

Klasifikacija umetničkih slika po stilovima korišćenjem CNN i multimodalnog pristupa

1. Definicija problema

Cilj projekta je razvoj modela koji može da klasifikuje umetničke slike prema umetničkom stilu (npr. realizam, barok, impresionizam...). Osnovni ulaz u model su slike umetničkih dela, dok se dodatni metapodaci kao što su godina nastanka i autor koriste da se poboljša tačnost klasifikacije. Problem je višeklasna klasifikacija sa 137 klasa, gde svaka instanca slike pripada tačno jednom stilu, što čini problem izazovnim zbog velikog broja klasa i potencijalne neravnoteže u raspodeli podataka.

2. Motivacija

Razlikovanje umetničkih stilova je zadatak od značaja za digitalne muzeje, obrazovne platforme i sisteme za pretragu kulturnog sadržaja. Automatska klasifikacija umetničkih dela može pomoći u:

- digitalizaciji i boljoj organizaciji muzejskih kolekcija,
- razvoju obrazovnih alata koji studentima pomažu da uče o umetnosti,
- unapređenju sistema preporuka (npr. preporuka sličnih dela)

3. Skup podataka

Koristiću WikiArt dataset: <https://huggingface.co/datasets/Artificio/WikiArt>

Ovaj dataset sadrži oko 103.000 umetničkih dela, kolona koju je cilj predvideti je *style*. Najvažniji atributi:

- **image** – vizuelni prikaz slike (ulaz u CNN),
- **style** – umetnički stil (ciljno obeležje, 137 klasa),
- **artist** – ime umetnika (opciono kao metapodatak),
- **date** – godina nastanka dela (opciono kao metapodatak),
- **genre** – žanr slike (dodatni kontekst, npr. portret, pejzaž).

4. Način pretprocesiranja podataka

Sve slike će biti dovedene na istu rezoluciju od 224×224 piksela i normalizovane tako da vrednosti piksela budu u standardizovanom opsegu pogodnom za treniranje neuronskih mreža. Podaci će biti podeljeni u trening, validacioni i test skup (70/15/15). Radi poboljšanja generalizacije korišće se metode augmentacije kao što su rotacija, horizontalno preslikavanje i promene osvetljenja.

Metapodaci će takođe biti pretprocesirani: godina će biti numerički normalizovana i umetnik i žanr će biti predstavljeni korišćenjem one-hot encoding-a ili label encoding-a.

5. Metodologija

Radiću projekat iz dve faze. U prvoj fazi korišću CNN za klasifikaciju umetničkog stila isključivo na osnovu vizuelnih karakteristika slike. U drugoj fazi, sistem će biti proširen multimodalnim pristupom, gde se pored vizuelnih odlika iz CNN-a uključuju i metapodaci (godina, umetnik i žanr). Dobijeni skup osobina se spaja i prosleđuje završnom klasifikatoru koji određuje umetnički stil.

Projekt će se testirati u dve verzije: samo CNN i CNN + metapodaci, da bi se ocenilo koliko metapodaci doprinose tačnosti.

6. Način evaluacije

Metrike:

- accuracy
- F1-score
- Confusion matrix

Poređenje performansi:

- samo CNN model,
- CNN + metapodaci model.

7. Tehnologije

Programski jezik: Python.

Biblioteke:

- **PyTorch** ili TensorFlow/Keras za implementaciju neuronskih mreža,
- **scikit-learn** za pretprocesiranje i evaluaciju,
- **pandas/numpy** za obradu metapodataka.

8. Relevantna literatura

- Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). *Image style transfer using convolutional neural networks*. Link: <https://ieeexplore.ieee.org/document/7780634>
- Tan et al. (2016): "Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification". Link: [https://www.researchgate.net/publication/305402238 Ceci n'est pas une pipe A Deep Convolutional Network for Fine-art Paintings Classification](https://www.researchgate.net/publication/305402238_Ceci_n'est_pas_une_pipe_A_Deep_Convolutional_Network_for_Fine-art_Paintings_Classification)
- <https://www.kaggle.com/datasets/steubk/wikiart/data>