

Analisi di immagini e video

Progetto: classificazione multiclasse e multilabel di trailer



Dipartimento di Ingegneria Informatica Modellistica
Elettronica e Sistemistica

Docenti

Prof. Giuseppe Manco

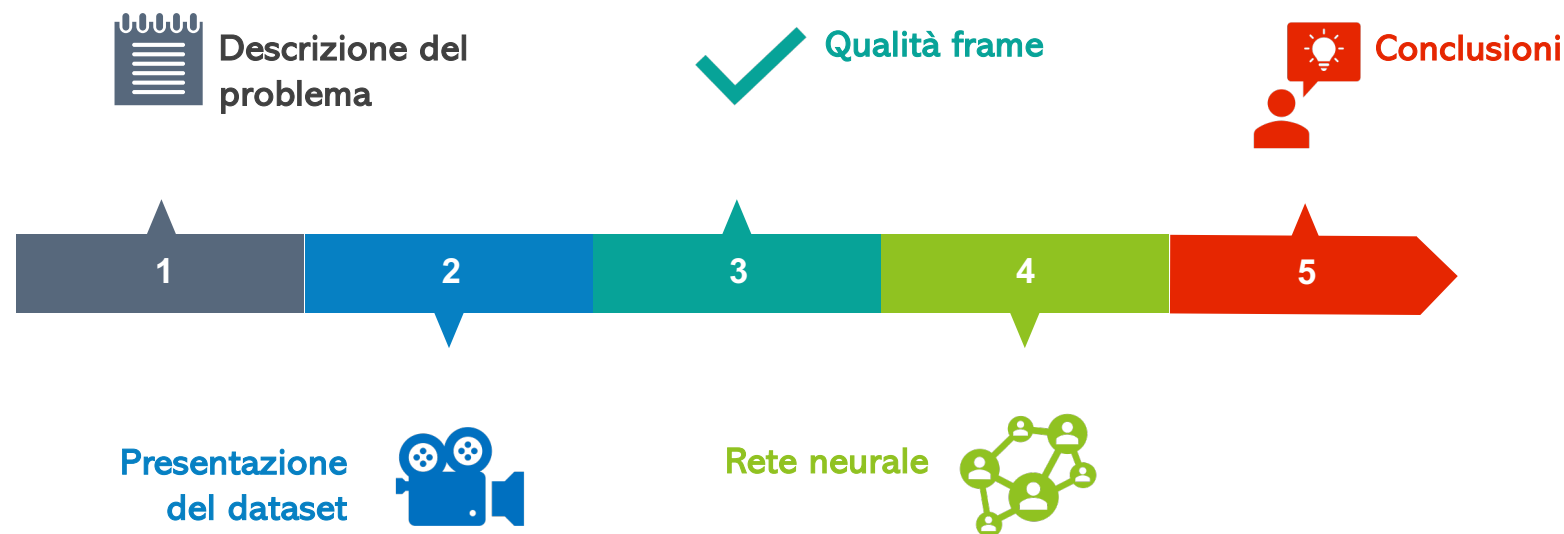
Prof. Francesco Sergio Pisani

Studenti

Claudia Grazioso 214510

Tommaso Ruga 214778

Sommario



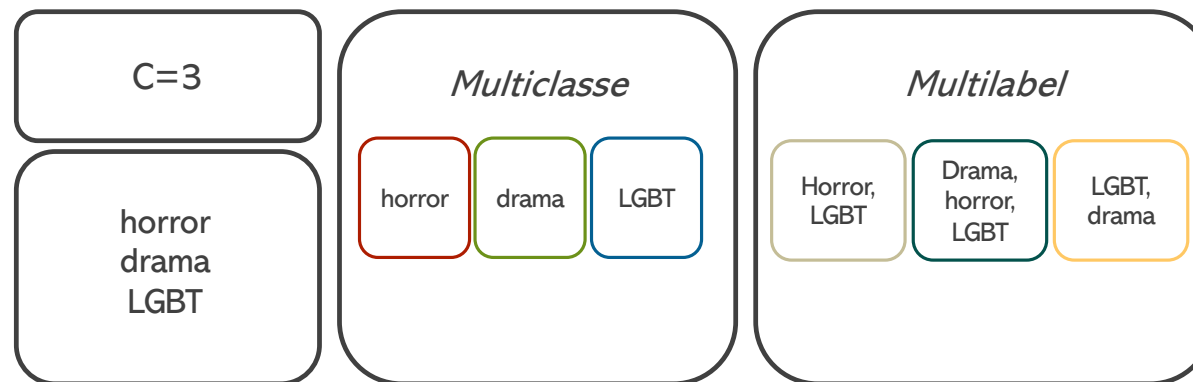


Descrizione
del problema

1

Obiettivo:

- Costruire un modello di classificazione
- Predire le classi associate ad ogni film

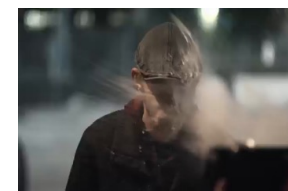
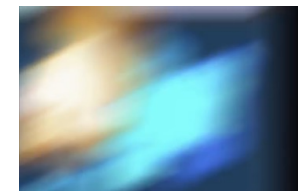


- Multiclasce: problema di classificazione in cui l'obiettivo è assegnare una classe agli oggetti scelta tra più di due classi
- Multilabel: problema di classificazione in cui ad ogni oggetto la classificazione assegna più di una classe

- **progetto_2021_dataset_labeled.tgz –TRAIN+TEST**
- progetto_2021_dataset_UNlabeled1.tgz
- progetto_2021_dataset_UNlabeled2.tgz
- progetto_2021_dataset_UNlabeled3.tgz

progetto_2021_dataset_labeled.tgz –TRAIN+TEST

- 5404 cartelle contenenti frame (360x240) in numero variabile estratti da trailer di film
 - 4292 di train
 - 1112 di test
- train_test_split_dict.json
 - *Esempio: tt4060866': ['crime', 'drama', 'action']*
- 85 label distribuite in modo sbilanciato





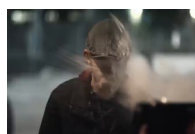
3

Valutare la qualità dell'immagine rappresentando *quantitativamente* la percezione umana della qualità

- Modello **BRISQUE**: utilizza solo i pixel dell'immagine per calcolarne le feature
- Calcola la luminescenza attraverso il coefficiente *mean subtracted contrast normalized (MSCN)*
 - *Le immagini naturali possiedono una certa regolarità in termini di distribuzione statistica della loro intensità. Questa regolarità viene meno in presenza di degradazioni dell'immagine. Misurando tali degradazioni si è in grado di valutare la qualità dell'immagine*
- Restituisce uno score compreso tra 0 e 100
 - 100 -> qualità pessima



37.39



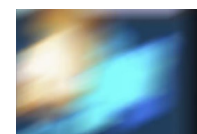
51.33



29.04



67.60



88.29



91.13

Soglia introdotta: immagini con score <60.0



Qualità
frame

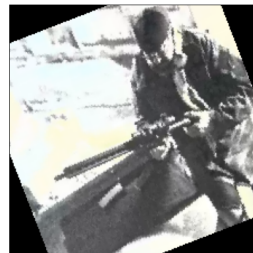
3

- Numero frame per film: 10
- Batch size: 32
- Data augmentation

```
transform = transforms.Compose([transforms.CenterCrop(224),  
                                transforms.RandomHorizontalFlip(p=0.5),  
                                transforms.RandomRotation(degrees=45),  
                                transforms.ToTensor()])
```



Frame originali

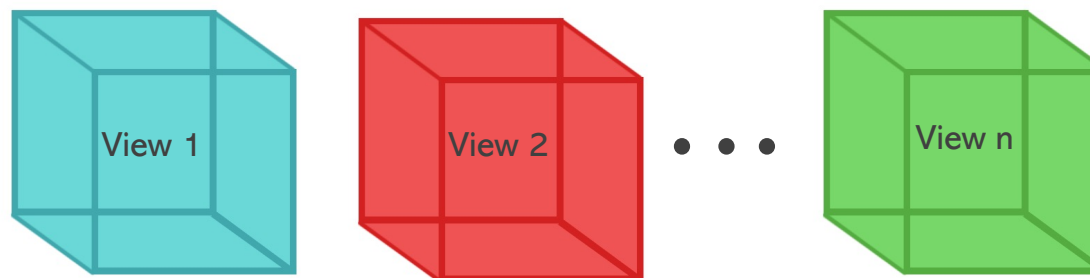


Frame generati



4

- Architettura di **rete neurale personalizzata** per gestire il problema dei **frame multipli**
- Utilizzo di un'architettura CNN pre-addestrata (*Resnet50*) come estrattore di feature
- Aggiunta di un blocco di classificazione costituito da livelli densi



Concatenazione di feature map lungo una dimensione



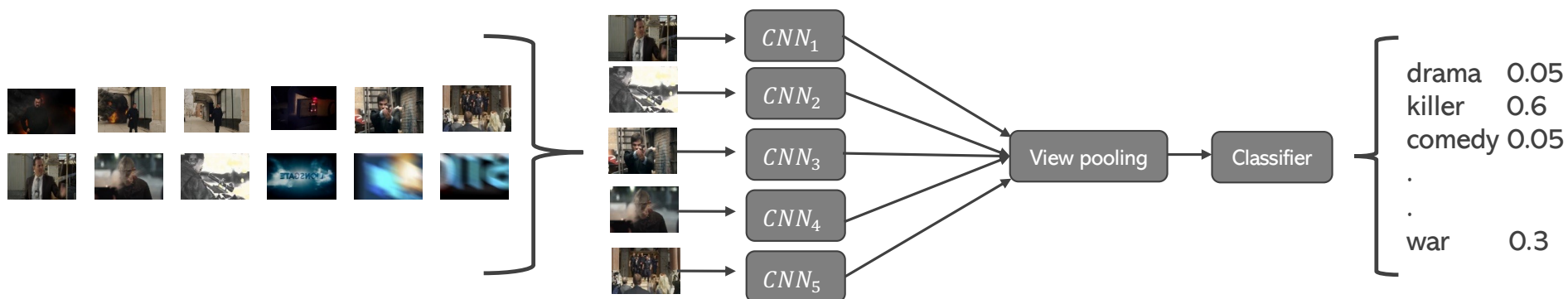
Rete neurale

4

```
class MVCNN(nn.Module):
    def __init__(self, num_classes=85, pretrained=True):
        super(MVCNN, self).__init__()
        resnet = models.resnet50(pretrained = pretrained)
        fc_in_features = resnet.fc.in_features
        self.features = nn.Sequential(*list(resnet.children())[:-1])
        self.classifier = nn.Sequential(
            nn.Dropout(),
            nn.Linear(fc_in_features, 2048),
            nn.ReLU(inplace=True),
            nn.Dropout(),
            nn.Linear(2048, 2048),
            nn.ReLU(inplace=True),
            nn.Linear(2048, num_classes)
        )

    def forward(self, inputs):
        inputs = inputs.transpose(0, 1)
        view_features = []
        for view_batch in inputs:
            view_batch = self.features(view_batch)
            view_batch = view_batch.view(view_batch.shape[0], view_batch.shape[1:].numel())
            view_features.append(view_batch)

        pooled_views= torch.mean(torch.stack(view_features),0)
        outputs = self.classifier(pooled_views)
        return outputs
```





Conclusioni



5

	precision	recall	f1-score	support
LGBT	0.11	0.54	0.18	61
action	0.18	0.98	0.31	191
action_comedies	0.00	0.00	0.00	6
adventure	0.12	0.79	0.20	75
alcohol_addiction	0.00	0.00	0.00	6
alien	0.00	0.00	0.00	12
animation	0.28	0.76	0.41	21
aviation	0.00	0.00	0.00	20
bank_robbery	0.00	0.00	0.00	19
...				
...				
thriller	0.14	0.98	0.25	139
time_travel	0.00	0.00	0.00	12
true_story	0.00	0.00	0.00	10
vampires	0.00	0.00	0.00	12
violence	0.00	0.00	0.00	20
war	0.22	0.33	0.27	58
wartime	0.13	0.65	0.21	49
western	0.15	0.24	0.18	33
zombies	0.00	0.00	0.00	15
micro avg	0.17	0.61	0.27	3167
macro avg	0.05	0.18	0.07	3167
weighted avg	0.17	0.61	0.25	3167
samples avg	0.17	0.65	0.26	3167



Grazie per l'attenzione