



Machine Learning & Cinema: Color grading style classification through Vectorscopes and CNNs

Master's Degree in Computer Engineering



Politecnico
di Torino



Table of Contents

- ▶ Introduction
- ▶ Data Preprocessing
- ▶ Experiments and Results
- ▶ Conclusions



Introduction

Vectorscope

The goal of the project is exploring the possibility of implementing vectorscopes in classification of movie frames by the Genre.



Figure 1: Raw image

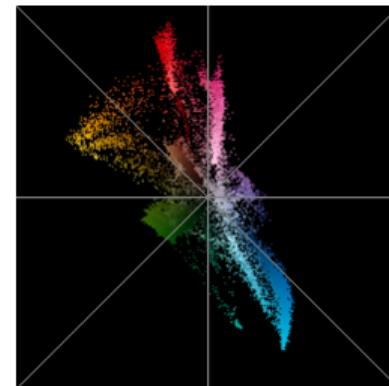


Figure 2: Vectroscope image



Introduction

CNNs

To predict the genre of a movie I'll be using various models.

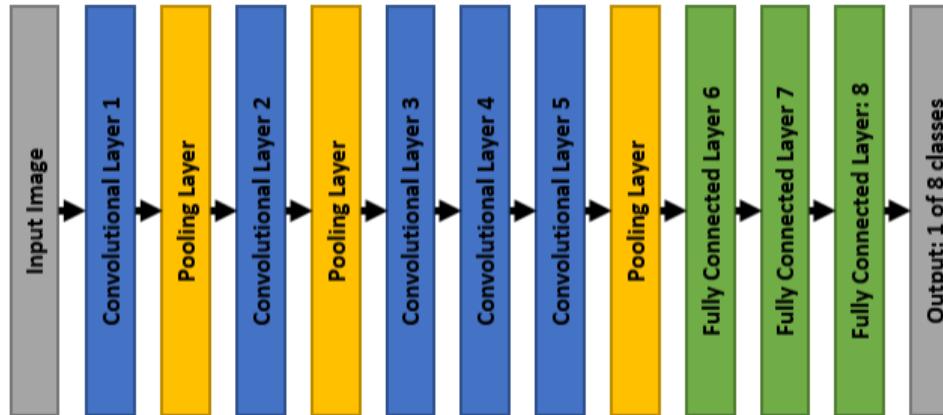


Figure 3: AlexNet architecture



Table of Contents

- ▶ Introduction
- ▶ Data Preprocessing
- ▶ Experiments and Results
- ▶ Conclusions



Data Preprocessing

Data was available thanks to [CondensedMovies repository](#).

videoid	year	clip_idx	clip_tot	upload_year	title	clip_name	imdbid
---------	------	----------	----------	-------------	-------	-----------	--------

Table 1: clips.csv header

imdbid	genre	country
--------	-------	---------

Table 2: movie_info.csv header



Data Preprocessing

Removing unavailable and gray scale clips

The final table contained 27076 links to YouTube video clips.

title	clip_name	year	videoid	genre
-------	-----------	------	---------	-------

Table 3: movieclips_all.tsv header

After verifying the validity of each video link, new links were appended to a new .tsv file, resulting in 17,723 movie clips.

In the final preprocessing step, I removed black and grey frames as they wouldn't be useful for vectorscope experiments, ultimately obtaining 16,262 links.



Data Preprocessing

Limiting the dataset

- Action and Adventure became Act_Adv,
- Thriller, Horror and Crime became Thr_Hor_Cri,
- Comedy and Romance became Com_Rom,
- Drama remained as it was.

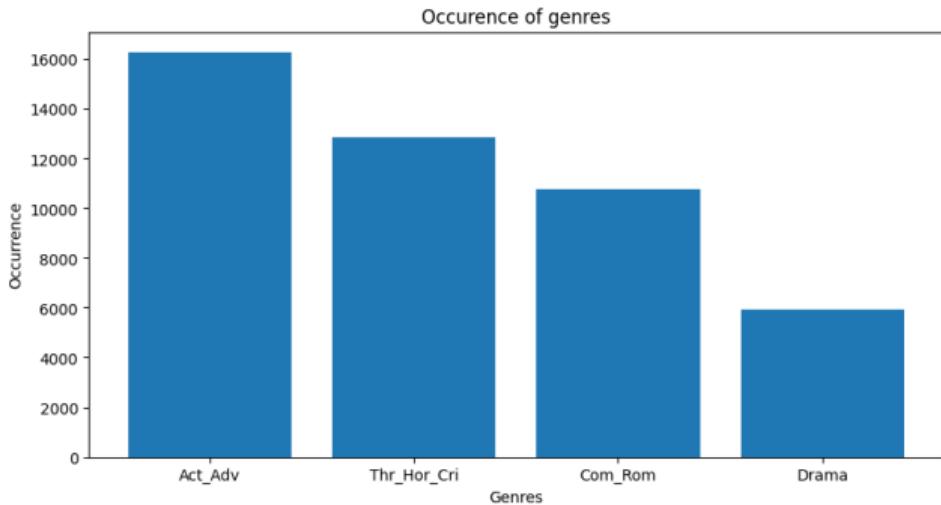


Figure 4: Data distribution after merging classes



Data Preprocessing

Binary classification

Two biggest classes:

- Action and Adventure - Act_Adv,
- Thriller, Horror and Crime - Thr_Hor_Cri.

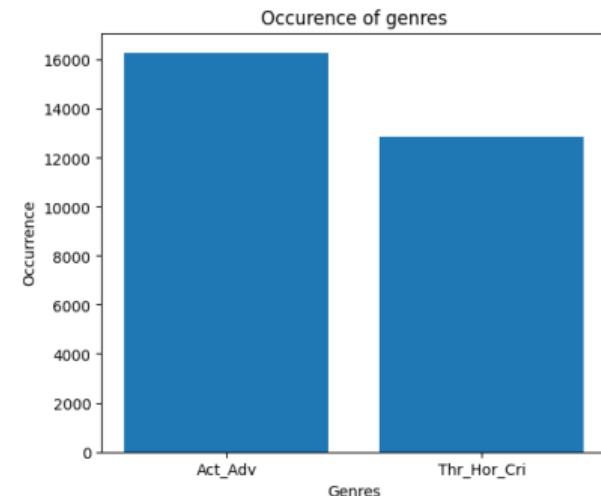


Figure 5: Data distribution for binary classification



Data Preprocessing

Collecting the data

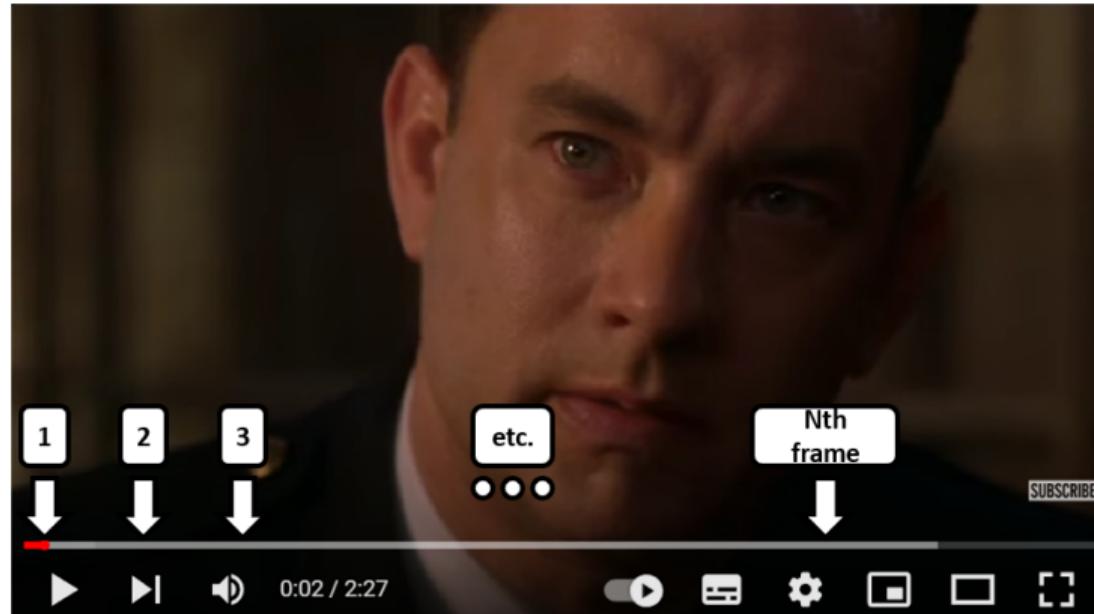


Figure 6: Frame sampling



Data Preprocessing

Training and testing split

“Random” data split:

- Frames from one clip are allowed to be present in both training and testing set,
- Less realistic scenario.

“Separate” data split:

- Frames from one clip are NOT allowed to be present in both training and testing set,
- More realistic scenario.



Table of Contents

- ▶ Introduction
- ▶ Data Preprocessing
- ▶ Experiments and Results
- ▶ Conclusions



Experiments and Results

Model selection

- Raw frames - ResNet-50
- Vectorscope representations - VGG-16

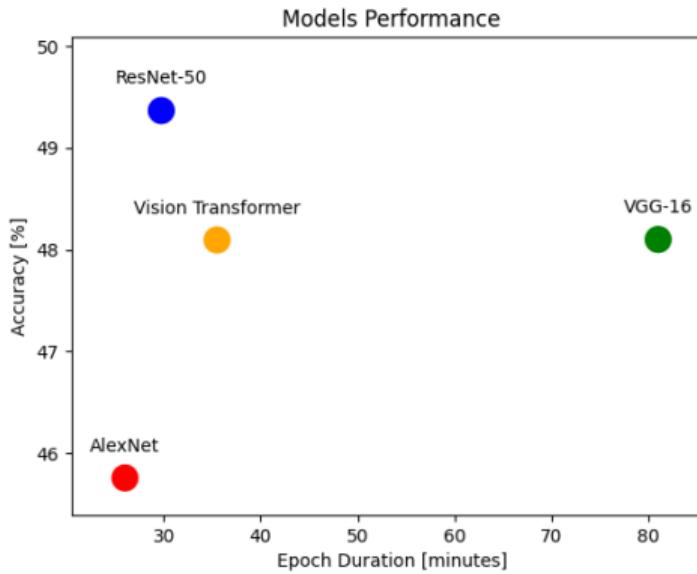


Figure 7: Model accuracies in relation to the duration of each epoch



Experiments and Results

Binary classification - raw frames

Raw frames,
“random” distribution, ResNet-50:

- Accuracy: 76.87%

Genre	Act_Adv	Thr_Hor_Cri
Act_Adv	1921	565
Thr_Hor_Cri	448	1446

Table 4: Confusion matrix

Genre	Precision	Recall	F1-score
Act_Adv	0.811	0.773	0.791
Thr_Hor_Cri	0.719	0.763	0.741
weight_avg	0.771	0.769	0.769

Table 5: Results



Experiments and Results

Binary classification - vectorscope representations

Vectorscope representations,
“random” distribution, VGG-16:

- Accuracy: 63.49%

Genre	Act_Adv	Thr_Hor_Cri
Act_Adv	1612	874
Thr_Hor_Cri	725	1169

Table 6: Confusion matrix

Genre	Precision	Recall	F1-score
Act_Adv	0.690	0.648	0.668
Thr_Hor_Cri	0.572	0.617	0.594
weight_avg	0.639	0.635	0.636

Table 7: Results



Experiments and Results

Binary classification - “separate” distribution

Raw frames,
“separate” distribution, ResNet-50:

- Accuracy: 70.22%

Genre	Precision	Recall	F1-score
Act_Adv	0.744	0.691	0.717
Thr_Hor_Cri	0.659	0.715	0.686
weight_avg	0.705	0.702	0.703

Table 8: Results

Vectorscope representations,
“separate” distribution, VGG-16:

- Accuracy: 60.09%

Genre	Precision	Recall	F1-score
Act_Adv	0.649	0.583	0.614
Thr_Hor_Cri	0.555	0.622	0.587
weight_avg	0.606	0.601	0.602

Table 9: Results



Experiments and Results

4 classes - "random" distribution

Raw frames,
"random" distribution, ResNet-50:

- Accuracy: 60.54%

Genre	Precision	Recall	F1-score
Drama	0.534	0.449	0.488
Act_Adv	0.677	0.688	0.682
Com_Rom	0.592	0.576	0.584
Thr_Hor_Cri	0.554	0.574	0.568
weight_avg	0.604	0.605	0.604

Table 10: Results

Vectorscope representations,
"random" distribution, VGG-16:

- Accuracy: 45.73%

Genre	Precision	Recall	F1-score
Drama	0.332	0.142	0.199
Act_Adv	0.537	0.508	0.522
Com_Rom	0.444	0.494	0.468
Thr_Hor_Cri	0.409	0.508	0.453
weight_avg	0.453	0.457	0.448

Table 11: Results



Experiments and Results

4 classes - "separate" distribution

Raw frames,
"separate" distribution, ResNet-50:

- Accuracy: 51.55%

Genre	Precision	Recall	F1-score
Drama	0.341	0.325	0.333
Act_Adv	0.621	0.592	0.606
Com_Rom	0.478	0.487	0.482
Thr_Hor_Cri	0.486	0.513	0.500
weight_avg	0.517	0.516	0.516

Table 12: Results

Vectorscope representations,
"separate" distribution, VGG-16:

- Accuracy: 43.81%

Genre	Precision	Recall	F1-score
Drama	0.330	0.187	0.239
Act_Adv	0.476	0.437	0.456
Com_Rom	0.405	0.428	0.416
Thr_Hor_Cri	0.444	0.537	0.486
weight_avg	0.434	0.438	0.432

Table 13: Results



Experiments and Results

Misclassified frames



Figure 8: Example of Act_Adv frame classified as Com_Rom



Experiments and Results

Misclassified frames



Figure 9: Example of Act_Adv frame classified as Thr_Hor_Cri



Experiments and Results

Misclassified frames



Figure 10: Example of Com_Rom frame classified as Drama



Experiments and Results

Misclassified frames



Figure 11: Example of Thr_Hor_Cri frame classified as Com_Rom



Table of Contents

- ▶ Introduction
- ▶ Data Preprocessing
- ▶ Experiments and Results
- ▶ Conclusions



Conclusions

- Categorizing a movie genre using only a single frame is challenging, but the inclusion of supplementary information could make it more achievable.
- Experiments with vectorscope representations of images show encouraging outcomes, not only in terms of computational efficiency but also in their adaptability.



Machine Learning & Cinema: Color grading style classification through Vectorscopes and CNNs

*Thank you for listening!
Any questions?*