Movie Genre Classification

Alaa Shehab
Bassant Ahmed


LUCK HAS NOTHING TO DO WITH IT.
lucky number slevin
PULP FICTION
LEON
UN FILM DE
tom hanks
saving private ryan
FROM THE VISIONARY DIRECTOR OF '300'
WATCHMEN
03.06.09
PANS LABYRINTH
GARY OLDMAN
BRUCE WILLIS
MILLA JOVOVICH
THE FIFTH ELEMENT
There is no future without it.
GO FOR THE KILL
KILL BILL
THE FOURTH FILM BY QUENTIN TARANTINO
MEL GIBSON
BRAVEHEART
OPENS WEDNESDAY, MAY 24
I AM LEGEND
DIE HARD 4.0
WEDNESDAY JULY 4
SEAN CONNERY
DIE LIGA DER AUSSERGEWÖHNLICHEN GENTLEMEN
Vom Regisseur von BLADE
FRANK MILLER'S
SIN CITY
PITCH BLACK
PLANET DER FINSTERNIS

Problem Statement

- Multi labeled movie genre classification based on posters analysis, which is achieved through extracting some elements from the poster such as the color, facial expression, objects and many more.

Dataset

- Kaggle dataset, contains 45,466 image collected from TMDB.
- Each entry in dataset contains info about the movie
  As shown below.

	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

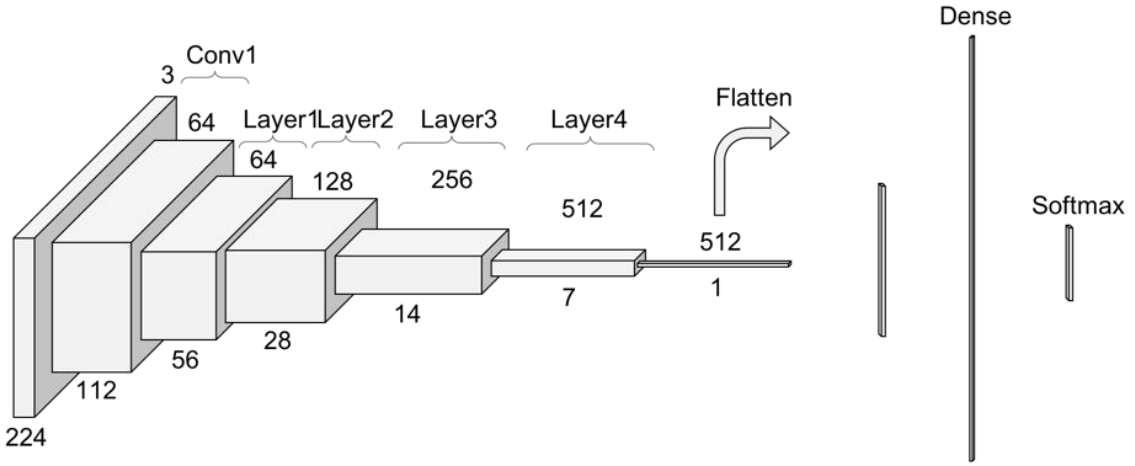
Data preprocessing

- Removed entries with NAN -> release date, posters or genres
- Removed all Grayscale movies and movies earlier than 1980
- Resized images to (224, 224) RGB images
- Pre processed images by scaling

Total number of movies for training and validation after preprocessing are 31k

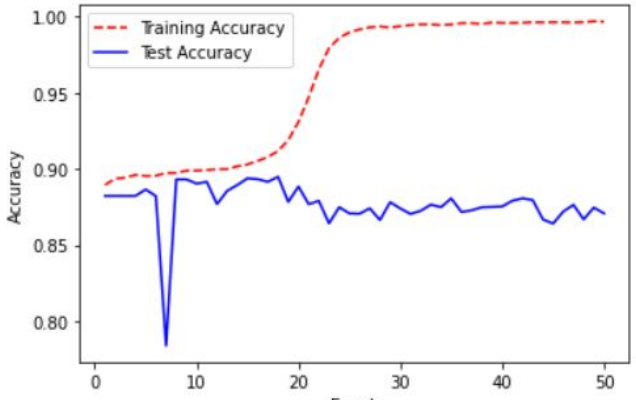
Baseline model (ResNet34)

- Trained on 31538 image
- Using Adam optimizer and loss function -> Binary cross-entropy

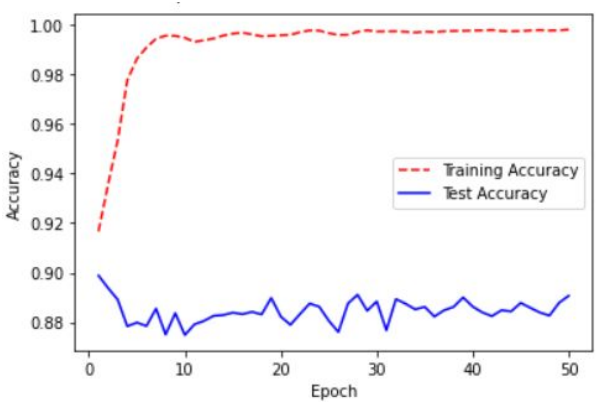


Results for baseline model

- **After running 50 epochs model reached an accuracy of 88% on validation dataset and overfitted the data.**



Split 1 of data - 15k



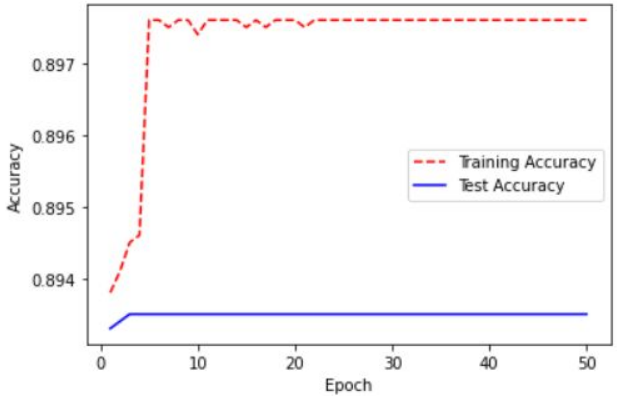
Split 2 of data - 15k

Proposed solutions

- Use ResNet18 / VGG16 to avoid overfitting data
- Use ResNet50 -> as recommended in some papers
- Increase dataset by data augmentation

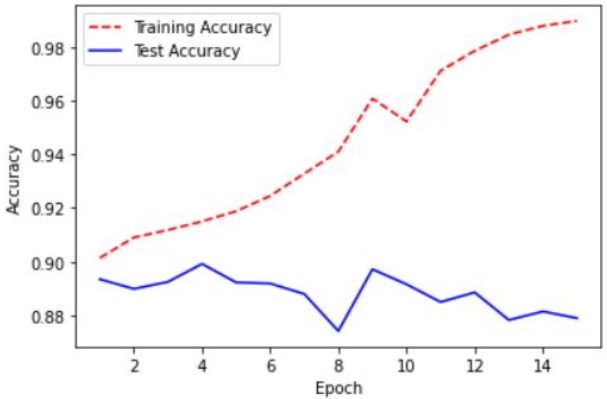
Transfer learning using VGG16

- After running 50 epochs model was stuck at an accuracy of 89.35% on validation dataset.
- Running VGG16 on 300 epoch didn’t change the results



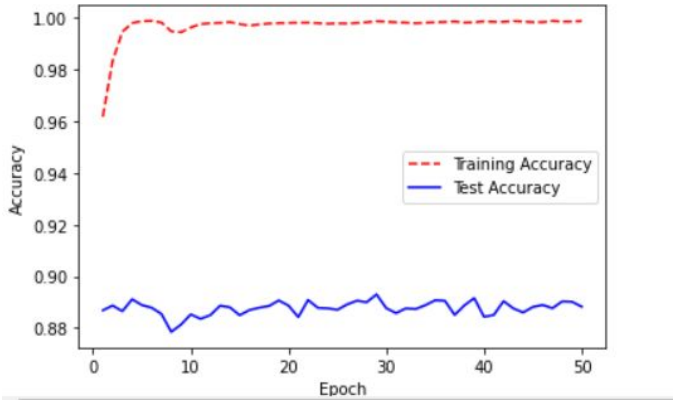
Transfer learning using ResNet18

- Early stopping when reaching epoch 15 due to overfitting
- Accuracy reached for validation dataset 87%



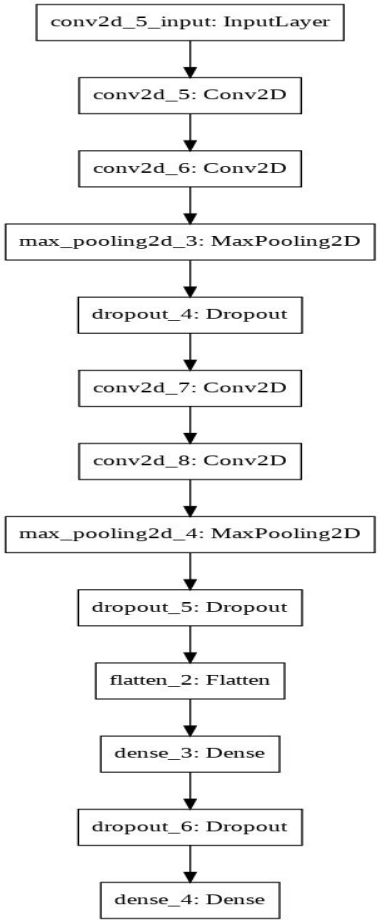
Transfer learning using ResNet50

- Overfitting data
- Accuracy reached for validation dataset 89%



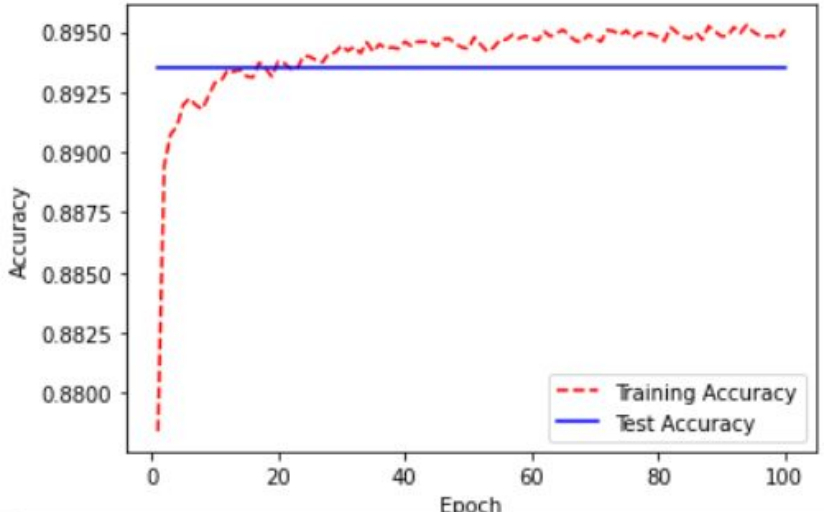
Custom Model

- Simple CNN model - 11 layers
- 2 Maxpool layers
- 3 Dropout layers to avoid overfitting



Custom Model results

- Stuck at 89.35% validation
- Gradient vanishing



Data Augmentation

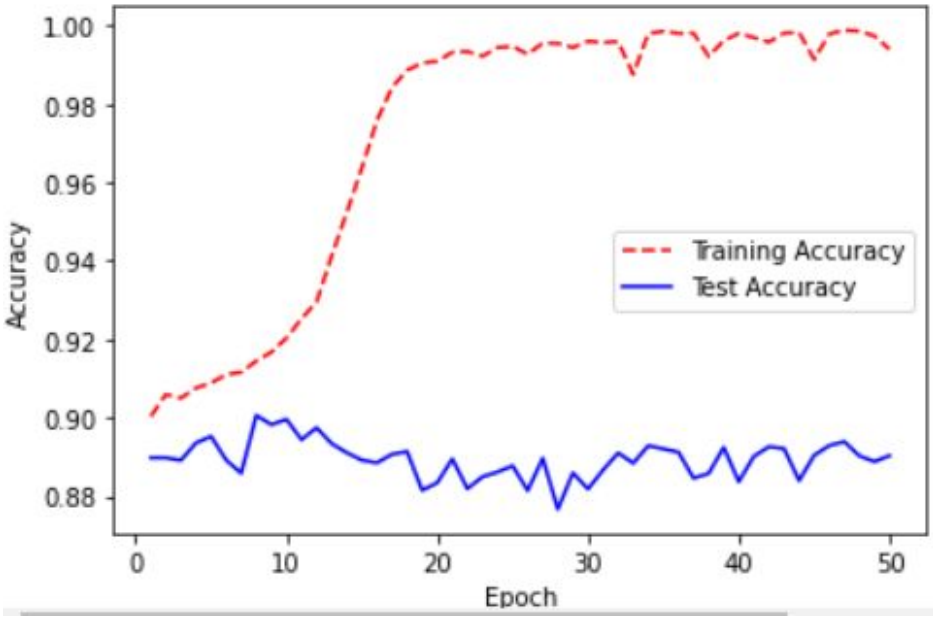
To increase and balance the dataset

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

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

Results after data augmentation & normalization

ResNet34 reached better accuracy 90.34% with early stopping at epoch 12



Team contribution

Team contribution

Bassant

- Baseline model deployment
- Splitting dataset into train and validation
- Creating data generator
- Data augmentation

Alaa

- Baseline model deployment
- Splitting data into 4 splits (sorted/randomized data)
- Generating data/labels csv from downloaded images
- Custom model architecture and deployment

Conclusion

Before data augmentation :

- ResNet34 overfitted the data performing 99.9% 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.

Conclusion

After augmentation and addition preprocessing

- 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.

Observations

- 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 consistent data ex : cartoon/different culture movies removed from dataset.

Observations

- ResNet34 had best performance
  - Although it overfitted the data, it achieved the highest accuracy of 90.34%
  - Overcame problem of vanishing gradient due to using residual blocks

Future work

- Preprocessing data to remove low resolution images and inconsistent data
- 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 AE to generate images and increase dataset instead of using keras ImageDataGenerator .

Demo

References

[1] The Movies Dataset. https ://www.kaggle.com/ the movies datase
[2] Kos, Pobar & Ipsic Automatic Movie Posters Classification into Genres.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 https://www.cs.ccu.edu.tw/~wtchu/papers/2017MUSA chu. pdf
[5] ResNet34 pre trained on imagenet and compatible with keras https://github.com/qubvel/classification_models