



Progetto di Analisi Immagini e Video

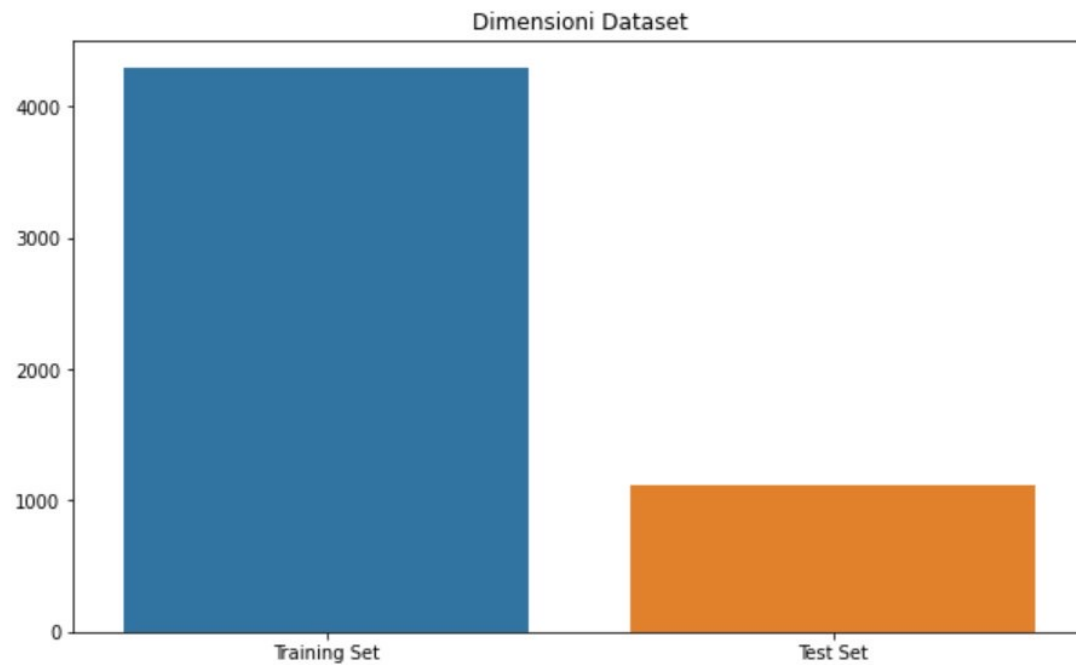
Alati Maria Francesca
Galimberti Giulia Katia

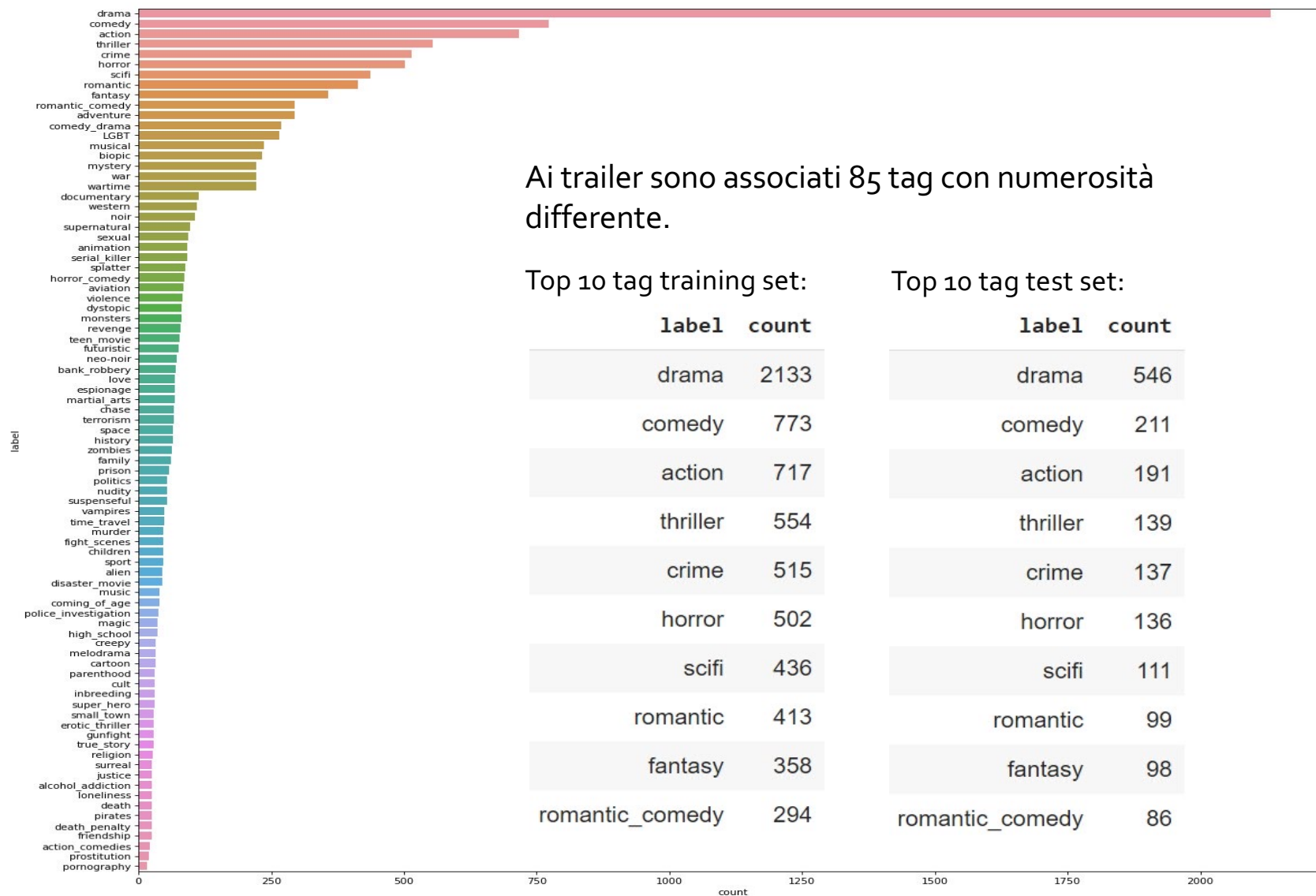
Introduzione

- Obiettivo: Classificazione multi-classe e multi-label di trailer di film
- Struttura del progetto:
 1. Data analysis
 2. Loading the dataset
 3. Loading the model
 4. Train
 5. Plot the loss
 6. Test Suite

Data analysis

- Il dataset utilizzato contiene 5404 trailer diviso in:
 - 4292 Training set
 - 1112 Test set





Loading the dataset

Dataset:

- Frame 100 x 100 pixel
- 15 frame per trailer

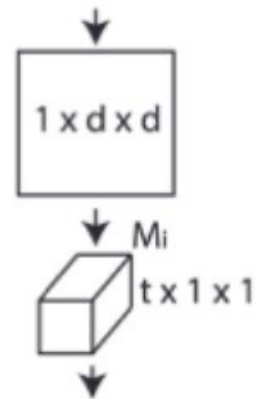
Dataloaders:

- Batch size = 4





3D conv



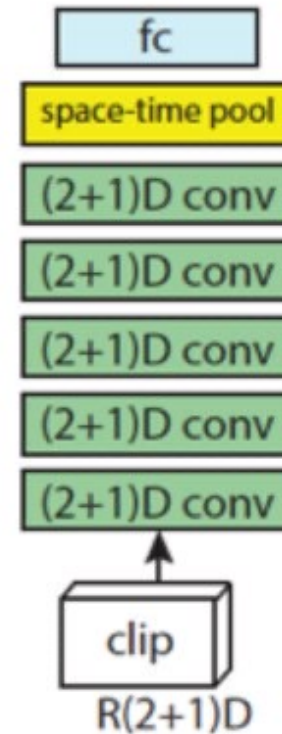
(2+1)D conv

R(2+1)D Model

L'idea fondamentale alla base di questa architettura è approssimare una convoluzione 3D utilizzando una convoluzione spaziale 2D seguita da una convoluzione temporale 1D.

R(2+1)D Model

- La R(2+1)D è composta da cinque (2+1)D blocchi convoluzionali.
- Poiché ogni blocco è composto da una convoluzione spaziale e una convoluzione temporale, la non linearità è doppia rispetto a una CNN 3D composta con lo stesso numero di blocchi.



Parametri:

```
lr = 1e-4

criterion = nn.BCEWithLogitsLoss()

optimizer = optim.SGD(train_params, lr=lr, momentum=0.8, weight_decay=5e-3)

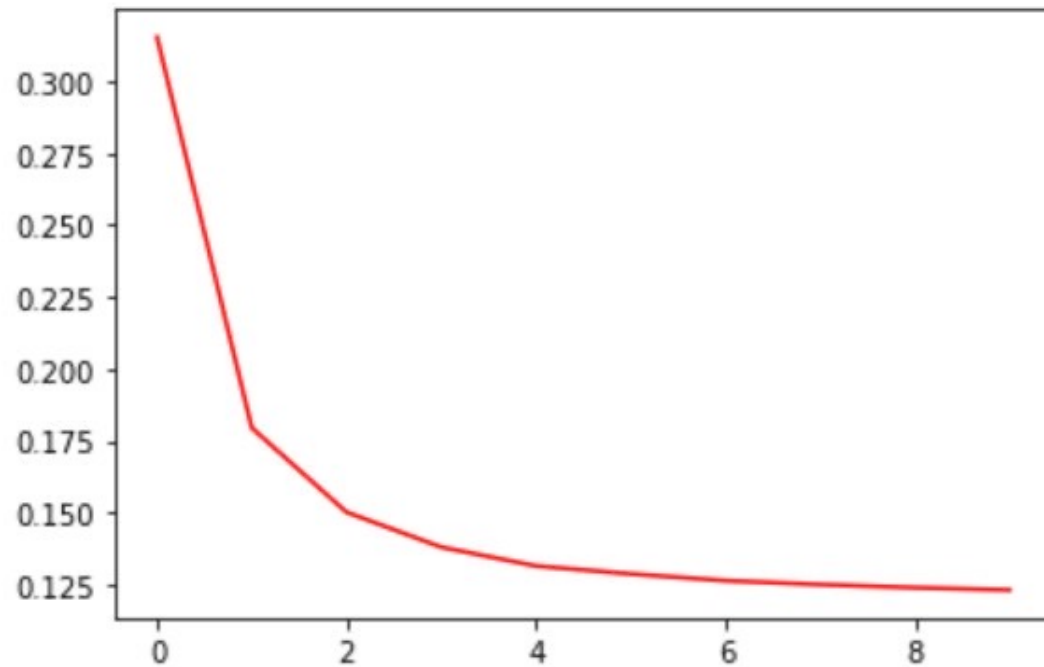
scheduler = optim.lr_scheduler.StepLR(optimizer, step_size=4, gamma=0.1)

epochs = 10
```

Train

Plot the loss

Di seguito si mostra l'andamento della Loss nelle 10 epoche eseguite



Test

Per la valutazione del modello è stato utilizzato il test set, il seguente Classification Report mostra i valori ottenuti

	precision	recall	f1-score	support
micro avg	0.16	0.56	0.25	3167
macro avg	0.02	0.12	0.03	3167
weighted avg	0.14	0.56	0.21	3167
samples avg	0.16	0.59	0.24	3167

The image features two teal-colored geometric shapes. On the left is a tall, narrow triangle pointing downwards. On the right is a large, solid rectangle. The text 'Grazie per l'attenzione!' is centered between these two shapes.

Grazie per l'attenzione!