Anime Genre Classification Using CNNs

Megan Chen MYC9078@NYU.EDU
Mimi Chen XC2063@NYU.EDU
Jack Lee JL11862@NYU.EDU

Milestone 1

1. Introduction and Motivation

The topic we are exploring is classifying anime genres based on their posters. In the world of anime, posters have long served as the gateway to entice audiences and provide valuable insight into the genre and theme of the show or film. We want to see if there are specific features of a poster that correspond to a certain genre. Since anime may belong to multiple genres, this would require a multi-label image classification model. Given the vast number of anime series produced annually, it would be fascinating to discern the recurring features that define each genre. Furthermore, since posters are typically the initial promotional materials for new anime releases, this analysis could potentially provide insights into the themes and content of these series.

2. Related Works

Movie Genre Classification based on Poster Images with Deep Neural Networks presented a system to classify movie poster images into genres. In this paper, the researchers utilized a deep neural network to consider visual appearance and object information, then a classifier was constructed to estimate the probabilities of a poster belonging to different genres. The model showed promising performance, but the classification performance can be improved upon by utilizing the metadata to investigate more movie properties.

A Study on CNN Transfer Learning for Image Classification propose a novel deep network structure to help the enhancement of a model by using micro-networks to stride over input images to produce a feature map. Some challenges faced by the researchers that we should take into considerations are computational power and time. As the epochs increased, the classification accuracy improved at the cost of considerable training time. The researchers proposed the utilization of GPU as opposed to CPU, which we will follow to improve our model accuracy and time efficiently.

Predicting Genre from Movie Posters explores whether elements of a poster allows a model to predict the movie's genre, implementing a Multi-Label K-Nearest Neighbors (ML KNN) and ResNet34 trained on the ImageNet Dataset. Some issues with this research include an imbalanced dataset, which included an overwhelming amount of dramas and very few TV movies. This paper also faced a memory and computational limitations that may have negatively affected the results.

3. Methodology

Data Augmentation In the case that our model overfits to our anime poster dataset, we are going to experiment with different data augmentation methods, like cropping the images or adjusting brightness/contrast of images to create new ones, thus synthetically increasing the amount of training data we have. Since our dataset only has 24k items, this will help introduce more diversity into the training data and avoid overfitting.

Learning Phase Referencing the process of Hossain et.al., our system consists of two states: training and testing.

1. *Training State*
   (a) The initial step of our Convolution Neural Network is converting the anime posters into a feature space to perform feature extraction. Our proposed model consists of **** layers. detail
   (b) For extracting features, the operation in the convolution layer is
      i. equation
   (c) In our kjfsdkfjds f
      i. equation
   (d) Finally, a fully connected layer of our network is represented as
      i. equation
2. *Testing State*
   (a) We evaluate the performance of

As shown by the study done by Barney and Kaya, ResNet34 was successful at predicting a movie poster's genres. Assuming that the ResNet architecture will also successfully predict anime poster's genres, we will base our model on ResNet50, pretrained on ImageNet-1K and finetuned on the anime dataset, experimenting with different frozen layers. We will also train a custom CNN model and compare the results of this to the ResNet50 model.

Evaluation The metrics we will use to evaluate the performance of our model are F1 score, recall, and top-k categorical accuracy.

4. Dataset

We will use a Kaggle Dataset from 2023 consisting of 24,905 posters. Each poster has an id, name (in original language), English name, other name, rating of anime, genres of the anime, synopsis, type (movie or tv series), number of episodes, and date aired.

5. Work Plan

Up until now, all team members have collectively decided on the project topic as well as the methodology we will implement. The upcoming weekly tasks will be distributed accordingly:

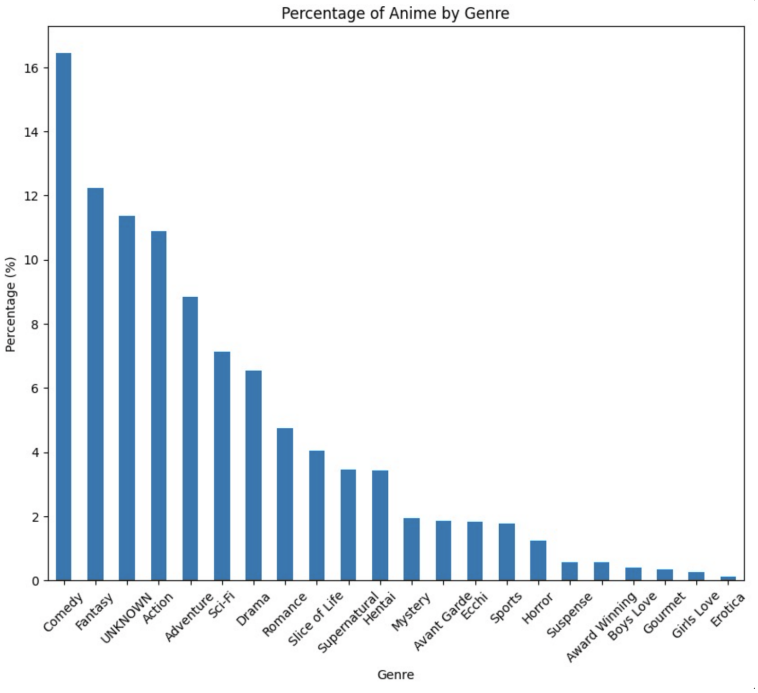

Figure 1: Class Distribution of Anime Dataset

Week	Megan
11/6	Finish cleaning data
11/13	Read and research about self-supervised and visually grounded pretraining
11/20	Understand architecture of ResNet-50
11/27	Evaluate the new ResNet-50 model on a validation set
12/4	Build initial CNN model
12/11	Evaluate CNN Model
12/18	Clean everything, make video, upload to GitHub

Week	Mimi
11/6	Finish cleaning data
11/13	Read and research about movie genre classification
11/20	Load ResNet-50 with weights pre-trained on ImageNet
11/27	Unfreeze some ResNet layers and train on the anime poster dataset
12/4	Use dataset to train the CNN
12/11	Compare CNN Model with the ResNet-50 Model
12/18	Clean everything, make video, upload to GitHub

Week	Jack
11/6	Finish cleaning data
11/13	Finish visualizing and preprocessing data (resizing, augmentation)
11/20	Pass the anime poster dataset through the ResNet-50 model to extract features
11/27	Attach new fully connected layers at the end of the ResNet-50 model
12/4	Evaluate the performance of the model
12/11	Test both models on unseen anime posters
12/18	Clean everything, make video, upload to GitHub

Figure 2: Sample of previous research results of object detection for movie posters (Chu & Guo, 2017

References

[1] Hussain, M., Bird, J.J., and Faria, D.R. (2019). A Study on CNN Transfer Learning for Image Classification. In *Advances in Computational Intelligence Systems* (Vol. 840, Advances in Intelligent Systems and Computing, pp. 16). Cham: Springer.

[2] Barney, G., and Kaya, K. (2019). Predicting Genre from Movie Posters. Note: Unpublished article.

[3] Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K., and Ghayvat, H. (2021). CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope. *Electronics*, 10(20), 2470.

[4] Hossain, N., Ahamad, M. M., Aktar, S., and Moni, M. A. (2021). Movie Genre Classification with Deep Neural Network using Poster Images. In *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)* (pp. 195-199). Dhaka, Bangladesh: IEEE.

[5] Unal, F. Z., Guzel, M. S., Bostanci, E., Acici, K., and Asuroglu, T. (2023). Multilabel Genre Prediction Using Deep-Learning Frameworks. *Applied Sciences*, 13(15), 8665.

[6] Wehrmann, J., and Barros, R. C. (2017). Movie genre classification: A multi-label approach based on convolutions through time. *Applied Soft Computing*, 61, 973-982.

[7] Chu, Wei-Ta, and Guo, Hung-Jui. (2017). Movie Genre Classification based on Poster Images with Deep Neural Networks. In *Proceedings of the Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes (MUSA2 '17)* (pp. 39–45). New York, NY, USA: ACM.