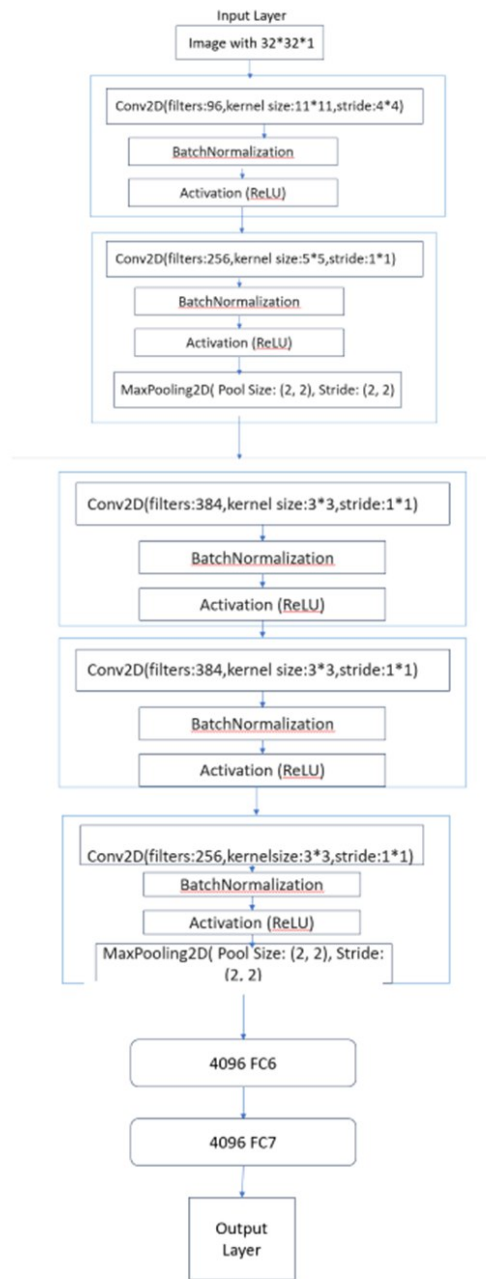


Figure 1: The CNN architecture

- **Color matters:** In the absence of distinguishable features, the CNN has to rely on color alone to classify covers. It's often observed that books of the same genre tend to have similar colors on their covers. For instance, "Self Help" books often sport white covers, while "Religion & Spirituality" books tend to feature yellow. Similarly, "Science & Math" books are commonly adorned with green covers, "Computers &

Technology" books with blue, "Medical Books" with red, and "Biographies & Memoirs" with black.

- **Tone matters:** The tone of the book can define the mood, so "Children's Books" commonly have designs with yellow or bright backgrounds and "Science Fiction & Fantasy" books usually have black or dark backgrounds.
- **Objects matter:** The imagery displayed on book covers often serves as the initial draw for potential readers. Consequently, it's not unexpected that the subject depicted on the cover influences how the book is categorized. For instance, featuring people on the cover is common among many genres, but the type of person or how the person is dressed determines how the book gets classified.
- **Layout matters:** The structure and layout of the book cover also makes a difference in the classification. Books with rectangular title boards, no matter the color, tended to be classified as "Law" and books with a large landscape photographs tended to be "Travel". This trend continued to other categories, such as "Cookbooks, Food & Wine" with a central image of food stretching to the edges of the cover, "Biographies & Memoirs" featuring close-up shots of people, and reference and textbooks containing solid color bands.
- **Text Matters:** The text style on the cover of a book affects the classification, revealing that relationships between text style and genre exist. For example, "Literature & Fiction" often uses expressive fonts to reveal messages about the book.

Book covers that follow design principle, are easily distinguished by this method. Total three level matching system is used here:

- Level-1: RGB Colour Histogram.
- Level-2: Structural Similarity Index Measure (SSIM).
- Level-3: FLANN matching using SIFT features.

It can be shown in the figure 2:

Through the utilization of the pipeline, we identify books that closely resemble the queried book cover, along with the matching percentage. In this procedure, the genre of the most closely matched book ideally aligns with the genre of the queried book cover.

For example, here the book cover of computers and technology is sent for query and it matched with four books which is shown in figure 3. Here the book cover is matched highest with a book in some features which are pointed in figure 4.

For another example, here the book cover of "test preparation" is sent for query and it matched with four books in figure 5. Here the book cover is matched highest with a book in some features which are pointed below in figure 6:

## 4 RESULTS:

In the first approach, after running 20 epochs, the accuracy was very less and it was 3.687% only. As books cover consists of different types of ambiguous things and objects so only running a single CNN model won't surpass it. So we tried CBIR, that performs very accurately for most of the books. As our dataset was very little for this kind of progress, the total accuracy was not calculated. For training the model different aspects we need a huge dataset for this and a high configuration machine to run the model feasibly.

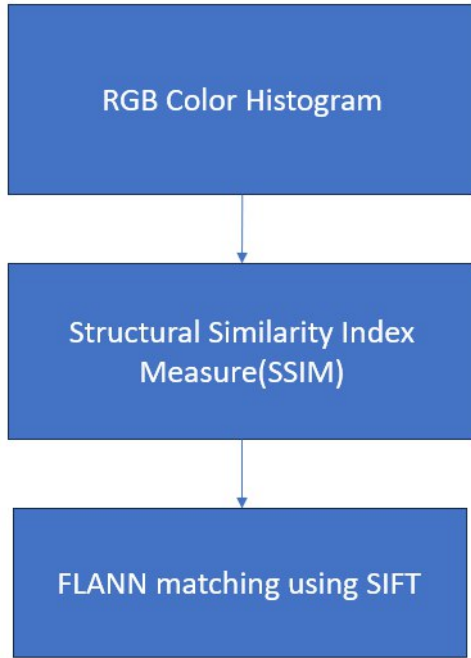


Figure 2: Three level matching system using CBIR



Figure 3: Match of a book cover of "computers and technology"

## 5 DISCUSSION:

From the above methods, it is evident that the first approach utilizing CNNs was not as efficient as anticipated. This inefficiency primarily stemmed from the limited size of our dataset, which hindered the model's ability to generalize effectively. Moreover, many



Figure 4: Feature matching of a book cover of "computers and technology"

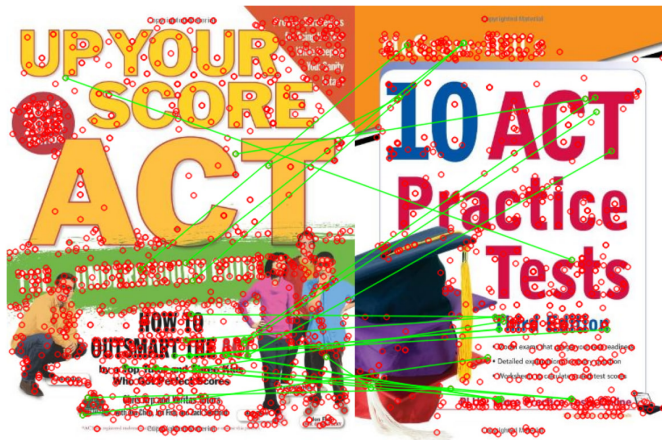


Figure 5: Match of a book cover of "test preparation"

book covers in our dataset exhibited ambiguity and did not adhere to established design principles. These factors contributed to reduced accuracy in genre classification when solely relying on CNNs.

However, upon employing the CBIR approach, we observed significant improvements in our results. By leveraging content-based image retrieval techniques, we were able to identify book covers that closely matched in visual characteristics, indicative of the same genre. This method allowed us to overcome the challenges posed by ambiguous and non-compliant book covers, as it focused on intrinsic visual features rather than relying solely on predetermined design principles.

As a result, we could confidently judge the genre of a book based on its cover, demonstrating the efficacy of the CBIR approach in mitigating the limitations encountered with the CNN-based method.



**Figure 6: Feature matching of a book cover of "test preparation"**

This shift in methodology underscores the importance of adaptability and exploring alternative techniques to address challenges in genre classification tasks.

## REFERENCES

- [1] Seyed Ali Amirshahi and Joachim Denzler. 2017. Judging aesthetic quality in paintings based on artistic inspired color features. In *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE, 1–8.
- [2] Eva Cetinic and Sonja Grgic. 2016. Genre classification of paintings. In *2016 international symposium elmar*. IEEE, 201–204.
- [3] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2015. A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576* (2015).
- [4] Brian Kenji Iwana, Syed Tahseen Raza Rizvi, Sheraz Ahmed, Andreas Dengel, and Seiichi Uchida. 2016. Judging a book by its cover. *arXiv preprint arXiv:1610.09204* (2016).
- [5] Yong-Bae Lee and Sung Hyon Myaeng. 2002. Text genre classification with genre-revealing and subject-revealing features. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*. 145–150.
- [6] Nicolas Scaringella, Giorgio Zoia, and Daniel Mlynek. 2006. Automatic genre classification of music content: a survey. *IEEE Signal Processing Magazine* 23, 2 (2006), 133–141.
- [7] George Tzanetakis and Perry Cook. 2002. Musical genre classification of audio signals. *IEEE Transactions on speech and audio processing* 10, 5 (2002), 293–302.