Movie genres prediction from poster

Alaa Shehab & Bassant Ahmed

Faculty of Engineering, Alexandria university

Introduction

A movie poster can convey a lot of details about the movie.

By seeing only the poster

- We can tell the genre and theme of the movie.
- We can predict the movie's rating and revenue.

In this project we propose a multi labeled movie genre classification based on posters analysis, which is achieved through extracting elements from the poster such as the poster color, facial expression, objects and many more.

The goal is to predict movie genre/genres by seeing only its poster.

Prediction (for Jumanji)
Action – Adventure - Comedy

Dataset

We used 'The Movies Dataset' on Kaggle, which is an ensemble of data collected from TMDB and GroupLens.

- The dataset contains metadata for all 45,466 movies listed in the Full MovieLens Dataset.
- Each entry in the dataset contains info about the movies, as shown below, from title, genres, rating, release year, poster, etc.

	adult	budget	id	imdb_id	original_title	popularity	poster_path	release_date
0	False	30000000	862	tt0114709	Toy Story	21.9469	/rhIRbceoE9lR4veEXuwCC2wARtG.jpg	1995-10-30
1	False	65000000	8844	tt0113497	Jumanji	17.0155	/vzmL6fP7aPKNKPRTFnZmiUfciyV.jpg	1995-12-15
2	False	0	15602	tt0113228	Grumpier Old Men	11.7129	/6ksm1sjKMFLbO7UY2i6G1ju9SML.jpg	1995-12-22

Data Preprocessing

We preprocessed the dataset as follows:

- Removed entries with NAN release date, posters or genres
- Removed all Grayscale movies and movies earlier than 1980
- Resized images to (24, 24) RGB images
- Pre-processed images by scaling pixels

Total number of movies for training and validation after preprocessing are 31538 image

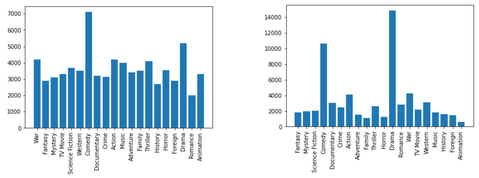
Figures shows a sample of images in the dataset

Data Augmentation

We added new images to balance the dataset and avoid overfitting by

- Random Zoom Augmentation
- Random Brightness Augmentation
- Random Rotation Augmentation

Total number of movies for training and validation after augmentation are 44182 image



Figures shows the dataset before and after augmentation

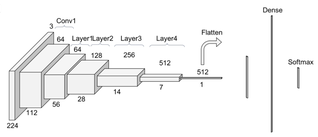
Models

Baseline ResNet34

Used pretrained ResNet34 network on imagenet and split data into 2 splits (22000, 22000)

We replaced the final softmax layer with a sigmoid layer with 20 classes and changed the loss function to Binary Cross Entropy Loss.

We trained using sorted/randomized data before and after data augmentation



ResNet18 - ResNet50V2 - VGG16

As ResNet34 was overfitting the data we used transfer learning on these models as well and split data into 2 splits (22000, 22000)

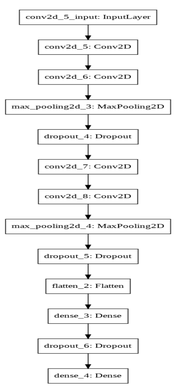
Using final layer as sigmoid, changing the loss function to Binary Cross Entropy Loss and using optimizer as Adam.

Custom Architecture

We used a simple CNN architecture with 11 layers

- 2 Maxpool Layers
- 3 Dropout Layers
- 512 Dense Layer

Used Adam optimizer and loss function as binary Cross entropy.



Discussion

- We trained on both 31k and 44k image datasets and had to split the data into 2 splits equally.
- All Models used binary cross entropy loss function with Adam and SGD(0.001) optimizers and having from 15-300 epochs where batch size is 32.
- Models performance was affected negatively by the following :
  - Splitting data into 2 splits instead of training on entire dataset.
  - Learning irrelevant features from the posters such as the words written on the poster
  - Having some images with bad resolution
  - Not having cartoon/different culture movies removed from dataset.
- All models were trained on both datasets of 31538 images and 44182 images, splitting data into 75% - 25% training and validation respectively.
- Splitting 44k images into 2 splits trained separately by loading model weight, achieved better accuracy than splitting 44k into 4 splits.

Results

Baseline model (Figure 1)
Training ResNet34 on 31538 image (Overfitting)
Reaching accuracy 88% on validation

Figure 1

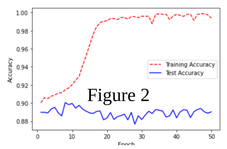


Figure 2

ResNet34 (Figure 2)
Trained on 44182 image (Overfitting)
Reaching accuracy 89% on validation

ResNet34(Figure 3)
Trained on 44182 image
Early stopping to prevent overfitting
Reaching accuracy 90.34% on validation

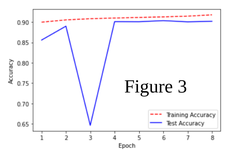


Figure 3

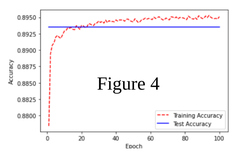


Figure 4

Custom Architecture (Figure 4)
Trained on both 31538 and 44182 dataset.
Best accuracy reached was 89.35%

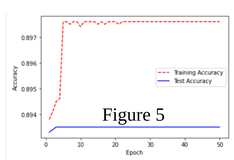


Figure 5

VGG16(Figure 5)
Trained on both 31538 and 44182 dataset.
Best accuracy reached was 89.35%

Conclusion

Before data augmentation

- ResNet34 overfitted the data performing 99.9% accuracy on training and 88% on validation
- VGG16 and the custom model achieved an accuracy of 91% on training and 89.35% on validation
- Although ResNet34 was overfitting, it overcame the vanishing gradient problem, where both VGG16 and the custom model were stuck at 89.35% validation accuracy after running 300 epoch with all optimizers, due to using residual blocks.

After data augmentation

- VGG16 and custom model had no change in performance
- ResNet34 had the best validation accuracy of 90.34% after running 8 epochs (early stopping to avoid overfitting) using 2 data splits each of 22k image.
- Predicting genres from posters requires additional pre-processing on images to ensure resolution of images and consistency of dataset.

Future Work

- Preprocessing data to remove low resolution images and movies that are inconsistent with the dataset i.e. cartoon movies, Indian and Turkish movies, etc.
- Training 44k images in 1 pass without splitting is seen to have achieved better result than trying to split them and retraining using saved weights.
- Use GANs or variational AED to generate images and increase dataset instead of using keras ImageDataGenerator to rotate, change brightness , and zoom int/out.

References

[1] The Movies Dataset. https://www.kaggle.com/rounakbanik/the-movies-datase

[2] Kos, Pobar & Ipsic - Automatic Movie Posters Classification into Genres 2015. https://www.researchgate.net/publication/282196711_Automatic_Movie_Posters_Classification_into_Genres

[3] Nirman Dave, Hampshire college – predicting movie genre from poster. https://github.com/nddave/Movie-Genre-Prediction/blob/master/paper/predicting-movie-genres-paper.pdf

[4] Chu & Guo - Movie Genre Classification based on Poster Images with Deep Neural Networks 2017. https://www.cs.ccu.edu.tw/~wtchu/papers/2017MUSA-chu.pdf

[5] ResNet34 pretrained on imagenet and compatible with keras https://github.com/qubvel/classification_models