# CONVOLUTIONAL NEURAL NETWORK-BASED MOVIE POSTER ANALYSIS

*Cole Balacek, Skylar Bi, Xingjian Du*

University of Rochester

## ABSTRACT

For this project our aim was to program a neural network that would take in the poster of a movie via a .jpg file and make a prediction on what the genre and score of the movie would be. To do this we decided to use a resnet18 neural network provided by pytorch. This model was trained using a dataset taken from kaggle.com and while it contained a number of details on each of the many movies it lists we only took into consideration the poster, the score, and genre of each movie.

***Index Terms*—** Cover song identification, instance-batch normalization (IBN), BNNeck, classification loss, triplet loss.

## 1. INTRODUCTION

Movie posters play a critical role in attracting audiences and conveying the general essence of a film. However, designing an effective movie poster that accurately represents a movie's genre and quality remains a challenging task for producers and graphic designers. Our project plans to tackle this issue by developing and training an AI model that is capable of automatically predicting a poster's genre and IMDB score solely based on the poster image. This tool can assist movie producers in creating posters that can effectively communicate the quality and genre of their films. This in turn will improve advertisement for the film and allow the general public to want to view the movie more.

The goal of our project is to create an artificial intelligence that would be able to take in the .jpg file of a movie poster and predict what the score and genre of the film will be. This will allow film studios to use our artificial intelligence model to make sure the posters they put out for their movies are ones that both effectively communicate the genre of the movie and are posters that will cause the general public to associate the movie with other high-quality movies. Doing so will more effectively market the movie to potential consumers, allowing the film to make more money.

In the field of movie genre prediction, significant strides have been made using both non-visual and visual promotional materials. Hoang et al. [1] explored the effectiveness of various machine learning models, including Naive Bayes and recurrent neural networks, and discovered that Gated Recurrent Units (GRU) achieved an accuracy of 80.5% in classifying genres based on plot summaries. Additionally, research by Makita et al. [2] demonstrated that a multinomial Bernoulli event model could predict movie genres from ratings with a 50% success rate. On the visual front, Wehrmann et al. [3] used Convolutional Neural Networks (CNNs) to classify genres from movie trailers, finding that CNNs outperformed traditional methods which constrained feature selection. Motivated by these findings, our project aims to advance CNN architecture to further enhance genre prediction capabilities.

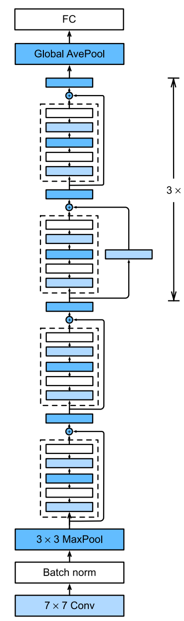


**Fig. 1**. The Structure of ResNet-18

In this project we decided to use a type of convolution neural network for our program. Convolution networks have been used extensively for machine learning algorithms that analyze images so it seemed like a good fit for our project. It works by putting the image through multiple layers of convolution in order to analyze it.

## 2. METHODS AND EXPERIMENTS

The first part of our process is extracting the needed information from our dataset. For our project, we used a dataset from Kaggle containing a large variety of movies. From this dataset we are only using the poster, score, and genre. This

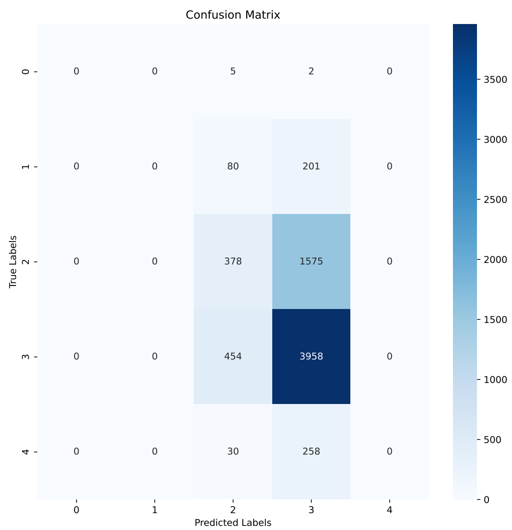


**Fig. 2**. The confusion matrix of movie score prediction

could lead to some problems as things like the director or film studio responsible for the film could have potential effects on the poster but we don't have access to that information and thus can't incorporate it into our algorithm. We have created a pytorch dataset subclass that will extract the poster jpg, score, and genre of each movie in our CSV file.

After that, we put our jpg through a transformer so we can make sure all of our posters have the same dimensions of 224 pixels by 224 pixels, are placed in a tensor, and are viewed in relation to the average color values of an image. We got the average value of an image from ImageNet. We also reduced our score range down from 0-10 to 0-4 by grouping the scores together. Scores from 0-2 were made to be 0, scores from 2-4 were made to be 1, scores from 4-6 were made to be 2, scores from 6-8 were made to be 3, and scores from 8-10 were made to be 4. This will increase the accuracy of our dataset and prevent us from having to deal with float values.

Then we separate our dataset into our training, evaluation, and testing datasets. From there we made an instance of a pytorch resnet18 model to use. A resnet network is a type of convolution neural network that puts the input through multiple layers of convolution to receive a result. However, this causes a problem where the accuracy goes down when processed through too many layers of convolution. The resnet neural network fixes this problem by using residual channels to use the original input throughout the model and in its final decision. We trained the model using our training dataset. We trained it in 10 epochs. As we fit the model we use the evaluation dataset to evaluate it. Once we had achieved our final model we used our test dataset to check its accuracy. We ended with a final accuracy of 66%. Our analysis in figure 2 reveals that scores of 1 and 2, as well as scores of 2 and 3, are more frequently confused in movie score prediction. This observation aligns with intuition, as the boundaries between these adjacent scores are likely less distinct.

| Model | mAP ↑ | mAUC ↑ |
|---|---|---|
| Baseline | 0.125 | 0.661 |
| + Pos Weight | 0.220 | 0.720 |
| + LR Warmup | 0.231 | 0.729 |
| + AdamW | 0.248 | 0.722 |

**Table 1**. Results of the ablation study for movie genre classification, detailing the incremental improvements in model performance metrics mAP (mean Average Precision) and mAUC (mean Area Under the Curve) through the successive introduction of modifications: positive class weighting, learning rate warmup, and AdamW optimizer.

For the genre classification task, we maintain the training settings used in score prediction. However, considering that each movie poster may have between one and four genre labels, we employ binary cross entropy loss for training the model. As shown in Table 2, the results of the ablation study for movie genre classification reveal incremental improvements in performance metrics mAP (mean Average Precision) and mAUC (mean Area Under the Curve). These enhancements result from successive modifications: the introduction of positive class weighting, learning rate warmup, and the AdamW optimizer.

## 3. LIMITATIONS

Due to several outside factors beyond our control that were not in the dataset, we ran into issues. Here are a few of the causes for variability and inaccuracy in our results:

Timing and Trends: The accuracy of our predictions may have been affected by the period in which the posters were released and the leading trends in poster design at that time. Styles, themes, and audience expectations will evolve over time, and our model may trip up on what is considered good in what period, causing issues in accuracy prediction.

Director, Producer, and Cast: The lack of information regarding the director, producer, and cast of the movies shown in the posters could have influenced our model's predictions. These factors play an important role in forming audience expectations and perceptions of a film, which in turn can be portrayed in the poster design. Without this contextual information, our model can struggle to accurately predict genre and score.

Related Intellectual Property (IP): Our dataset could not include information about other intellectual properties (IP) connected to the movies depicted in the posters. Cross-promotional materials, tie-ins with existing franchises, or adaptations of popular works can influence poster design and audience expectations. The lack of this data may have limited the model's ability to make accurate predictions.

Poster Design Team: The team responsible for designing the posters may also have a significant impact on the effectiveness of the genre and quality of the poster. Without information about the design team's style, expertise, and past works, our model may have difficulty recognizing patterns and making accurate predictions based on the poster design alone.

In summary, the absence of key contextual information such as release timing, director, producer, cast, related IP, and poster design team details may have caused variability and inaccuracy in our model's predictions.

## 4. CONCLUSION

In conclusion, our project demonstrates the feasibility of using AI to predict movie scores based on poster images. While the accuracy of the program was not spectacular, it was reasonable given the limited information we were able to provide to the program. Nevertheless, we were able to successfully create a neural network that was able to train itself, evaluate its own training, and test its accuracy using the provided dataset.

## 5. CONTRIBUTIONS OF AUTHORS

The following outlines the specific contributions of each author to the development and execution of this project.

- **Xingjian Du**: Engaged in the implementation of Score Prediction and Genre Classification. Contributed to the writing of the Methods and Experiments sections.
- **Cole Balacek**: Engaged in the implementation of Score Prediction. Contributed to the writing of the Introduction, Methods, and Experiments sections.
- **Skylar Bi**: Engaged in the implementation of Genre Classification. Contributed to the writing of the Methods, Experiments and Limitations sections.

## 6. REFERENCES


[1] Quan Hoang, "Predicting movie genres based on plot summaries," *arXiv preprint arXiv:1801.04813*, 2018.

[2] Eric Makita and Artem Lenskiy, "A multinomial probabilistic model for movie genre predictions," *arXiv preprint arXiv:1603.07849*, 2016.

[3] Jônatas Wehrmann and Rodrigo C Barros, "Convolutions through time for multi-label movie genre classification," in *Proceedings of the Symposium on Applied Computing*, 2017, pp. 114–119.