

Određivanje žanra filma na osnovu postera

Vladimir Jovin SW-30/2018

Fakultet tehničkih nauka u Novom Sadu



Problem

Poster su jedan od ključnih delova filmske industrije. Njihova osnovna funkcija je da zainteresuju i privuku pažnju potencijalnih gledalaca. S obzirom na način na koji ljudi interpretiraju različite elemente poput boja, pejzaža, facijalnih ekspresija, i slično, mogu da naprave pretpostavke o tematici filma. Ovo dovodi do pretpotavke da je moguće da boje, tekstura, kao i strukturalni elementi slike, nose neke inherente relacije koje bi mogle da se koriste u svrhe algoritama mašinskog učenja

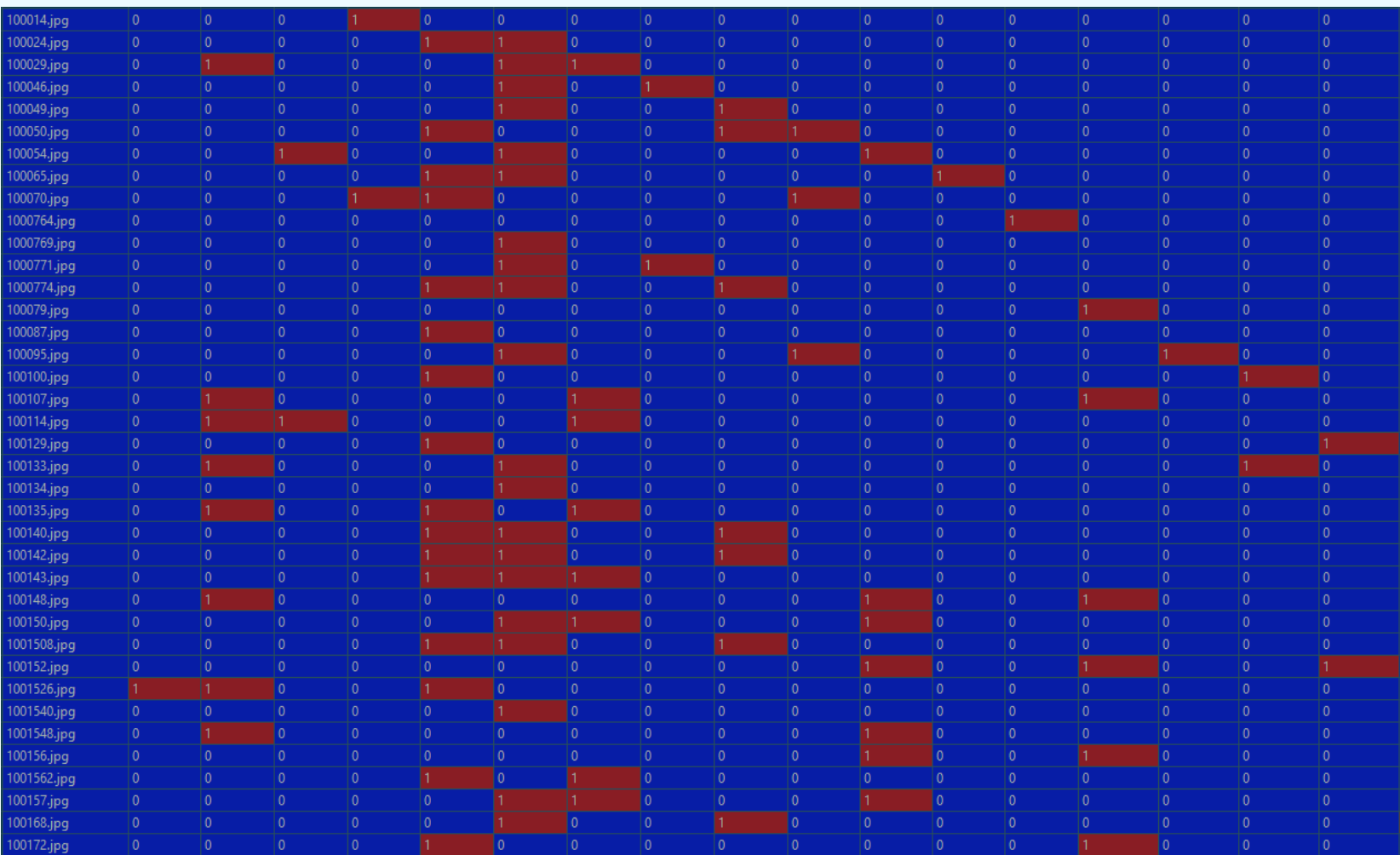
Skup podataka

Korišćen je javno dostupni skup podataka sa Kaggle sajta* koji sadrži 39470 jedinstvenih naslova filmova. Broj mogućih kategorija koje filmovi mogu imati je 28. Svaki naslov odgovara minimalno jednom žanru, odnosno maksimalno trima različitim žanrovima. Prvi korak u izradi projekta jeste preuzimanje i skladištenje ovih slika. Za svaku sliku je provereno da li je link do slike validan i nakon preuzimanja slike proverava se da li je slika validna. Ukoliko nije izbacuje se.

* <https://www.kaggle.com/neha1703/movie-genre-from-its-poster>

Multi-hot encoding

Žanrovi, s obzirom na činjenicu da su kategorični, moraju da se konvertuju u neku numeričku formu da bi se reprezentovali. Primenjen je princip **one-hot encodinga**, koji konvertuje kategoričke labele u vektore binarnih vrednosti. U slučaju da film jeste nekog žanra vrednost će biti 1, u slučaju da nije biće 0.



Arhitekture modela

Arhitektura mreže prikazana je ispod. Slike se učitavaju u dimenzijama (100, 75, 3(RGB)). Zbog limitacija u resursima, prevashodno ramu, nasledjen je keras.utils.Sequence i implementiran namenski generator za učitavanje slika. Poslednji Dense sloj ima **28 jedinica**, što odgovara 28 mogućih klasa, odnosno žanrova. Ovo je multi-label kasifikacija, što znači da je potrebno izračunati nezavisne verovatnoće za svaku od 28 mogućih žanrova. Ovo se obezbeđuje korišćenjem **sigmoidne** funkcije kao aktivacioni funkcije u poslednjem sloju. S obzirom da klase nisu međusobno isključive, ne koristi se categorical cross-entropy kao loss, već se koristi **binary cross-entropy**.

Train
25451

Validation
7272

Test
3636

Arhitektura prvog modela

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 96, 71, 32)	2432
max_pooling2d (MaxPooling2D)	(None, 48, 35, 32)	0
dropout (Dropout)	(None, 48, 35, 32)	0
conv2d_1 (Conv2D)	(None, 46, 33, 32)	9248
max_pooling2d_1 (MaxPooling2	(None, 23, 16, 32)	0
dropout_1 (Dropout)	(None, 23, 16, 32)	0
conv2d_2 (Conv2D)	(None, 21, 14, 64)	18496
max_pooling2d_2 (MaxPooling2	(None, 10, 7, 64)	0
dropout_2 (Dropout)	(None, 10, 7, 64)	0
conv2d_3 (Conv2D)	(None, 8, 5, 64)	36928
max_pooling2d_3 (MaxPooling2	(None, 4, 2, 64)	0
dropout_3 (Dropout)	(None, 4, 2, 64)	0
flatten (Flatten)	(None, 512)	0
dense (Dense)	(None, 128)	65664
dropout_4 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 64)	8256
dropout_5 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 28)	1820

Arhitektura drugog modela

VGG16 model ima pre-trenirane težine (Imagenet) bez klasifikacionog sloja. Svi slojevi, osim poslednja 4, su zamrznuti. Potpuno povezani klasifikacioni sloj je dodat, takođe i sigmoid, iz ranije pomenutih razloga, sa 28 jedinica.

Layer (type)	Output Shape	Param #
vgg16 (Functional)	(None, 3, 2, 512)	14714688
flatten_2 (Flatten)	(None, 3072)	0
dense_5 (Dense)	(None, 1024)	3146752
dropout_7 (Dropout)	(None, 1024)	0
dense_6 (Dense)	(None, 28)	28700

Problemi i rezultati

Primećena je **nebalansiranost** skupa podataka. Naime 48% naslova su žanra "Drama", dok je naredni žanr po stepenu reprentovanosti komedija sa 30%, a pritam ima dosta žanrova koji imaju jako malu količinu pojava, poput Reality-TV sa 2 pojave, Talk-Show sa 6 pojava, News sa 78... Ovo bi se verovatno moglo popraviti korišćenjem resampling tehnika u nekom budućem radu.

Korektna identifikacija podzumeva da je bar jedan žanr, od 3 sa najvećom verovatnoćom uspešno identifikovan.

Prvi model je postigao **73.89%** tačno identifikovanih postera

Drugi model je postigao **80.79%** tačno identifikovanih postera



Action
0.51

Fantasy
0.32

Horror
0.15

Action
0.43

Thriller
0.17

Crime
0.16

