

PRIMJENA DUBOKOG UČENJA ZA DETEKCIJU STUPNJA SLIČNOSTI INDIVIDUALNOG UMJETNIČKOG STILA

Manhart, Katarina

Master's thesis / Specijalistički diplomski stručni

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Algebra University College / Visoko učilište Algebra**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:225:152296>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-11-05**



Repository / Repozitorij:

[Algebra University - Repository of Algebra University](#)



VISOKO UČILIŠTE ALGEBRA

DIPLOMSKI RAD

**PRIMJENA DUBOKOG UČENJA ZA
DETEKCIJU STUPNJA SLIČNOSTI
INDIVIDUALNOG UMJETNIČKOG STILA**

Katarina Manhart

Zagreb, veljača 2023.

„Pod punom odgovornošću pismeno potvrđujem da je ovo moj autorski rad čiji niti jedan dio nije nastao kopiranjem ili plagiranjem tuđeg sadržaja. Prilikom izrade rada koristila sam tuđe materijale navedene u popisu literature, ali nisam kopirala niti jedan njihov dio, osim citata za koje sam navela autora i izvor, te ih jasno označila znakovima navodnika. U slučaju da se u bilo kojem trenutku dokaže suprotno, spremna sam snositi sve posljedice uključivo i poništenje javne isprave stečene dijelom i na temelju ovoga rada“.

U Zagrebu, 1.2.2023.

Katarina Manhart

Predgovor

Veliko hvala profesoru i mentoru dr.sc. Leu Mršiću na vodstvu i pomoći pri završetku pisanja ovog diplomskog rada. #FTW

Zahvaljujem se i dr.sc. Hrvoju Jerkoviću, na podršci i razumijevanju tijekom studija i prilikom pisanja ovog rada, posebno jer me je upravo on prvi upoznao s pojmovima kao što su umjetna inteligencija i strojno učenje. Hvala i ostalim profesorima i asistentima na podatkovnom smjeru, na uloženom trudu i nesebičnom prijenosu znanja.

Želim se zahvaliti Algebri, a posebno Milanu i kolegama iz prijavnog ureda, na velikoj dozi motivacije i prilici da krenem potpuno novim karijernim putem.

Hvala kolegama iz *BASF SE*, naročito Dennisu i Ines, koji brinu da moje stručno znanje raste svakog dana.

Hvala svim članovima moje obitelji i najbližim prijateljima – najprije na strpljenju, a potom i zato što su mi bili velika podrška u procesu školovanja. Ovaj trenutak ne bi bio moguć bez vas.

Na kraju, od srca HVALA mom zaručniku Mateju. Znaš da si ti moj najveći oslonac.

Temeljem članka 8. Pravilnika o diplomskom radu i diplomskom ispitu na diplomskom studiju Visokog učilišta Algebra sačinjena je ova

Potvrda o dodjeli diplomskog rada

kojom se potvrđuje da studentica Katarina Manhart, JMBAG 0130282185, OIB 43562413388 u šk. godini 2022./2023., studij: Diplomski studij - Primijenjeno računarstvo, od strane povjerenstva za provedbu diplomskog ispita, dana 13.01.2023. godine, ima odobrenu izradu diplomskog rada

s temom: **Primjena dubokog učenja za detekciju stupnja sličnosti individualnog umjetničkog stila**

i sažetkom rada: Ovaj rad istražuje u kojoj mjeri algoritam za duboko učenje može prepoznati karakteristike i stupanj sličnosti individualnog umjetničkog stila. Teorijski dio rada opisuje najvažnije istraživačke smjerove u području računalnih znanosti koji mogu doprinijeti umjetničkoj zajednici. Objašnjava se koncept konvolucijskih neuronskih mreža koje su posebno prilagođene za procesiranje slika. Zatim su izdvojene one značajke mreža koje na temelju prethodnih istraživanja pokazuju najveći potencijal za povećanje točnosti prilikom kreiranja modela za prepoznavanje stupnja sličnosti individualnog umjetničkog stila. Praktični dio rada prikazuje korištenje transfernog učenja, pri čemu se prilagođavaju tri moderne arhitekture za duboko učenje: ResNet50, ResNeXt50_32x4d i densenet121. Iako analiza ne pokazuje velike razlike u primjeni modela, prema metrici točnosti najbolje rezultate pokazuje ResNet50.

Mentor je: Leo Mršić.

Odobrenjem diplomskog rada studentici je omogućen upis kolegija "Diplomski rad" te je sukladno članku 8. Pravilnika o diplomskom radu i diplomskom ispitu dužan najkasnije do početka nastave zimskog semestra u sljedećoj školskoj godini, uspješno obraniti diplomski rad uspješnim polaganjem diplomskog ispita.

U protivnom studentica može zatražiti novog mentora/icu i temu te ponovo upisati kolegij "Diplomski rad" budući da rad koji nije predan i obranjen na diplomskom ispitu u roku određenom Pravilnikom diplomskom radu i diplomskom ispitu prestaje vrijediti. Izrada novog diplomskog rada se izvodi sukladno rokovima određenima za školsku godinu u kojoj je studentici određen novi mentor/ica i dodijeljen novi diplomski rad.

Potpis studentice:

Potpis mentora:

Potpis predsjednika
povjerenstva:

Ova potvrda izdaje se u 4 (četiri) primjerka od kojih 3 (tri) idu kao prilog diplomskom radu.

Sažetak

Ovaj rad istražuje u kojoj mjeri algoritam za duboko učenje može prepoznati karakteristike i stupanj sličnosti individualnog umjetničkog stila. Teorijski dio rada opisuje najvažnije istraživačke smjerove u području računalnih znanosti koji mogu doprinijeti umjetničkoj zajednici. Objasňuje se koncept konvolucijskih neuronskih mreža koje su posebno prilagođene za procesiranje slika. Zatim su izdvojene one značajke mreža koje na temelju prethodnih istraživanja pokazuju najveći potencijal za povećanje točnosti prilikom kreiranja modela za prepoznavanje stupnja sličnosti individualnog umjetničkog stila. Praktični dio rada prikazuje korištenje transfernog učenja, pri čemu se prilagođavaju tri moderne arhitekture za duboko učenje: *ResNet50*, *ResNeXt50_32x4d* i *densenet121*. Iako analiza ne pokazuje velike razlike u primjeni modela, prema metrici točnosti najbolje rezultate pokazuje *ResNet50*. Iz tog je razloga uvršten u prototip web rješenja za procjenu stupnja sličnosti individualnog umjetničkog stila. Prototip donosi ideju za novi alat koji mjeri stupanj sličnosti ili utjecaj autora u području umjetnosti.

Ključne riječi: vizualna umjetnost, klasifikacija slika, duboko učenje, konvolucijske neuronske mreže, transferno učenje

Abstract

This paper investigates the extent to which a deep learning algorithm can recognize features and compare the degree of similarity of individual artistic style. The theoretical part of the paper describes new trends in the field of computer science that can contribute to the artistic community. The paper explains concept of convolutional neural networks, which are specially adapted for image processing. It lists features that had the greatest impact on the results of research in domains of author classification and recognition of individual artistic style. Practical part of the paper shows the usage of transfer learning, while adapting three modern architectures for deep learning: *ResNet50*, *ResNeXt50_32x4d* i *densenet121*. Although the comparison does not show major differences in the models, *ResNet50* shows the best results, at least according to the accuracy metric. For this reason, it was included in the web application for the assessment of the individual artistic style. This prototype can inspire idea for a new tool for assessing the influence of individual artistic style and authors in the field of art.

Keywords: visual arts, painting classification, deep learning, convolutional neural networks, transfer learning

Sadržaj

	Uvod	1
	Prethodna istraživanja u području računalnih znanosti na temu vizualnih umjetnosti..	3
1.	1.1. Automatska klasifikacija slika.....	2
	1.2. Automatsko prepoznavanje i detekcija objekata na slikama.....	4
	1.3. Prepoznavanje emocija	5
	1.4. Prijenos stila i generiranje nove umjetnosti.....	5
2.	Primjena konvolucijskih neuronskih mreža za prepoznavanje stupnja sličnosti individualnog umjetničkog stila	7
	2.1. Konvolucijske neuronske mreže.....	7
	2.2. Važnost plitkih slojeva i semantičkog sadržaja za detekciju stupnja sličnosti individualnog umjetničkog stila	9
	2.3. Kontrola prekomjernog prilagođavanja na bazi umjetničkih djela	9
3.	2.3.1. Upravljanje veličinom baze podataka.....	10
	2.3.2. Dropout sloj	11
	2.3.3. L1 i L2 regularizacija	12
	Izrada rješenja za prepoznavanje stupnja sličnosti individualnog umjetničkog stila ..	13
	3.1. Prikupljanje podataka i definiranje uzorka.....	14
	3.2. Istraživačka analiza i priprema podataka.....	16
4.	3.3. Osnovna ResNet arhitektura i njezina obilježja	19
	3.3.1. Treniranje i modifikacija ResNet modela za prepoznavanje autora.....	21
	3.3.2. Evaluacija modela.....	22
	Unaprjeđenje modela za prepoznavanje autora umjetničkog djela	23
4.1.	Korištenje ResNeXt arhitekture.....	23

4.1.1.	Treniranje i modifikacija ResNeXt modela.....	24
4.1.2.	Evaluacija modela.....	24
4.2.	Korištenje DenseNet arhitekture	25
4.2.1.	Treniranje i modifikacija DenseNet modela.....	26
4.2.2.	Evaluacija modela.....	26
	Prototip web rješenja za detekciju stupnja sličnosti individualnog umjetničkog stila	28
5.1.	Izrada prototipa koristeći Streamlit	28
5.2.	Upogonjavanje modela koristeći Google Colab.....	30
5.3.	Primjena modela u stvarnim situacijama – <i>crash test</i>	33
	Smjernice za buduća istraživanja	43
6.	Zaključak	44
	Popis kratica	47
	Popis slika.....	48
	Popis tablica.....	49
	Popis kodova	50
	Literatura	51

Uvod

Umjetnost prati čovjeka od njegovih najranijih početaka te danas predstavlja izrazito važan dio kulture i društva. Umjetnička djela predstavljaju poseban način dokumentiranja i očuvanja povijesti. Mogu biti napravljena od različitih materijala, medija i tehnika te predstavljaju univerzalni način izražavanja kreativnosti i emocija. Klasifikacija umjetničkih djela u pravilu se prepušta stručnjacima, međutim, neki od najutjecajnijih slikara imaju individualni umjetnički stil kojeg lako prepoznaju i potpuni laici. Takvu vještinu prepoznavanja ljudi obično smatraju jako prirodnom te ju teško opisuju riječima. U literaturi zapravo ne postoji jedinstvena definicija individualnog umjetničkog stila, ali gotovo uvijek postoje preklapanja u najvažnijim karakteristikama koje ga određuju; kao što su način na koji umjetnik koristi formu, boje i kompoziciju; sklonost prema korištenju određenih metoda i tehnika te odabir motiva koji se nalazi u pozadini samog umjetničkog djela (*Art Styles Explained - Explore Popular Styles of Art! — Art Is Fun, 2023*). Ovaj rad istražuje je li koncept individualnog umjetničkog stila moguće promatrati matematički te u kojoj mjeri umjetna inteligencija može prepoznati nešto što je toliko povezano s ljudskom intuicijom i osjećajima. Sadržaj je podijeljen u dva dijela.

Prvi, teorijski dio rada, upoznaje čitatelja s pregledom modernih istraživanja koja nastoje povezati računalne znanosti i vizualnu umjetnost. Istraživanja su grupirana u četiri istraživačka smjera, a zajedničko im je da sva koriste duboko učenje - oblik strojnog učenja i umjetne inteligencije koji najviše podsjeća na način na koji ljudski mozak stječe nova znanja. Istaknuto je područje automatske klasifikacije slika kojem pripada ovaj rad. U drugom poglavlju detaljno su predstavljeni fundamentalni koncepti konvolucijske neuronske mreže. Radi se o arhitekturi dubokog učenja koja redovito daje najbolje rezultate prilikom analize i obrade slika. Opisane su pojedinačne karakteristike konvolucijske neuronske mreže na koje je potrebno pripaziti prilikom razvoja modela ovog tipa.

Drugi, praktični dio rada, bavi se izradom programskog rješenja za detekciju stupnja sličnosti individualnog umjetničkog stila. Prilikom izrade aplikacije postavljena su tri istraživačka cilja. Prvi cilj je prilagoditi algoritam za duboko učenje da što jasnije prepozna jedinstveni, individualni umjetnički stil deset odabranih autora.

Za potrebe istraživanja prikupljene su fotografije umjetničkih djela korištenjem aplikacijskog programskog sučelja internetske stranice *WikiArt*. Kako bi rad imao što veću vrijednost i manju pristranost, u uzorak su uključena djela deset najutjecajnijih slikara prema istraživanju „*Interakcije i utjecaj svjetskih slikara iz reducirane Google matrice Wikipedija mreža*“ (El Zant et al., 2018). Individualni umjetnički stil promatran je kao set vizualnih karakteristika koje povezuju različita umjetnička djela jednog autora. Upravo taj skup karakteristika čini svakog autora jedinstvenim i prepoznatljivim.

Pretpostavka je da će konvolucijska neuronska mreža biti u stanju prepoznati individualni umjetnički stil, ukoliko je njegove najvažnije karakteristike moguće promatrati matematički. Na primjer, potezi kistom ili olovkom mogu sadržavati geometrijske oblike i udaljenosti, dok je boje moguće klasificirati kroz njihove RGB kodove. Prethodna istraživanja pokazuju da je ovakvo prepoznavanje moguće do neke mjere, ali broj istraživanja koje se bave ovom temom još uvijek je izrazito ograničen. Na temelju pregleda postojećih istraživanja, za ovu svrhu odabrana je prilagođena rezidualna neuronska mreža (u daljnjem tekstu *ResNet*). Posebnost ove mreže je da po svojoj strukturi i funkciji oponaša piramidalne stanice koje postoje u kori ljudskog mozga. Završni rezultat algoritma za duboko učenje je model za prepoznavanje individualnog umjetničkog stila, upravo ono što je algoritam naučio.

Ukoliko kreirani model može prepoznati umjetnički stil, prema drugom cilju rada nastoji se povećati uspješnost modela korištenjem i optimizacijom dvije novije duboke arhitekture: *ResNeXt* i *DenseNet*. Ideja je usporediti više modela dubokog učenja te odabrati najoptimalniji za prepoznavanje individualnog umjetničkog stila. Rezultati komparativne analize modela prikazani su u sklopu evaluacije. Do sada u sličnim istraživanjima nisu obuhvaćene novije arhitekture dubokog učenja.

Treći i najvažniji cilj rada je izrada web rješenja koje će svakom korisniku omogućiti postavljanje vlastite digitalne fotografije te usporedbu stupnja sličnosti umjetničkog stila različitih autora. Rezultat je postotak ili procjena, u kojoj mjeri prikazani umjetnički stil nalikuje individualnom umjetničkom stilu nekog od deset najutjecajnijih autora.

Ova ideja inspirirana je činjenicom da autori često žele nalikovati nekom drugom autoru ili školi. Cilj im je naglasiti utjecaj drugih umjetničkih djela, izraziti prestiž koji proizlazi iz nasljeđa njihovog umjetničkog uzora ili preraditi određeni umjetnički motiv u vizualno prepoznatljivom stilu. Razvijeni alat pokazuje matematičku mjeru takvog utjecaja, što ga čini potpuno novim, jedinstvenim i relativno objektivnim alatom za procjenu umjetnosti.

Prethodna istraživanja u području računalnih znanosti na temu vizualnih umjetnosti

1. U ovom poglavlju istaknuta je važnost umjetne inteligencije te su definirani osnovni pojmovi iz područja računalnih znanosti koji su potrebni za razumijevanje ovog rada. S obzirom da je fokus rada prepoznavanje stupnja sličnosti individualnog umjetničkog stila, predstavljena su istraživanja koja spajaju područje računalnih znanosti sa vizualnom umjetnosti. Relevantna istraživanja podijeljena su u četiri istraživačka smjera.

Posljednjih godina, dolazi do velikih otkrića i iznimnog napretka u istraživanjima i primjenama umjetne inteligencije (u stranoj literaturi engl. *Artificial Intelligence* ili *AI*). Iako ne postoji službeno prihvaćena definicija umjetne inteligencije, John McCarthy ju 1956. definira kao znanstvenu disciplinu koja se bavi izgradnjom računala čije se ponašanje može tumačiti kao inteligentno. Ova disciplina danas je prihvaćena kao jedna od glavnih predvodnica četvrte industrijske revolucije. Na primjer, 2021. predviđeno je da će veličina tržišta umjetne inteligencije iznositi više od 93,523 milijarde američkih dolara te da će ta brojka nastaviti značajno rasti (Li et al., 2021).

Najveći rast je vidljiv u području dubokog učenja, grani umjetne inteligencije koja se danas smatra jednom od najistraživanijih znanstvenih tema te pronalazi primjenu u gotovo svim industrijama, uključujući i vizualne umjetnosti (Li et al., 2022). Duboko učenje odnosi se na skup tehnika strojnog učenja koje koriste hijerarhijske strukture za klasificiranje i organiziranje podataka (Castellano & Vessio, 2021). Posebnost dubokog učenja u odnosu na klasično strojno učenje je mogućnost računalnih modela da nauče koje su značajke važne za rješavanje problema direktno iz podataka, na sličan način na koji to čini ljudski mozak. Mozak se sastoji od velikog broja neurona koji rade paralelno da bi obradili podatke iz okoline. Slično, računalni modeli u području dubokog učenja koriste neuronske mreže – složene strukture sastavljene od mnogo slojeva koje procesiraju informacije u svrhu učenja značajki i prepoznavanja uzoraka (Adedokun, 2019). Na temelju rezultata obrade podataka, algoritmi obično mogu donijeti informiranu pretpostavku koja se temelji na najvećoj vjerojatnosti. Pretpostavke se često kontinuirano prilagođavaju u procesu treninga, odnosno, čini se da algoritmi ove vrste „uče iz vlastitih pogrešaka“ (Adedokun, 2019). Umjetnost predstavlja posebni oblik individualnog izražavanja koji se smatra izrazito intuitivnim, te se

je cijela domena vizualnih umjetnosti dugo vremena činila odvojenom i nedostižnom za računala. U posljednjih nekoliko godina, tehnološki napredak u kombinaciji sa smanjenjem troška obrade računalnih podataka dovodi do veće dostupnosti velikih digitalnih kolekcija vizualne umjetnosti i otvara nove prilike za istraživače u području računalnih znanosti koji žele pomoći umjetničkoj zajednici (Castellano & Vessio, 2021).

Posebno je široka primjena metoda dubokog učenja, zbog čega je u ovom poglavlju moguće obuhvatiti tek nekoliko najistaknutijih istraživačkih smjerova. To su:

- Automatska klasifikacija slika
- Automatsko prepoznavanje i detekcija objekata na slikama
- Prepoznavanje emocija na slikama
- Prijenos stila i generiranje nove umjetnosti

Ovaj rad može se svrstati u područje automatske klasifikacije slika. Zanimljivi nalazi iz preostala tri smjera, opisani su kako bi potaknuli generiranje novih ideja i istraživanja koje spajaju ova dva izrazito dinamična područja.

1.1. Automatska klasifikacija slika

Automatska klasifikacija jedan je od najčešćih problema za čije se rješavanje koriste neuronske mreže. Radi se o procesu kategorizacije osoba ili stvari u određene grupe prema nekom obilježju. U području vizualnih umjetnosti računalni modeli mogu predviđati neku karakteristiku kao što je umjetničko razdoblje kojem rad pripada, korištena tehnika ili dominantni motiv. Primjer klasifikacijskog problema je i prepoznavanje individualnog umjetničkog stila ili autora umjetničkog djela, istraživački problem koji je inspirirao ovaj rad.

Prepoznavanje autora umjetničkog djela na temelju individualnog umjetničkog stila čini se kao kompleksan zadatak. To je moguće objasniti čestom pojavom niske inter-varijabilnosti među različitim umjetnicima i visoke intra-varijabilnosti u opusu istog umjetnika (Castellano & Vessio, 2021). Različiti umjetnici koji pripadaju istom vremenskom periodu mogu imati sličan slikarski stil, dok je *Pablo Picasso* najpoznatiji primjer autora koji je tijekom karijere imao više stvaralačkih faza koje se značajno razlikuju po umjetničkom stilu.

U slučaju potrebe za klasifikacijom ove vrste, obično se pouzdamo u povjesničare umjetnosti koji posjeduju potrebnu stručnost. Takvi pojedinci prolaze detaljnu obuku kako bi se upoznali sa različitim umjetnicima i individualnim umjetničkim stilovima. Na tragu takvog pristupa, prva istraživanja u području računalnih znanosti također su se oslanjala na ručno definirane općenite značajke slika kao što su transformacije invarijantnih značajki mjerila (engl. *scale-invariant feature transforms*, skraćeno SIFT) ili histogrami usmjerenih gradijenata (engl. *histograms of oriented gradients* ili skraćeno HOG) (Viswanathan, 2017). Takav pristup je bio izrazito kompliciran i računski težak; proračuni su nerijetko zahtijevali veliku količinu vremena te nisu bili isplativi za računala male snage (Goyal et al., 2021). Uz navedene općenite značajke, prijašnja istraživanja nastojala su konstruirati značajke specifične za područje vizualnih umjetnosti, kao što je prepoznavanje karakterističnih poteza kistom (Viswanathan, 2017). Na primjer, Li et al. u tu svrhu koriste kombinaciju detekcije rubova i segmentacije temeljene na grupiranju (Viswanathan, 2017). U sljedećem valu istraživanja autori su eksperimentirali sa metodama kao što su metoda potpornih vektora (engl. *Support Vector Machines ili SVM*) ili lijeni klasifikatori (engl. *k-Nearest Neighbors ili k-NN*) (Viswanathan, 2017).

Najnovija istraživanja napuštaju ideju prethodno definiranih značajki te problem klasifikacije najčešće rješavaju uz pomoć dubokog učenja. Catellano & Vessio ističu da do najveće revolucije dolazi pojavom konvolucijskih neuronskih mreža koje su posebno prilagođene za procesiranje slika (2021). Takve mreže imaju sposobnost automatskog izdvajanja značajki slika koje u pravilu nadmašuju tradicionalne ručno izrađene značajke stručnjaka (Hong & Kim, 2017). Radi velikog značaja i boljeg razumijevanja, konvolucijske neuronske mreže su izdvojene i detaljno opisane u drugom poglavlju rada.

Modeli dubokog učenja koji se bave automatskom klasifikacijom slika imaju više primjena u području umjetnosti. Mogu pomoći stručnjacima prilikom analize slika te značajno ubrzati proces katalogizacije umjetničkih djela, osobito ukoliko je potrebno brzo stvoriti ili proširiti digitalnu kolekciju. U turističkom sektoru javljaju se aplikacije koje mogu automatski prepoznati karakteristike umjetničkog djela, kako bi na primjer, posjetiteljima muzeja pružile relevantne informacije. Izrazito uspješan model u ovoj domeni potencijalno bi se mogao koristiti za detekciju krivotvorina, poznati problem u području povijesti umjetnosti.

1.2. Automatsko prepoznavanje i detekcija objekata na slikama

Sljedeći istraživački trend grupira radove koji se bave temom automatskog prepoznavanja i detekcije objekata na slikama. Ovaj problem je nešto složeniji od tradicionalne automatske klasifikacije slika budući da modeli dubokog učenja uz prepoznavanje objekta na slici nastoje odrediti i njegovu točnu lokaciju (Castellano & Vessio, 2021). To pokazuju crtanjem odgovarajućeg pravokutnog okvira oko identificiranog objekta (Slika 1.1).



Slika 1.1 Usporedba klasifikacije i detekcije objekta

Istraživanja koja se bave ovom temom mogu se podijeliti u dvije skupine (Castellano & Vessio, 2021). Prva uključuje modele kao što su konvolucijske mreže bazirane na regijama (engl. R-CNN), koji najprije predlažu regiju, a zatim svaki prijedlog klasificiraju u različite kategorije. Druga skupina se odnosi na modele u grupaciji YOLO (engl. *You Only Look Once*) koji određenu sliku pogledaju samo jednom da bi odredili koji se objekti nalaze na njoj i gdje točno. Ove metode detekciju objekta objedinjuju u jednu konvolucijsku mrežu te detekciju svode na klasifikacijski ili regresijski model (Zhao et al., 2019). Najvažniji problem koji je otkriven prilikom razvoja modela za automatsko prepoznavanje i detekciju problema je problem unakrsnog prikaza - potrebno je prepoznati i pronaći predmet neovisno o tome je li fotografiran, naslikan ili nacrtan (Castellano & Vessio, 2021). Većina rješenja pretpostavlja da će korisnik odabrati fotografiju, međutim takva istraživanja često nemaju dobru moć generalizacije.

Ovi algoritmi nerijetko se primjenjuju kada je potrebno pronaći objekt na velikom broju slika ili podržati analizu povjesničara umjetnosti. Na primjer, često pitanje stručnjaka u području

umjetnosti je kada se je određeni objekt prvi puta pojavio na slici ili kako se je prikaz određenog objekta mijenjao kroz vrijeme (Catellano & Vessio, 2021).

1.3. Prepoznavanje emocija

Umjetnost se smatra odrazom unutarnjeg stanja i emocija umjetnika te ju svatko doživljava na svoj način. Istraživanja u području vizualnih umjetnosti koja se bave prepoznavanjem emocija često koriste *WikiArt Emotions*, skup podataka koji uključuje 4105 fotografija umjetničkih djela uz napomene o emocijama koje su izazvane kod promatrača (Mohammad & Kiritchenko, 2018). Ovaj skup podataka sadrži dvadeset kategorija emocija, uključujući i neutralno stanje. Osim emocija, napomene se odnose i na prikaz lica te koliko su se umjetnine svidjele osobi koja ih je promatrala. Na primjer, napomene uz fotografiju Mona Lise navode da se u promatračima često bude osjećaji sreće, ljubavi i povjerenja, dok prosječna ocjena ili stupanj sviđanja iznosi 2,1 u rasponu od -3 do 3) (Mohammad & Kiritchenko, 2018). Čini se da umjetnička djela mogu prenositi određene osjećaje i emocije, štoviše, istraživanje Cetinic et al. pokazuje da fino podešavanje neuronskih mreža koje su prethodno trenirane za prepoznavanje scene i predviđanje osjećaja ima veću uspješnost prilikom klasifikacije stila u odnosu na mreže koje su prethodno trenirane za detekciju objekata (2018). Ovo je jedan od najnovijih trendova istraživanja u području vizualnih umjetnosti, međutim, prikupljena znanja se već koriste za razvoj automatskih sustava koji detektiraju emocije izazvane umjetničkim djelom, ali i za razvoj automatskih sustava koje mogu prenesti stil fotografije u emocionalnom smislu ili stvoriti potpuno novu umjetnost koja u promatraču izaziva željeni odgovor (Mohammad & Kiritchenko, 2018).

1.4. Prijenos stila i generiranje nove umjetnosti

Duboko učenje je uspješno i u prijenosu umjetničkog stila s jedne fotografije na drugu. Cilj istraživanja koja se bave prijenosom stila je stvaranje novih slika koje će oponašati stil postavljenog uzora dok istovremeno zadržavaju semantički sadržaj druge slike (Mikołajczyk & Grochowski, 2018). Odvajanje sadržaja slike od njezinog stila nije jasno definiran problem. Teško je odrediti što čini umjetnički stil slike - jesu li to potezi kistom, kolekcija boja, pojava određenih geometrijskih oblika ili kompozicija scene i odabir motiva umjetničkog djela (Gatys et al., 2016). Na primjer, autori navode da je kod poznatog

umjetničkog djela kao što je „*Zvezdana noć*“ Vincenta van Gogha teško zamisliti prijenos stila koji ne bi uključivao zvijezde u pozadini.

Najnoviji trend u umjetničkoj zajednici su generativni modeli koji koriste umjetne neuronske mreže za stvaranje novog sadržaja. U sklopu ovog trenda, najpoznatija je arhitektura generativne suparničke mreže (GAN), koju prvi predlaže *Ian Goodfellow* 2014. godine. GAN je složena arhitektura sastavljena od dvije mreže, koje se međusobno natječu. Prva mreža, generator, stvara lažne slike. Druga mreža, diskriminator, prima stvarne i lažne slike s namjerom da ih što bolje razlikuje. Obje mreže se istovremeno treniraju: kada generator savršeno pogodi distribuciju stvarnih podataka, diskriminator je maksimalno zavarano te predviđa jednaku vjerojatnost (0.5) za sve ulazne slike (Castellano & Vessio, 2021). Kao rezultat, računalo stvara nova umjetnička djela.

Razvoj generativnih modela omogućuje istraživanja koja ispituju kreativnost. Iako ljudi nisu direktno uključeni u proces razvoja modela, ljudski proizvodi se često koriste u procesu učenja. Na primjer, u radu (Elgammal et. al, 2017) u tu svrhu predlažu se kreativne suparničke mreže (CAN), varijaciju originalne generativne arhitekture koja nastoji stvoriti nova umjetnička djela kroz maksimalno odstupanje od ustaljenih stilova i minimalno odstupanje od distribucije umjetnosti (Castellano & Vessio, 2021). Model nastoji stvoriti umjetnost koja je nova, ali ne „previše nova“, uzimajući u obzir pretpostavku da stručnjaci također koriste vlastita prethodna iskustva i stečena znanja da bi razvili kreativnost (Castellano & Vessio, 2021). Modeli ovog tipa značajno ubrzavaju razvoj u industriji igara, gdje je na temelju nekoliko slika moguće generirati izgled igrice u stilu umjetnika. U ovom slučaju kompletno crtanje ilustratora bilo bi izrazito dugotrajno.

Primjena konvolucijskih neuronskih mreža za prepoznavanje stupnja sličnosti individualnog umjetničkog stila

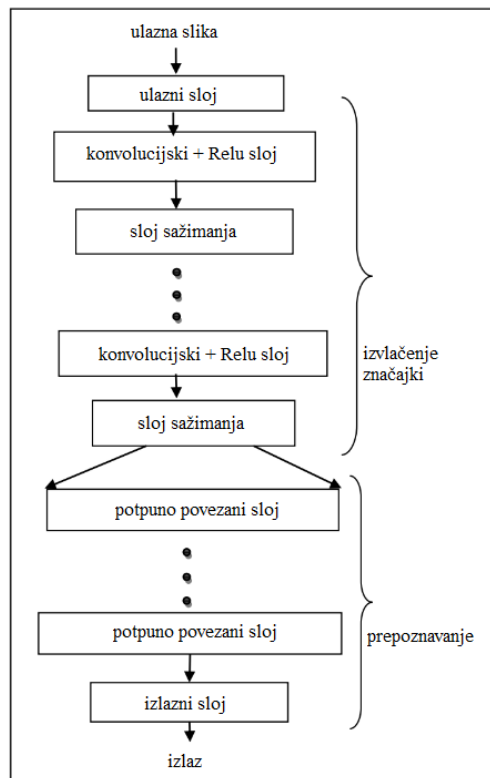
2.

U ovom poglavlju predstavljene su temeljni koncepti konvolucijskih neuronskih mreža, arhitekture za analizu i obradu slika. Opisano je kako izgleda početna arhitektura te je objašnjena uloga različitih gradivnih blokova. Prikazani osnovni pojmovi zatim su smješteni u kontekst rješavanja problema prepoznavanja individualnog umjetničkog stila. Naredna potpoglavlja izdvajaju pojedinačne karakteristike konvolucijske neuronske mreže koje na temelju prethodnih istraživanja mogu utjecati na uspješnost prepoznavanja autora ili individualnog umjetničkog stila. Ukoliko je model u stanju točno prepoznati autora ili individualni umjetnički stil, moguće je posredno mjeriti stupanj sličnosti umjetničkog stila dva različita autora. Stupanj sličnosti predstavlja matematičku vrijednost, odnosno postotak sličnosti između vizualnih karakteristika nove fotografije u odnosu na izračunati skup karakteristika koji je model prepoznao kao individualni umjetnički stil. Dobivenu vrijednost moguće je interpretirati kao mjeru utjecaja jednog autora na umjetničko djelo drugog.

2.1. Konvolucijske neuronske mreže

Konvolucijske neuronske mreže ili CNN posebna su vrsta neuronskih mreža čija se arhitektura često prikazuje kao niz slojeva te je izrazito prilagođena za procesiranje slika (Slika 2.1). Goyal et al. opisuju tipične gradivne blokove konvolucijske neuronske mreže (2021). Na **ulazu** može biti monokromatska slika ili slika u boji. Ukoliko se radi o slici u boji, slika će biti prikazana kao vektor s tri dimenzije - visina, širina i kanal boje (crvena, zelena, plava). U sljedećem koraku **konvolucijski slojevi** obično obrađuju vrlo malene dijelove slike koji su posloženi u male filtere. Svaki filter prikuplja uzorak sa slike i sprema ga u formatu matrice. Proces se ponavlja dok se ne obradi svaki piksel, kao najmanji grafički element slike. Rezultat je jedna velika matrica koja predstavlja različite značajke slike prikupljene u konvolucijskom sloju. Veličina matrice može biti smanjena korištenjem slojeva sažimanja. **Slojevi sažimanja** izvlače najveću vrijednost iz svake podmatrice, tako da se veličina smanjuje bez gubitka važnih značajki. Novi oblik zatim prolazi kroz potpuno

povezani sloj koji je odgovoran za prepoznavanje značajki i obično slikama dodjeljuje klasu ili oznaku. Zatim slika prolazi kroz **izlazni sloj** koji procjenjuje ispravnost klase ili oznake za svaku ulaznu sliku.



Slika 2.1 Tipična arhitektura konvolucijske neuronske mreže (Goyal et al., 2021)

Uz navedene slojeve, u arhitekturi konvolucijskih neuronskih mreža u pravilu se koriste i **funkcije aktivacije**. Primjer je **ReLU** prijenosna funkcija. Koristi se za postavljanje graničnih vrijednosti u raspon od nule do z , pri čemu se sve negativne vrijednosti tretiraju kao nula. Bez sličnih funkcija, ulazni podaci bi prolazili kroz sve slojeve mreže koristeći isključivo linearne funkcije.

U domeni vizualnih umjetnosti, konvolucijska neuronska mreža konstruirana na ovaj način može naučiti prepoznati individualni umjetnički stil autora. Filteri mreže prilagođavaju se karakterističnim obilježjima umjetnika te odgovaraju na prisutnost tih obilježja na ulaznoj slici umjetničkog djela. Baš kao i u navedenom primjeru, arhitektura mreže obično uključuje potpuno povezane slojeve koji su naslagani na vrh konvolucijskih i slojeva sažimanja, dok je finalni odgovor mreže jedinstveni koeficijent pouzdanosti da se radi o određenom autoru

(Castellano & Vessio, 2021). Visoka vrijednost znači da postoji prisutnost snažnog umjetničkog stila, dok niska vrijednost označava da je sličnost slaba ili nepostojeća. Istraživanje Mondal et al. ovaj problem oblikuje u binarni, odgovarajući na pitanje pripada li umjetničko djelo određenom autoru ili mreži ostalih umjetnika te postiže prosječnu točnost od oko 85% (2021). Slični rezultati vidljivi su i prilikom predviđanja radi li se o indijskom ili međunarodnom umjetničkim djelu (Mondal et al., 2021).

2.2. Važnost plitkih slojeva i semantičkog sadržaja za detekciju stupnja sličnosti individualnog umjetničkog stila

Konvolucijski slojevi neuronske mreže izdvajaju različite značajke slike. Filteri niske razine ili plitki slojevi pomažu u prepoznavanju boja i linija, dok značajke visoke razine pronalaze rubove i kutove (Goyal et al., 2021). Umjetnički stil je snažno određen bojama i potezima koji se mogu matematički prikazati kroz korištenje geometrijskih oblika. U skladu s tim, Qu et al. dobivaju značajno bolje rezultate te zaključuju da je treniranje plitkih slojeva modela u pravilu odgovorno za razumijevanje površinskog umjetničkog stila slike, dok su duboki slojevi u pravilu odgovorni za detekciju objekata na slici (2016; Best Artworks of All Time, 2019). S druge strane, istraživanje Karayev et al. pokazuje da neuronske mreže koje su prethodno učile na setu podataka iz domene prepoznavanja i detekcije objekta zadržavaju učinkovitost i prilikom predviđanja umjetničke škole kojoj djelo pripada (2013; Castellano & Vessio, 2021). Autori ovo ponašanje objašnjavaju opažanjem da prepoznavanje predmeta uvelike ovisi i o izgledu predmeta na slici (2013; Castellano & Vessio, 2021). Odnosno, čini se da je umjetnički stil jako ovisan o semantičkom sadržaju umjetničkog djela.

2.3. Kontrola prekomjernog prilagođavanja na bazi umjetničkih djela

Budući da neuronske mreže mogu koristiti milijune ili bilijune parametara za učenje, sklone su problemu prenaučivosti ili prekomjernog prilagođavanja (engl. *overfitting*). Ova statistička pogreška dolazi do izražaja kada matematička funkcija u pozadini modela postane pretjerano usklađena sa skupom točaka koji se koristi za učenje. Model počinje pratiti detalje

i šum koji se nalaze u podacima za učenje umjesto stvarnih odnosa, što negativno utječe na rezultate dobivene na novim podacima koji ne sadrže istu pogrešku. Model koji je uspješan u praksi mora pokazivati dobre rezultate na viđenim primjerima, ali treba i posjedovati moć generalizacije na neviđene primjere. Metode kojima se to postiže u praksi su poznate pod krovim imenom regularizacija (Goodfellow et al., 2016). Trening uspješnog modela ne staje kod pronalaska dovoljno kompleksne neuronske mreže koja ima dovoljno slojeva i parametara, najbolje rezultate obično postižu modeli koji su uz to i pravilno regularizirani (Goodfellow et al., 2016).

Sljedećih nekoliko odlomaka objašnjava najvažnije metode regularizacije koje su u prethodnim istraživanjima imale najveći utjecaj na rezultate prepoznavanja autora ili individualnog umjetničkog stila. To su upravljanje veličinom baze podataka, *dropout* sloj i L1/L2 regularizacija. Svim navedenim tehnikama je zajedničko da smanjuju pogrešku na skupu za testiranje, međutim, mogu dovesti do povećanja pogreške na skupu za učenje (Goodfellow et al., 2016). Povećana pogreška na skupu za učenje može dovesti do suprotnog fenomena poznatog kao podnaučenost ili engl. *underfitting*, situacije gdje je model prejednostavan u odnosu na stvarnu funkciju, što dovodi do loših predikcija i na viđenim i na neviđenim primjerima.

2.3.1. Upravljanje veličinom baze podataka

Najučinkovitiji trik za povećanje sposobnosti generalizacije modela je treniranje na većoj količini podataka. U praksi, broj umjetničkih djela po autoru je ograničen. Broj ulaznih podataka može se dodatno povećati prikupljanjem slika iz više različitih izvora ili kreiranjem novih umjetničkih djela uz pomoć prijenosa stila ili generativnih suparničkih mreža (Mikołajczyk & Grochowski, 2018). Treća opcija je umjetno povećanje količine podataka augmentacijom skupa za treniranje. Metode augmentacije povećavaju varijabilnost u skupu podataka za treniranje, što dodaje šum (Goodfellow et al., 2016). Najpoznatije augmentacije ili transformacije slika uključuju rotacije, promjenu veličine slike, centriranje, normalizaciju te prilagođavanje zasićenja i kontrasta.

Tablica 2.1 prikazuje nekoliko dostupnih primjera iz PyTorch dokumentacije:

Tablica 2.1 Primjer transformacija slika koje su dostupne u PyTorch biblioteki (PyTorch, 2022)

Transformacija	Funkcija
<i>CenterCrop</i> (veličina)	Izrezuje zadanu sliku u sredini.
<i>ColorJitter</i> ([svjetlina, kontrast, ...])	Nasumično mijenja svjetlinu, kontrast, zasićenost ili nijansu slike.
<i>Grayscale</i> ([broj izlaznih kanala])	Pretvora sliku u sive tonove.
<i>Pad</i> ([popunjavanje, ispunjenje, način popunjavanja])	Popunjava sliku sa svih strana zadanom vrijednošću.
<i>RandomHorizontalFlip</i> ([p])	Vodoravno okreće zadanu sliku nasumično sa zadanom vjerojatnošću.
<i>RandomRotation</i> (stupnjevi[, interpolacija, ...])	Rotira sliku po kutu.
<i>RandomVerticalFlip</i> ([p])	Okomito okreće zadanu sliku nasumično sa zadanom vjerojatnošću.
<i>Resize</i> (veličina [, interpolacija, maksimalna veličina, ...])	Promijenite veličinu ulazne slike na zadanu veličinu.

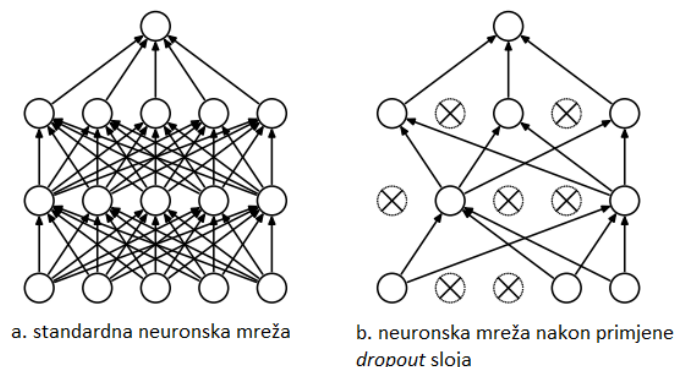
Navedene metode augmentacije rade se samo na skupu podataka za učenje, budući da mogu pomoći da model postigne bolju moć generalizacije i robusnost (Chollet, 2021). Nema ih svrhe koristiti na skupovima za predviđanje i validaciju.

Metode augmentacije pokazuju posebno dobre rezultate u istraživanjima koja se bave detekcijom objekta na slikama (Goodfellow et al., 2016).

2.3.2. Dropout sloj

Uvođenje *dropout* sloja dovodi do velikog iskoraka u metodama regularizacije (Srivastava et al., 2014). *Dropout* sloj prilikom učenja nasumično gasi pojedine neurone i tako dodaje

šum. Model je prisiljen za slične primjere tražiti više različitih putanja do sličnog izlaza, što obično povećava mogućnost generalizacije (Srivastava et al., 2014) (Slika 2.2).



Slika 2.2 Model s primjenom *dropout* sloja (Srivastava et al., 2014)

Chollet (2021) navodi kako se u praksi *dropout* sloj definira kroz postotak težina koje će se ugasiti tijekom učenja te se dodaje neposredno prije potpuno povezanog sloja. Kao i kod metoda augmentacije, niti ovaj oblik regularizacije ne koristi se na skupovima podataka za predviđanje i validaciju.

2.3.3. L1 i L2 regularizacija

L1 regularizacija ili *Lasso* regularizacija je suma apsolutnih vrijednosti svih težina u modelu (Goodfellow et al., 2016). Težine uvijek mogu biti pozitivne ili negativne.

L1 regularizacija dodaje se u funkciju gubitka te se težine kažnjavaju na temelju apsolutne vrijednosti parametara skaliranih za neku vrijednost (Goodfellow et al., 2016).

L2 regularizacija ili *Ridge* regularizacija je najkorištenija metoda regularizacije koja kažnjava velike težine. Potiče se da mreža podjednako koristi sve ulaze te da težine teže prema nuli (Goodfellow et al., 2016). Budući da istraživanja pokazuju da primjena *L1* i *L2* regularizacije u pravilu značajno ne utječe na rezultate istraživanja u domeni prepoznavanja umjetničkog stila, ove tehnike nisu korištene prilikom izrade praktičnog rješenja (Viswanathan, 2017).

Važno je napomenuti da pretjerano korištenje metoda regularizacije također može dovesti do loših predikcija mreže (Goodfellow et al., 2016).

Izrada rješenja za prepoznavanje stupnja sličnosti individualnog umjetničkog stila

3. U ovom poglavlju opisana je izrada rješenja za prepoznavanje stupnja sličnosti individualnog umjetničkog stila. Prilikom izrade, dana je posebna pažnja prikupljanju podataka te definiciji uzorka, budući da metode odabira mogu imati značajan utjecaj na rezultate. Rješenje je napisano u programskom jeziku *Python*, a modeli su vezani uz aplikacijski okvir *PyTorch* (PyTorch, 2022). Budući da je u stvarnosti broj umjetničkih djela po autoru izrazito ograničen, u uzorak je uključen mali dio prikupljenih podataka te je za detekciju stila korišteno transferno učenje. Transferno učenje počinje sa slučajnim inicijalizacijama, ali od mreže koja je trenirala na nekom povezanom zadatku (Stevens, 2020). Ova metoda se često koristi od praksi jer ubrzava proces dolaska do rezultata i traži puno manje ulaznih podataka (Stevens, 2020). U ovom radu prilagođene su tri moderne arhitekture (*ResNet*, *ResNeXt* i *DenseNet*) koje su prethodno trenirane na tisuću *ImageNet* kategorija. Težine dobivene na ImageNet-u su prenesene budući da prethodna istraživanja u domeni vizualnih umjetnosti pokazuju da su motivi ili sadržaj slike važna odrednica umjetničkog stila (Gatys et al., 2016; Karayev et al., 2013; Castellano & Vessio, 2021). Svi eksperimenti su provedeni u *Google Colab-u*, koristeći *2 x vCPU* i *NVIDIA Tesla T4 GPU*, kao i besplatni prostor u oblaku. *Google Colab* odabran je za izradu praktičnog rješenja jer *GPU* prilikom treninga neuronskih mreža može skratiti vrijeme treniranja mreže više od 50 puta u odnosu na implementaciju samo uz *CPU* (Hong & Kim, 2017). Prikazani kôd pokazuje *Python* biblioteke potrebne za reprodukciju modela (Kôd 3.1):

```
[ ] from datetime import datetime
import time
import os
import copy

import torch
import torch.nn as nn
import torch.optim as optim
from torch.optim import lr_scheduler
import torchvision
from torchvision import datasets, models, transforms

import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

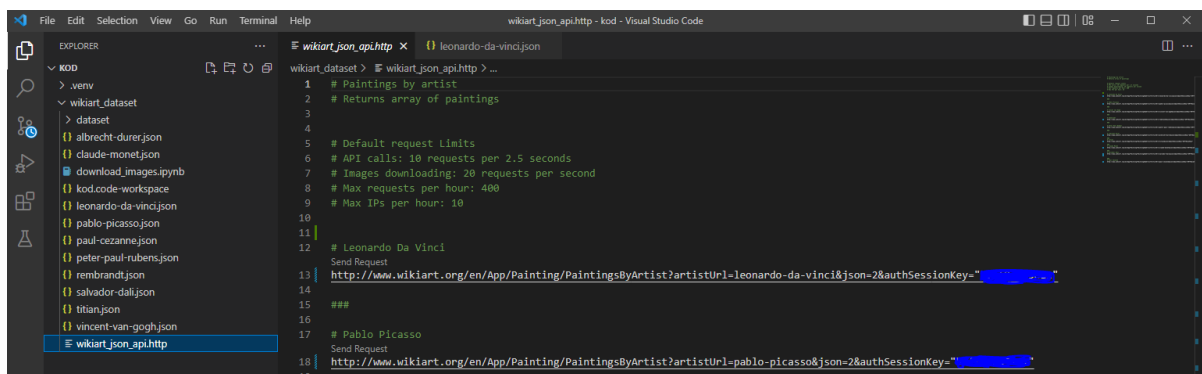
Kôd 3.1 Potrebne *Python* biblioteke za reprodukciju transfernog učenja

3.1. Prikupljanje podataka i definiranje uzorka

Za potrebe ovog rada, prikupljene su fotografije s internetske stranice *WikiArt*, najveće javno dostupne enciklopedije koja prikuplja informacije u području vizualnih umjetnosti. Stranica sadrži kopije originalnih umjetničkih djela u niskoj rezoluciji, ali i bogat set meta podataka kao što su umjetnički stil, period ili kolekcija. Sadržaj stranice neprestano raste pa u trenutku pisanja ovog rada, *WikiArt* sadrži više od 250,000 umjetničkih djela koja su pripisana oko 3000 različitih autora (*WikiArt.org - Visual Art Encyclopedia, n.d.*). Castellano & Vessio (2021) u opsežnom pregledu literature navode kako je upravo ova stranica najčešća početna točka u slučaju potrebe za kreiranjem vlastitog seta podataka u istraživačkim projektima koji su povezani s umjetnosti. S obzirom da je krajnji cilj rada izrada web rješenja za procjenu individualnog umjetničkog stila, u uzorak su uvršteni radovi deset najutjecajnijih slikara prema globalnoj rang listi El Zanta et al., redom to su **Leonardo da Vinci, Pablo Picasso, Vincent Van Gogh, Rembrandt Van Rijn, Peter Paul Rubens, Albrecht Durer, Titian, Claude Monet, Salvador Dali i Paul Cézanne** (2018).

Podaci su prikupljeni pomoću javno dostupnog aplikacijskog programskog sučelja *WikiArt Api – WikiArt.org* (u daljnjem tekstu API) te korištenjem ekstenzije *REST client* unutar aplikacije *Visual studio code*.

Visual studio code je tekst editor razvijen od strane Microsofta koji uključuje podršku za traženje grešaka u kôdu, alate za upravljanje verzijama kôda, označavanje sintakse te pametno dovršavanje i preuređivanje kôda (suradnici Wikipedije, 2022). Program je moguće prilagoditi instalacijom dodatnih ekstenzija ili proširenja. Ekstenzija *REST client* omogućuje jednostavno slanje HTTP zahtjeva (Slika 3.1), kao i izravan pregled odgovora u programu, koji zatim može biti lokalno spremljen na osobno računalo.



```
1 # Paintings by artist
2 # Returns array of paintings
3
4
5 # Default request Limits
6 # API calls: 10 requests per 2.5 seconds
7 # Images downloading: 20 requests per second
8 # Max requests per hour: 400
9 # Max IPs per hour: 10
10
11
12 # Leonardo Da Vinci
13 Send Request
14 http://www.wikiart.org/en/App/Painting/PaintingsByArtist?artistUrl=leonardo-da-vinci&json=2&authSessionKey="
15 ###
16
17 # Pablo Picasso
18 Send Request
19 http://www.wikiart.org/en/App/Painting/PaintingsByArtist?artistUrl=pablo-picasso&json=2&authSessionKey="
20
```

Slika 3.1 Primjer korištenog http zahtjeva za prikupljanje podataka

Pregledom API dokumentacije, odabran je zahtjev „*Paintings by Artist*“ čiji odgovor vraća niz slika specificiranog autora u *.json* formatu (*Wikiart Api – WikiArt.org, n.d.*). Odgovor servera za svakog umjetnika je spremljen na računalo, a zatim su slike preuzete u *.jpg* ili *.png* formatu pomoću *Python* skripte *download_images.ipynb* (Kôd 3.2)

```
download_images.ipynb • | leonardo-da-vinci.json
wikiart_dataset > download_images.ipynb > # List of API responses for 10 most influential artists
+ Code + Markdown | ▶ Run All | Clear Outputs of All Cells | Restart | Interrupt | Variables | Outline ...

import os
import json
import uuid
import urllib.request

[7]

# List of API responses for 10 most influential artists
responses = ['leonardo-da-vinci.json', 'pablo-picasso.json', 'vincent-van-gogh.json', 'rembrandt.json', 'peter-paul-rubens.json',
            'albrecht-durer.json', 'titian.json', 'claude-monet.json', 'salvador-dali.json', 'paul-cezanne.json']

for file in responses:

    # The path to the output folder with downloaded images
    parent_dir = "C:\\Users\\Katarina\\Desktop\\kod\\wikiart_dataset\\dataset\\"

    # Extract name of the artist and create appropriate folder directory
    artist_dir = os.path.basename(file)
    artist_dir = artist_dir.replace(".json", "")
    artist_dir += "\\"
    path = os.path.join(parent_dir, artist_dir)
    print(path)
    os.mkdir(path)

    # Read a list of works
    with open(file, 'r', encoding='utf8') as f:
        list_of_works = json.load(f)

        for artwork in list_of_works:

            # Grab artwork url
            url = artwork['image']

            # Determine image type
            if ".png" in url:
                extension = ".png"
            else:
                extension = ".jpg"

            # Create unique file name
            image_name = uuid.uuid4().hex

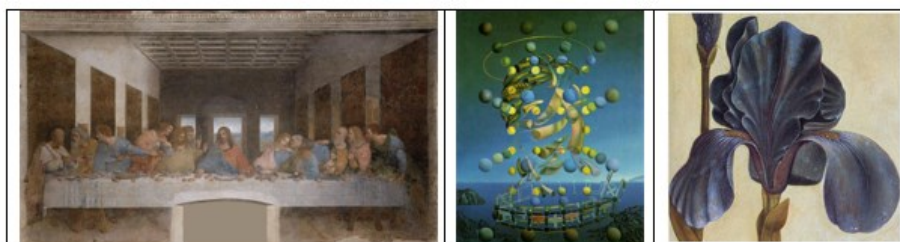
            # Create complete image path
            image = path + image_name + extension
            print(image)

            try:
                urllib.request.urlretrieve(url, image)
            except:
                pass

[28]
```

Kôd 3.2 *Python* kôd za preuzimanje slika u *.jpg* i *.png* formatu

Sve prikupljene fotografije su zatim prenesene u *Google drive*, servis za pohranu i dijeljenje datoteka u oblaku. Na ovaj način prikupljeno je 8.159 fotografija deset autora (Slika 3.2).



Slika 3.2 Primjer nekoliko fotografija umjetničkih djela preuzetih s *WikiArt-a*

3.2. Istraživačka analiza i priprema podataka

Preuzete fotografije uključivale su umjetnička djela izrađena različitim tehnikama. Iz tog razloga, prikupljeni podaci razdvojeni su u dvije mape. Čista verzija skupa podataka uključuje samo radove koji koriste kompleksnije slikarske i crtačke tehnike, dok su jednostavnije skice, poster i skulpture isključeni iz analize (Slika 3.3). Pretpostavka je da autori mogu biti skloni češćem korištenju određenih boja ili poteza što neuronska mreža može matematički prepoznati kao umjetnički stil, koristeći distribuciju boja ili ponavljanje geometrijskih oblika. Odabrane tehnike bi trebale imati najizraženiji umjetnički stil, a ujedno i predstavljaju najviše proučavani dio vizualne umjetnosti (Castellano i Vessio, 2021).



Slika 3.3 Primjer nekoliko fotografija koje su isključene iz analize

U teorijskom dijelu rada spominje se da analizu može otežati **niska inter-varijabilnost** među različitim umjetnicima. Na primjer, **umjetnički pokret** je tendencija ili stil u umjetnosti s određenom zajedničkom filozofijom ili ciljem, koju slijedi skupina umjetnika tijekom određenog vremenskog razdoblja (obično nekoliko mjeseci, godina ili desetljeća) (*Art movement*, 2022). Kako bismo prepoznali utjecaj ove pojave, svakom autoru je dodijeljen pripadajući umjetnički pokret. Pretpostavka je da će modeli raditi najviše pogrešaka u razlikovanju autora koji pripadaju istom ili sličnom umjetničkom pokretu kao što su Vincent Van Gogh (postimpresionizam), Paul Cézanne (postimpresionizam) i Claude Monet (impresionizam); Leonardo da Vinci (visoka renesansa), Titian (visoka renesansa) i Albrecht Durer (sjevernoeuropska renesansa) ili Rembrandt Van Rijn (barok) i Peter Paul Rubens (barok). Odabrani skup podataka također sadrži djela Pabla Picassa, umjetnika koji je poznat po **visokoj intra-varijabilnosti** i promjenama vlastitog umjetničkog stila. Tu je također moguće očekivati manju točnost klasifikacije. Tablica 3.1. sadrži detaljniji opis istraživačkog uzorka.

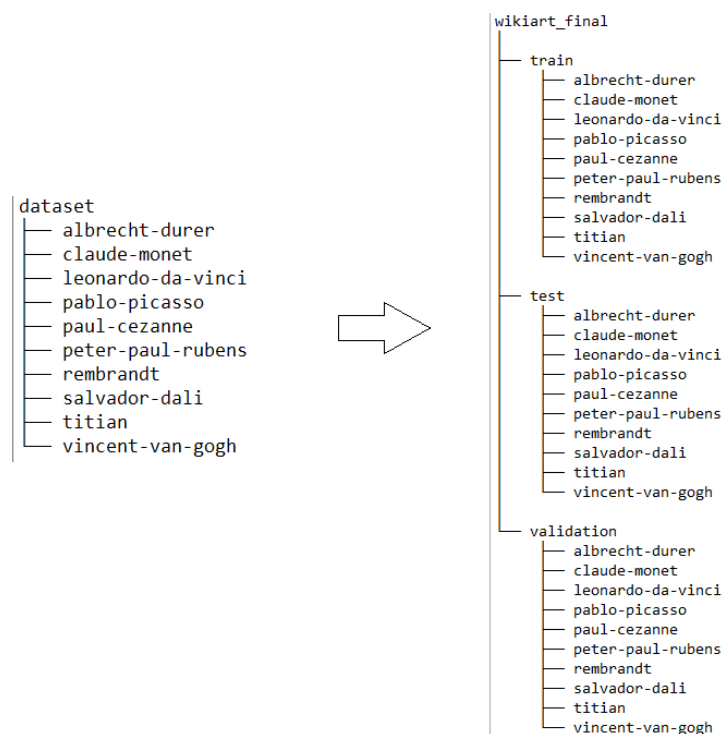
Tablica 3.1 Pregled dostupnog istraživačkog uzorka

Rang	Slikar	Umjetnički pokret	Preuzeto	Isključeno
1	Leonardo da Vinci	Visoka renesansa	172	29
2	Pablo Picasso	Kubizam	989	10
3	Vincent Van Gogh	Postimpresionizam	1896	80
4	Rembrandt Van Rijn	Barok	697	0
5	Peter Paul Rubens	Barok	609	1
6	Albrecht Durer	Sjevernoeuropska renesansa	789	23
7	Titian	Visoka renesansa	236	7
8	Claude Monet	Impresionizam	1355	0
9	Salvador Dali	Nadrealizam	838	24
10	Paul Cézanne	Postimpresionizam	578	2

Podatke je potrebno pripremiti kako bi se mogli koristiti za duboko učenje. Od 8.159 prikupljenih fotografija, samo je 1400 korišteno za trening i izradu modela za detekciju individualnog umjetničkog stila. Ovakav pristup je odabran zbog specifičnosti istraživačkog pitanja. Stvaranje umjetnosti je kreativan proces te se autori često značajno razlikuju po broju umjetničkih djela. Nekim autorima su potrebni mjeseci da bi dovršili jedno djelo. Na primjer, ulju na platnu je potrebno nekoliko tjedana da bi se u potpunosti osušilo. S druge strane, moguće je završiti nekoliko crteža unutar tjedan dana. Moguće je i da su najutjecajniji autori produktivniji od drugih, procjena je da će prosječni umjetnik tijekom karijere stvoriti najmanje jednu do sto slika. Budući da je broj umjetničkih djela po autoru gotovo uvijek ograničen, cilj je stvoriti model koji bi unatoč tome imao što veću moć generalizacije koristeći tehnike za kontrolu prekomjernog prilagođavanja. Osnovni podaci su podijeljeni u tri skupa: **skup za učenje, provjeru i ispitivanje** (Slika 3.4). Slike iz skupa za učenje (engl. *train*) korištene su za trening parametara modela, s ciljem da model što bolje prati podatke iz uzorka. Za ovaj skup je nasumično odabrano 100 umjetničkih djela po autoru.

Slike iz seta za provjeru koriste se za prilagođavanje hiperparametara modela (engl. *validation*). Za ovaj skup je nasumično odabrano 20 umjetničkih djela po autoru.

Slike iz skupa za ispitivanje ili testiranje koristimo za evaluaciju modela. Ovaj skup podataka predstavlja primjer novih fotografija umjetničkih djela koji model nije prethodno vidio, zbog čega mogu pružiti realističniju mjeru uspješnosti (engl. *test*). Za ovaj skup je također nasumično odabrano 20 umjetničkih djela po autoru. Podjela je napravljena pomoću biblioteke *splitfolders* i nalazi se u *Python* skripti *dataset_split.ipynb* (Kôd 3.3).



Slika 3.4 Struktura pripremljenog skupa podataka

```
Train - validation - test split

1) import splitfolders

2) input_folder = "C:\\Users\\Katarina\\Desktop\\kod\\wikiart_dataset\\dataset_clean\\"

3) # split val/test with a fixed number of items
   splitfolders.fixed(input_folder, output="C:\\Users\\Katarina\\Desktop\\kod\\wikiart_dataset\\wikiart_final\\",
                      seeds=2, fixed=(100, 20),
                      oversample=False, group_prefix=None)

4)
```

Kôd 3.3 Podjela podataka na skup za učenje, testiranje i validaciju.

Prethodno trenirani modeli očekuju ulazne slike koje su normalizirane na jednak način, na primjer, grupe fotografija koje koriste RGB model boja u obliku (3 x visina x širina), gdje visina i širina fotografije moraju iznositi najmanje 224 (PyTorch, 2022). Slike se dodaju u rasponu od nula do jedan te su nakon toga trebaju biti normalizirane koristeći specifične vrijednosti aritmetičke sredine [0.485, 0.456, 0.406] i standardne devijacije [0.229, 0.224, 0.225] (PyTorch, 2022). U nastavku je prikazana priprema ulaznih parametara za sva tri analizirana modela (Kôd 3.4).

```
# Data augmentation and normalization on training data
# Normalization for validation and test
data_transforms = {
    'train': transforms.Compose([
        transforms.Resize((224, 224)),
        transforms.RandomHorizontalFlip(),
        transforms.RandomRotation(15),
        transforms.ToTensor(),
        transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
    ]),
    'val': transforms.Compose([
        transforms.Resize((224, 224)),
        transforms.ToTensor(),
        transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
    ]),
    'test': transforms.Compose([
        transforms.Resize((224, 224)),
        transforms.ToTensor(),
        transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
    ]),
}

[ ] data_dir = 'drive/MyDrive/diplomski_rad/dataset_split'
dir_types = ['train', 'val', 'test']
image_datasets = {x: datasets.ImageFolder(os.path.join(data_dir, x),
                                             data_transforms[x])
                  for x in dir_types}

[ ] dataloaders = {x: torch.utils.data.DataLoader(image_datasets[x], batch_size=4,
                                                  shuffle=True, num_workers=2)
                  for x in dir_types}

[ ] dataset_sizes = {x: len(image_datasets[x]) for x in dir_types}
dataset_sizes
{'train': 1000, 'val': 200, 'test': 200}

class_names = image_datasets['train'].classes
class_names
['albrecht-durer',
 'claude-monet',
 'leonardo-da-vinci',
 'pablo-picasso',
 'paul-cezanne',
 'peter-paul-rubens',
 'rembrandt',
 'salvador-dali',
 'titian',
 'vincent-van-gogh']

[ ] # utilize GPU
device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")
```

Kôd 3.4 Kôd koji prikazuje ulazne parametre za sva tri modela.

3.3. Osnovna ResNet arhitektura i njezina obilježja

ResNet ili rezidualna neuronska mreža prilagođeni je oblik konvolucijske neuronske mreže koji je razvijen kako bi se ublažio problem iščezavajućih gradijenata. Ukoliko mreža ima veći broj slojeva, povratna propagacija slojeva ima tendenciju lošije prosljeđivati

informacije nižim slojevima. Gradijenti za svaki parametar mreže trebaju pasti na nulu tek na kraju treniranja, kada je mreža naučila što treba. Ukoliko se to dogodi ranije u početnoj fazi ili na sredini treniranja to može zaustaviti ili značajno usporiti proces treniranja.

Osnovna prednost ove arhitekture u odnosu na prijašnje je uvođenje veza prečaca identiteta (*identity shortcut connection*) koje mogu preskočiti skupinu slojeva mreže. Povezuju se dva udaljena sloja bez uključivanja slojeva između njih. Takve prečice grade rezidualne blokove ili blokove identiteta, koji imaju mogućnost brzog prijenosa aktivacije sloja na dublji sloj u neuronskoj mreži. Skup ovih arhitektura (Slika 3.5) prvi je put predstavljen u radu *Deep Residual Learning for Image Recognition* (He et al., 2016).

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10 ⁹	3.6×10 ⁹	3.8×10 ⁹	7.6×10 ⁹	11.3×10 ⁹

Slika 3.5 Arhitekture dubokih rezidualnih mreža predložene za ImageNet (He et al., 2016)

Na temelju rezultata prethodnih istraživanja, ovaj rad koristi ResNet50, konvolucijsku neuronsku mrežu koja je duboka pedeset slojeva. S obzirom na smanjeni uzorak koji je korišten za istraživanje, preuzeti su i rezultati parametara koji su dobiveni prethodnim treniranjem na više od milijun slika iz ImageNet skupa podataka. Zahvaljujući gotovim parametrima mreža odmah može klasificirati slike u tisuću kategorija, kao što su tipkovnica, olovka ili različite vrste životinja. Pretpostavka je da takvo znanje može doprinijeti rješavanju zadanog klasifikacijskog problema, budući da je likovni motiv jedan od važnijih obilježja umjetničkog stila. Motivi ili sadržaj umjetničkog djela nerijetko odgovaraju karakteristikama umjetnika i nose njegov umjetnički potpis (*Wikipedia contributors*, 2013).

3.3.1. Treniranje i modifikacija ResNet modela za prepoznavanje autora

```
▶ model_ft = models.resnet50(pretrained=True)
  num_ftrs = model_ft.fc.in_features
  # generalized
  model_ft.fc = nn.Linear(num_ftrs, len(class_names))

  model_ft = model_ft.to(device)

  criterion = nn.CrossEntropyLoss()

  # Observe that all parameters are being optimized
  optimizer_ft = optim.SGD(model_ft.parameters(), lr=0.001, momentum=0.9)

  # Decay LR by a factor of 0.1 every 7 epochs
  exp_lr_scheduler = lr_scheduler.StepLR(optimizer_ft, step_size=7, gamma=0.1)
```

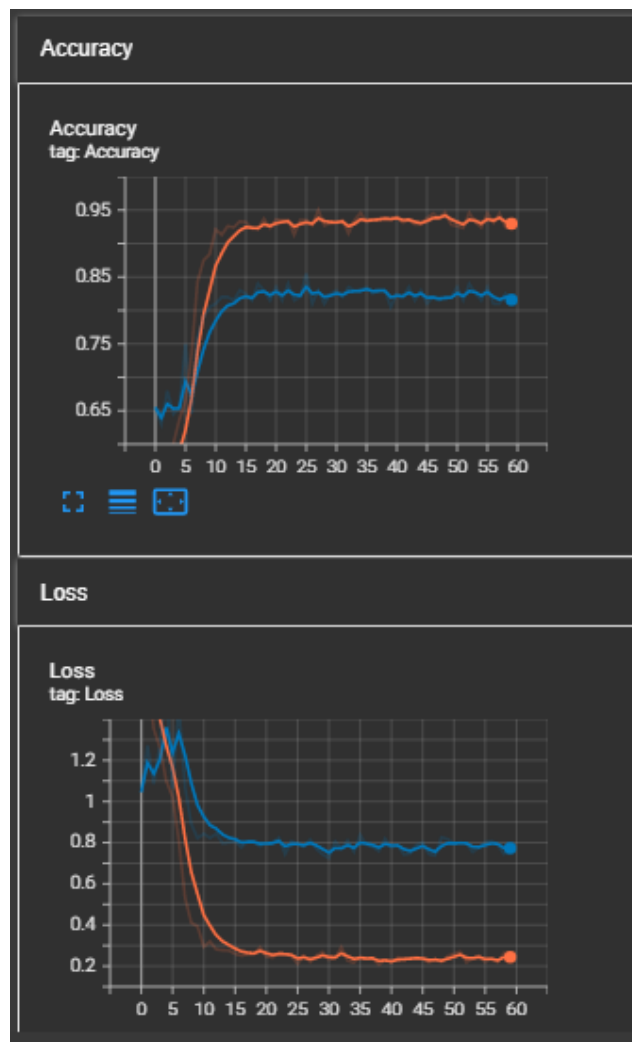
Kôd 3.5 Pozivanje Resnet50 modela.

Steven set al. navode da velike neuronske mreže sa stotinama parametara mogu imati dobru moć generalizacije na malenim skupovima podataka s dosta pogreške (2020). Sva tri modela su trenirana kroz 60 epoha. Parametar stope učenja (engl. *learning rate*) je na početku postavljen na 0.001 (Kôd 3.5). Stopa učenja se zatim za svaku grupu parametara smanjuje za *gammu* prilikom svake epohe i veličine koraka (*step size*) (*Introduction to PyTorch*, 2022). *cs231* predlaže da se krene s većim parametrom stope učenje jer brže optimizira u odnosu na „najbolji paramater učenja“, no u nekom trenu je nužno smanjiti parametar stope učenja kako koraci koji se odvijaju ne bi bili preveliki – funkcija bez toga ne bi mogla doseći minimum (2022).

Funkcija gubitka (engl. *loss*) pokazuje kolika je udaljenost od točnih predviđanja na izlazu korištene neuronske mreže, pri čemu niža vrijednost označava bolji rezultat. Prosječna vrijednost funkcije gubitka prati se tijekom treninga kako bi se mogao pratiti tijek učenja (CS231n: Deep Learning for Computer Vision, 2022). U ovu svrhu, korištena je funkcija gubitka unakrsne entropije.

3.3.2. Evaluacija modela

Budući da se radi o klasifikacijskom problemu, korisno je da rezultati procjene modela poprimaju vrijednosti između nule i jedinice kako bi mogli predstavljati vjerojatnost za svaku klasu. Budući da se radi o multi-klasifikacijskom problemu korištena je *Softmax* funkcija. Koristimo ju kada svakoj slici može biti dodijeljena jedna kategorija.



Slika 3.6 Prikaz točnosti i funkcije gubitka za model ResNet50.

Prikazom točnosti i funkcije gubitka vidljiv je problem prekomjernog prilagođavanja. Vidimo da skup podataka za učenje pokazuje izrazito visoku točnost u odnosu na skup za validaciju. Funkcija gubitka na validacijskom skupu podataka ne teži nuli. Predmetni graf pokazuje da korišteni model nije optimalan za detekciju individualnog umjetničkog stila (Slika 3.6). U pravilu, kada mreža pokuša učiti iz izrazito malog skupa podataka, nastojat će obuhvatiti sve varijacije podataka umjesto generalnog trenda.

Unaprjeđenje modela za prepoznavanje autora umjetničkog djela

4. U ovom odlomku rezultati dobiveni na modelu ResNet50 uspoređuju se s dvije modernije arhitekture - ResNext i DenseNet. Usporedba ne pokazuje značajne razlike među modelima, odnosno tip arhitekture mreže sam po sebi ne dovodi do razlika prilikom detekcije individualnog umjetničkog stila.

4.1. Korištenje ResNeXt arhitekture

ResNext modeli predstavljaju novu generaciju rezidualnih neuronskih mreža te se smatraju učinkovitijima i točnijima. Prvi put su predstavljene u radu *Aggregated Residual Transformations for Deep Neural Networks* (Xie et al., 2017) (Slika 4.1).

stage	output	ResNet-50	ResNeXt-50 (32×4d)
conv1	112×112	7×7, 64, stride 2	7×7, 64, stride 2
		3×3 max pool, stride 2	3×3 max pool, stride 2
conv2	56×56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128, C=32 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256, C=32 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
conv4	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512, C=32 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
conv5	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 1024 \\ 3 \times 3, 1024, C=32 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	global average pool 1000-d fc, softmax	global average pool 1000-d fc, softmax
# params.		25.5×10^6	25.0×10^6
FLOPs		4.1×10^9	4.2×10^9

Slika 4.1 Usporedba ResNet-50 modela s ResNeXt50_32x4d arhitekturom (Xie et al., 2017).

ResNeXt arhitektura uvodi novi hiperparametar **kardinalnost** za podešavanje kapaciteta modela. Xie et al. opisuju kardinalnost kao veličinu skupa transformacija koja postaje važan faktor uz dimenzije dubine i širine (2017). Također, ovaj oblik arhitekture koristi split-

transform-merge paradigmu koja se prvi put pojavljuje u *Inception* modelu (Xie et al., 2017). Koriste se razgranati putevi unutar jednog modula, odnosno, transformacije se zbrajaju u mreži.

4.1.1. Treniranje i modifikacija ResNeXt modela

```
▶ model_ft = models.resnext50_32x4d(pretrained=True)
  num_ftrs = model_ft.fc.in_features
  # generalized
  model_ft.fc = nn.Linear(num_ftrs, len(class_names))

  model_ft = model_ft.to(device)

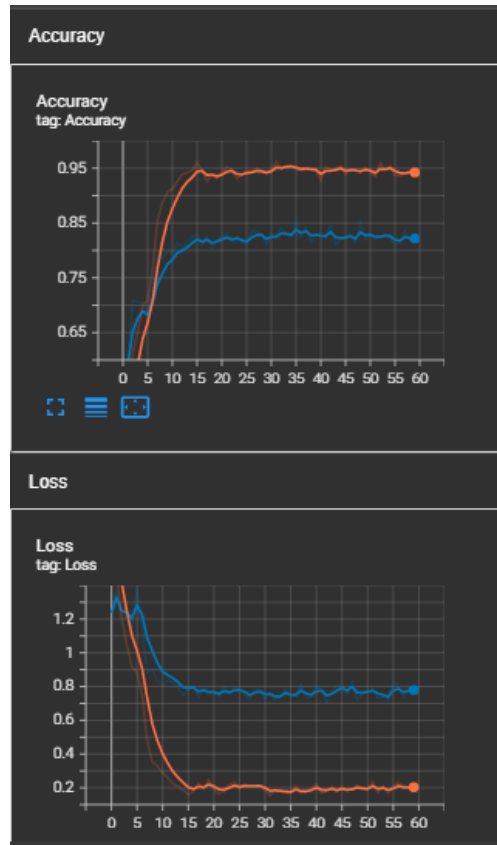
  criterion = nn.CrossEntropyLoss()

  # Observe that all parameters are being optimized
  optimizer_ft = optim.SGD(model_ft.parameters(), lr=0.001, momentum=0.9)

  # Decay LR by a factor of 0.1 every 7 epochs
  exp_lr_scheduler = lr_scheduler.StepLR(optimizer_ft, step_size=7, gamma=0.1)
```

Kôd 4.1 Pozivanje ResNeXt50_32x4d modela.

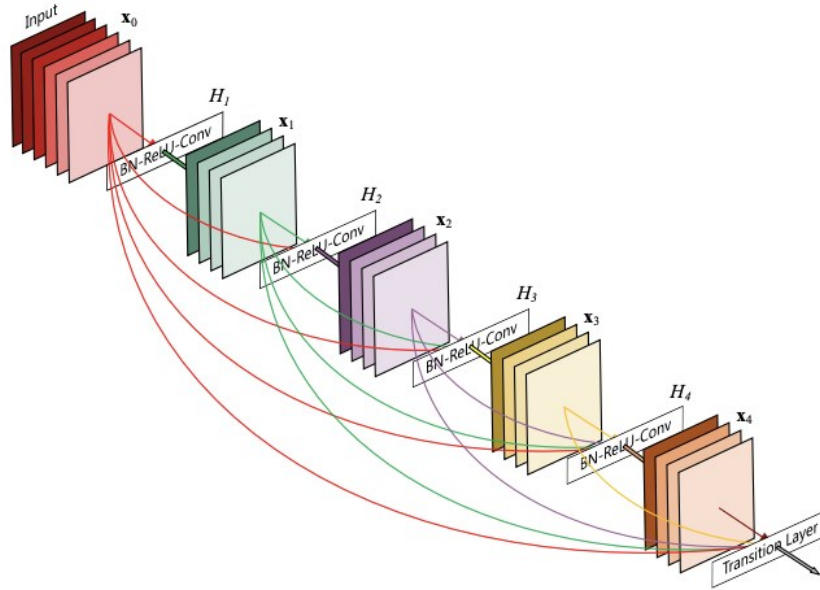
4.1.2. Evaluacija modela



Slika 4.2 Prikaz točnosti i funkcije gubitka za model resnext50_32x4d.

4.2. Korištenje DenseNet arhitekture

DenseNet ili guste neuronske mreže uvedene su u radu *Densely Connected Convolutional Networks* (Huang et. al., 2017). Karakteristika ovih modela je da je svaki sloj povezan sa svakim drugim bez povratnih veza (engl. *feed-forward*). Ova vrsta mreža ima jedan ulazni i izlazni sloj te više skrivenih slojeva neurona (Slika 4.3).



Slika 4.3 Prikaz slojeva u DenseNet mreži (Huang et al., 2017)

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Slika 4.4 Predložene DenseNet arhitekture za ImageNet (Huang et al., 2017).

DenseNet arhitektura je također osmišljena kako bi se ublažilo smanjenje točnosti uzrokovano problemom iščezavajućih gradijenata u dubokim neuronskim mrežama. Zbog duljeg puta između ulaznog i izlaznog sloja, može se dogoditi da informacija nestane prije nego što stigne na odredište. Osnovna razlika između ResNet i DenseNet arhitekture je da ResNet arhitekture koriste zbrajanje za povezivanje svih prethodnih mapa značajki dok ih DenseNet sve direktno spaja (Huang et al., 2017) (Slika 4.4).

4.2.1. Treniranje i modifikacija DenseNet modela

```
model_ft = models.densenet121(pretrained=True)
num_fts = model_ft.classifier.in_features
# generalized
model_ft.classifier = nn.Linear(1024, len(class_names))

model_ft = model_ft.to(device)

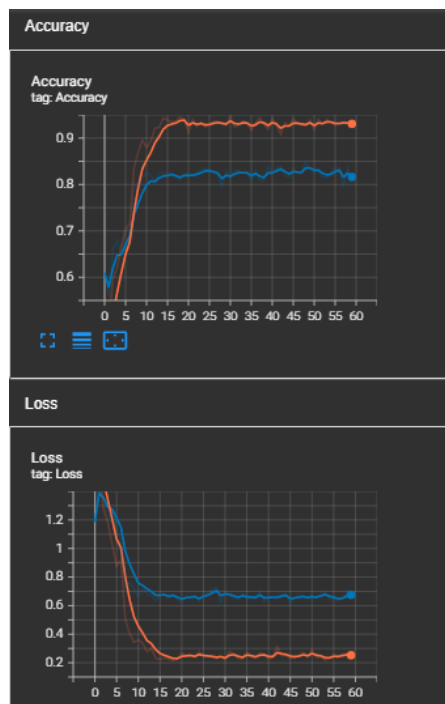
criterion = nn.CrossEntropyLoss()

# Observe that all parameters are being optimized
optimizer_ft = optim.SGD(model_ft.parameters(), lr=0.001, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_ft, step_size=7, gamma=0.1)
```

Kôd 4.2 Pozivanje *DenseNet* modela..

4.2.2. Evaluacija modela



Slika 4.5 Prikaz točnosti i funkcije gubitka za model densenet121.

Tablica 4.1 Usporedba rezultata treniranja različitih modela.

Model	Točnost (%)
resnet50	0.8550
resnext50_32x4d	0.8500
densenet121	0.8500

Točnost (engl. *accuracy*) odnosi se na udio točno klasificiranih primjera u skupu svih primjera. Iako je ova metrika dovoljna za odabir modela koji će se prikazivati u jednostavnoj web aplikaciji za detekciju individualnog umjetničkog stila, u praksi se rijetko koristi samostalno jer ne daje potpunu sliku o rezultatima analize.

Uz točnost, Dalbello Bašić i Šnajder definiraju i ostale najvažnije metrike za klasifikacijske probleme:

- **Preciznost** (engl. *precision*) odnosi se na udio točno klasificiranih primjera u skupu pozitivno klasificiranih primjera.
- **Odziv** (engl. *recall*) je udio točno klasificiranih primjera u skupu svih pozitivnih primjera. Često se naziva i osjetljivost. Želimo da odziv bude što veći (na primjer, želimo otkriti što veći broj slika koje su pod utjecajem Picassa)
- **Specifičnost** (engl. *specificity*) je udio točno klasificiranih primjera u skupu svih negativnih primjera. Ako je specifičnost 100% onda su sva Picassova umjetnička djela prepoznati kao njegov umjetnički stil. Manja specifičnost znači da je dio umjetničkih djela pod utjecajem Pabla Picassa pripisan drugim autorima.

Uz sve navedeno, često se prikazuje i **matrica zabune** (engl. *confusion matrix*).

Prototip web rješenja za detekciju stupnja sličnosti individualnog umjetničkog stila

5. Treći cilj rada je izrada web rješenja koje će korisniku omogućiti postavljanje vlastite fotografije umjetničkog djela i vizualnu usporedbu umjetničkog stila u odnosu na deset najutjecajnijih autora. Ova ideja inspirirana je procjenama stručnjaka, gdje je nerijetko moguće pročitati da djela mnogih autora nastaju pod snažnim utjecajem nekog određenog umjetnika ili umjetničkog pokreta. U analizi umjetnosti, umjetnička djela se rijetko promatraju kao izolirana tvorevina, stručnjaci ih obično proučavaju u širem kontekstu, uključujući utjecaj i poveznice među različitim umjetničkim školama (Castellano & Vessio, 2021). Prototip web rješenja je izrađen u programskom jeziku *Python* pomoću radnog okvira *Streamlit* te je upogonjen pomoću servisa *Google Colab*. U ovom se odlomku detaljno opisuju koraci izrade, prikazan je izgled sučelja te kratka evaluacija kroz nekoliko praktičnih primjera. U pozadini aplikacije je model koji koristi ResNet50 arhitekturu budući da je pokazao najbolje rezultate po mjerama evaluacije.

5.1. Izrada prototipa koristeći Streamlit

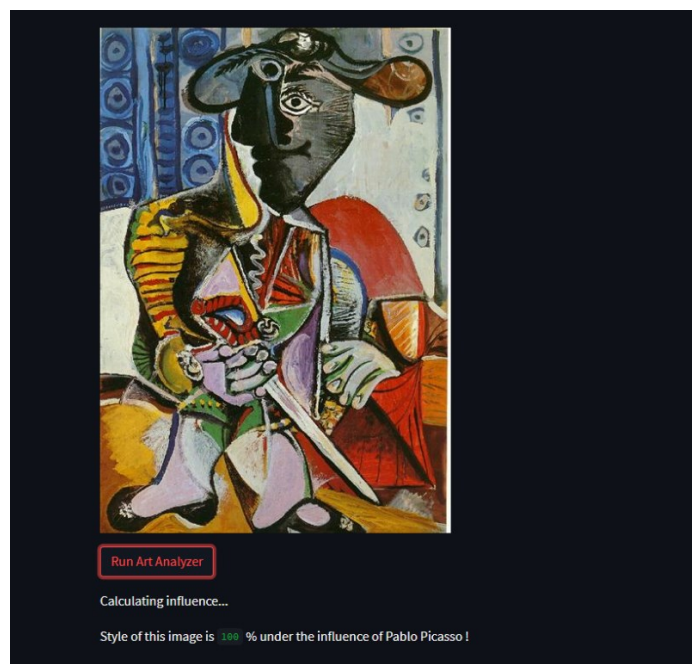
Streamlit je javno dostupna *Python* biblioteka specijalizirana za jednostavno stvaranje i dijeljenje prototipa web aplikacija iz područja strojnog učenja i znanosti o podacima (*Streamlit docs*, 2022).

U aplikaciju je moguće dodati vlastitu fotografiju digitalnog umjetničkog djela u *.jpg* ili *.png* formatu. To je moguće napraviti ispuštanjem datoteke u aplikaciju ili pretraživanjem direktorija na računalu. Slika 5.1 prikazuje sučelje aplikacije koje korisnik vidi prilikom otvaranja.

Nakon što korisnik odabere fotografiju, ista mu se prikazuje na ekranu (Slika 5.2). Potom slijedi transformacija fotografije kako bi odgovarala pretpostavkama modela za duboko učenje. Klasifikacija se odvija u pozadini za stil deset najutjecajnijih autora, najbolji rezultat ili najjači odgovor prema funkciji *Softmax* množi se sa sto kako bi bio u formatu postotka.



Slika 5.1 Sučelje aplikacije.



Slika 5.2 Primjer rezultata klasifikacije pomoću aplikacije *Art Analyzer*.

5.2. Upogonjavanje modela koristeći Google Colab

Kako bi se izbjeglo preuzimanje datoteka na osobno računalo, aplikaciju je moguće reproducirati i pokrenuti koristeći *Google Colab*. U radu (Nneji, 2022) opisani su potrebni koraci za pokretanje ili upogonjavanje modela:

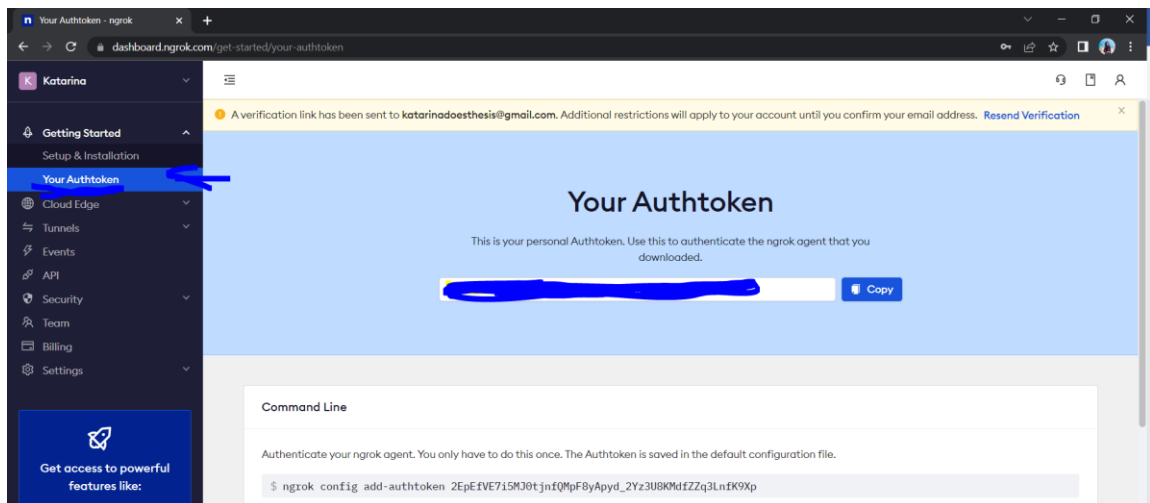
Prvo je potrebno instalirati *streamlit* i *pyngrok* biblioteke koristeći *Google Colab* (Kôd 5.1).

```
!pip install -q streamlit
!pip install pyngrok
```

Kôd 5.1 Instalacija biblioteka.

Zatim je potrebno izraditi besplatni korisnički račun na stranici ngrok.com.

Nakon prijave, potrebno je potražiti i kopirati vlastiti token (Slika 5.3)



Slika 5.3 Izgled koda za autentifikaciju.

Pokreće se autentifikacija pomoću naredbi u *Google Colab-u* (Kôd 5.2)

```
from pyngrok import ngrok
!./ngrok authtokens token
public_url = ngrok.connect(port='8501')
public_url
```

Kôd 5.2 Naredbe za autentifikaciju

Napomena: Riječ **token** potrebno je zamijeniti s vlastitim kopiranim kôdom.

Kao sljedeći korak pokreće se *python* datoteka koja sadrži kôd koji se izvodi u pozadini web aplikacije (Kôd 5.3).

```
%%writefile streamlit_app.py
import streamlit as st
import io
import os
from PIL import Image
import torch
import torch.nn as nn
import torch.nn.functional as F
from torch.utils.data import DataLoader
from torchvision import datasets, transforms, models

def load_image():
    uploaded_file = st.file_uploader(label='Pick an image for
analysis', type=['jpg', 'png'])
    if uploaded_file is not None:
        image_data = uploaded_file.getvalue()
        st.image(image_data)
        return Image.open(io.BytesIO(image_data))
    else:
        return None

def load_model():
    model = models.resnet18(pretrained=True)
    num_ftrs = model.fc.in_features
    model.fc = nn.Linear(num_ftrs, 10)
    model_save_name = 'ResNet50.pt'
    path = F"drive/MyDrive/diplomski_rad/{model_save_name}"
    model.load_state_dict(torch.load(path,
map_location='cpu'))
    model.eval()
    return model

def predict(model, categories, image):
    preprocess = transforms.Compose([
        transforms.Resize((224, 224)),
        transforms.ToTensor(),
        transforms.Normalize([0.485, 0.456, 0.406], [0.229,
0.224, 0.225])
```

```

    ])
    input_tensor = preprocess(image)
    input_batch = input_tensor.unsqueeze(0)

    with torch.no_grad():
        output = model(input_batch)

        probability = torch.nn.functional.softmax(output, dim=1)
        top_p, top_class = probability.topk(1, dim = 1)
        st.write('Style of this image is ',
round(top_p.item()*100), '% under the influence of ',
categories[top_class], '!')

def main():
    st.title('Welcome to Art Analyzer!')
    st.subheader('A deep learning powered tool for style
analysis.')
    artists =
Image.open('drive/MyDrive/diplomski_rad/artists_all.PNG')
    st.image(artists, use_column_width=True)
    st.subheader('Upload a photo and see how much the style
of your picture resembles one of the most influential
artists.')

    model = load_model()
    categories = ['Albrecht Durer', 'Claude Monet', 'Leonardo
da vinci', 'Pablo Picasso', 'Paul Cézanne', 'Peter Paul
Rubens', 'Rembrandt', 'Salvador Dali', 'Titian', 'Vincent van
Gogh']
    image = load_image()
    result = st.button('Run Art Analyzer')

    if result:
        st.write('Calculating influence...')
        predict(model, categories, image)

if __name__ == '__main__':
    main()

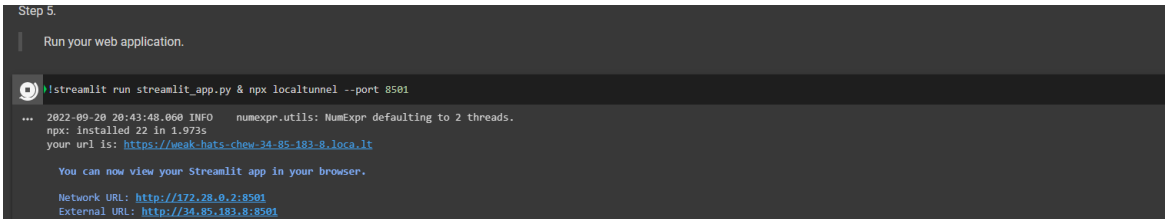
```

Kód 5.3 Kód u pozadini prototipa web aplikacije

Aplikaciju se zatim pokreće iz *Google colab*-a uz pomoć sljedeće naredbe:

```
!streamlit run /content/streamlit_app.py & npx localtunnel --port 8501
```

Link za pokretanje aplikacije bi trebao izgledati slično kao na slici (Slika 5.4):



```
Step 5.
Run your web application.

!streamlit run streamlit_app.py & npx localtunnel --port 8501
...
2022-09-20 20:43:48.060 INFO numexpr.utils: NumExpr defaulting to 2 threads.
npx: installed 22 in 1.973s
your url is: https://weak-hats-chew-34-85-183-8.localtunnel.net
You can now view your Streamlit app in your browser.
Network URL: http://172.28.0.2:8501
External URL: http://34.85.183.8:8501
```

Slika 5.4 Izgled veze za pokretanje aplikacije.

5.3. Primjena modela u stvarnim situacijama – *crash test*

Za razliku od testiranja na drugim umjetničkim djelima koja su također preuzeta sa internetske stranice *WikiArt*, u ovom poglavlju prikazani su rezultati primjene modela u stvarnim situacijama. Prikazano je ponašanje modela u odnosu na šest različitih skupina testiranih objekata, kako bi se istražila potencijalna širina primjene. Promatraju se sljedeće skupine:

- fotografije umjetničkih djela deset autora iz uzorka,
- fotografije više umjetničkih djela koje imaju zajednički motiv,
- fotografije autora koji su potvrdili da su svoja umjetnička djela radili pod jakim utjecajem drugog autora,
- fotografije predmeta na kojima je moguće prepoznati individualni umjetnički stil,
- fotografije objekata koji nisu povezani niti s jednim od autora.

Prvu skupinu testiranih objekata čine digitalne fotografije umjetničkih djela deset autora iz uzorka (Tablica 5.1). Fotografije iz ove skupine odabrane su nasumičnim odabirom prilikom *Google* pretraživanja tih autora te dolaze sa različitih internetskih stranica. Neke od fotografija sadrže prikaz umjetničkih djela koja su obuhvaćena istraživačkim uzorkom, međutim većina fotografija prikazuje umjetnička djela koja nisu korištena prilikom izrade i

testiranja modela. Sve fotografije značajno se razlikuju od originalnih po stupnju kontrasta, osvjetljenja, rezoluciji, veličini i sličnom. Ovaj test pokazuje da je model u pravilu u stanju točno prepoznati autora ili individualni umjetnički stil bez obzira na izvor, format ili kvalitetu fotografije.

Druga skupina testiranih objekata sadrži fotografije više umjetničkih djela koje imaju zajednički motiv, a razlikuju se po individualnom umjetničkom stilu (Tablica 5.2). Na primjer, u ovu skupinu ulazi fotografija s motivom Mona Lise koja je naslikana u Picassovom ili Monetovom stilu. Cilj ovakvog testa je provjeriti u kojoj mjeri se model oslanja na prepoznavanje vizualnog stila u odnosu na detekciju objekta ili motiva. S obzirom da su prilikom transfernog učenja, korištene početne težine prilagođene za detekciju objekata, ovaj test pokazuje da je izvršena uspješna prilagodba modela. Model je u stanju uspješno odvojiti individualni umjetnički stil od motiva i objekata.

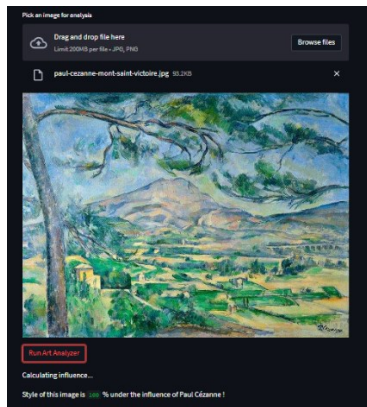
Treću skupinu testiranih objekata čine fotografije autora koji su potvrdili da su svoja umjetnička djela radili pod jakim utjecajem drugog autora (Tablica 5.3). Ovdje je model testiran na umjetničkim djelima koje bi promatrač mogao pripisati jednom od autora iz uzorka. Ovaj oblik testiranja najvažniji je za primjenu ovog rada, budući da pokazuje može li model mjeriti utjecaj jednog autora na drugog, odnosno stupanj sličnosti individualnog umjetničkog stila. Rezultati pokazuju da je model u stanju točno prepoznati utjecaj izabranih autora na pojedino umjetničko djelo. Ipak, model pokazuje prevelik stupanj sigurnosti prilikom procjene, često je to maksimalan postotak sigurnosti koji bi se trebao pojavljivati jedino u slučaju da se analizira djelo izvornog autora.

U četvrtoj skupini promatraju se rezultati modela na fotografijama koje ne prikazuju umjetničke slike, ali je na njima svejedno moguće prepoznati individualni umjetnički stil nekog od autora (Tablica 5.4). Na navedenim fotografijama prikazani su objekti iz svakodnevnog života: majice, namještaj, moda, filmska scena i slično. U ovoj skupini najčešće su testirani promotivni materijali sa slikom umjetničkog djela ili objekti koje su autori stvarali kao izraz divljenja odabranim umjetnicima. Prilikom provođenja ovog testa, zabilježene su različite razine točnosti, no model je i dalje, kod većine fotografija, u mogućnosti točno prepoznati individualni umjetnički stil. Na primjer, testirana je fotografija na kojoj je prikazana majica s otisnutom Van Gogh-ovom slikom. U ovom slučaju umjetničko djelo zauzima približno četvrtinu analizirane slike te sama fotografija uključuje i okolinu kao - stol, biljku i slično. Dobiveni stupanj sličnosti je 44% u odnosu na stil Van Gogha, što se u ovom slučaju može smatrati dobrim rezultatom s obzirom na distrakcije koje

je model morao analizirati. Ova skupina pokazuje potencijal za široku primjenu modela u svakodnevnom životu.

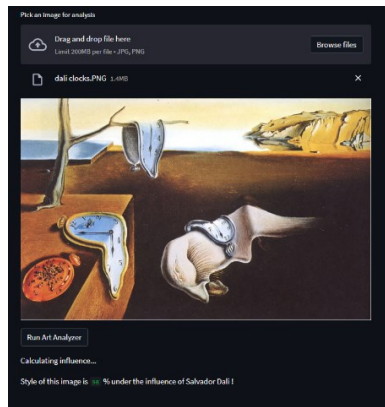
U petoj skupini testirano je kako će se model ponašati ako korisnik testira objekte za koje se pretpostavlja da nisu povezani niti s jednim od autora (Tablica 5.5). U ovoj skupini u aplikaciju su dodane slike čaše mlijeka, logotipi ili predmeti koji predstavljaju čest slikarski model, kao što je košara s voćem. Vidljivo je da je model razvijen tako da uvijek traži stupanj sličnosti s autorima iz uzorka te da ponudi najviši stupanj sličnosti koji može pronaći, bez obzira na stvarni značaj utjecaja. U navedenim situacijama, model obično pokazuje niži stupanj sigurnosti. Na primjer, iako je na prvu teško uočiti povezanost između fotografije čaša mlijeka i Dalijeveg opusa, od svih analiziranih autora njegov opus sadrži najviše fotorealističnih umjetničkih djela.

Tablica 5.1 Prva skupina: umjetnička djela autora iz uzorka, iz drugih izvora



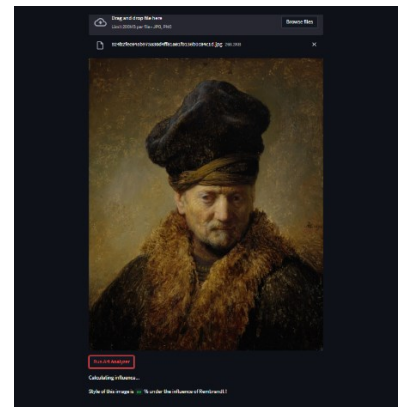
Paul Cézanne – Mont Saint-Victore

Izvor: <https://impressionistarts.com/>



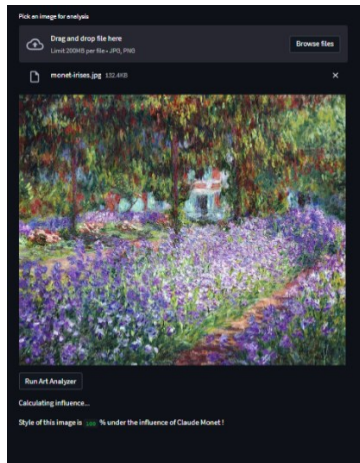
Salvador Dalí – The Persistence of Memory

Izvor: <https://arthive.com/publications/>



Rembrandt: Portrait of a man

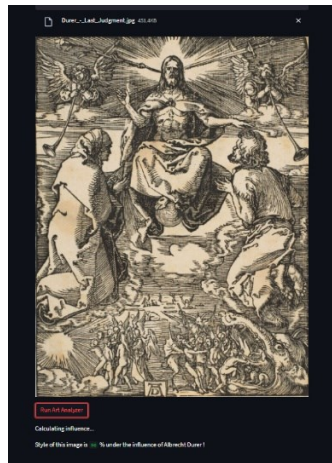
Izvor: <https://www.ngprague.cz/>



Claude Monet – Irises in Monet's Garden

Izvor:

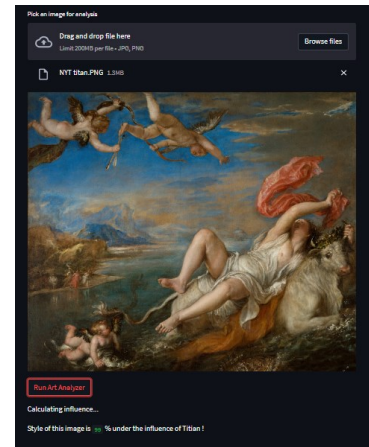
https://www.artfactory.com/art_appreciation/great-artists/claude-monet/claude-monet.html



Albrecht Dürer – The Last Judgement

Izvor:

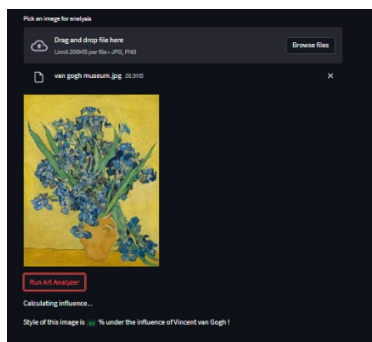
<https://www.illustrationhistory.org/artists/albrecht-duerer>



Titian – The Rape of Europa

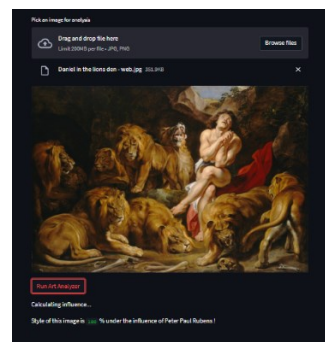
Izvor:

<https://www.nytimes.com/2021/08/12/arts/design/titian-isabella-stewart-gardner-museum-review.html>



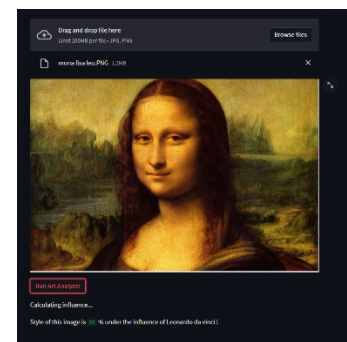
Vincent Van Gogh – Irises

Izvor: <https://www.vangoghmuseum.nl/en/>



Peter Paul Rubens – Daniel in the Lions' Den

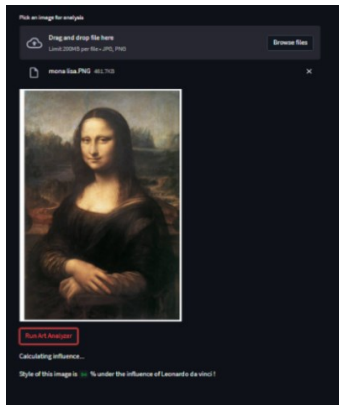
Izvor: <https://ago.ca/exhibitions/early-rubens>



Leonardo da Vinci – Daniel in the Lions' Den

Izvor: <https://niiod.com/the-10-most-famous-artworks-of-leonardo-da-vinci/>

Tablica 5.2 Druga skupina: isti motivi u različitim umjetničkim stilovima

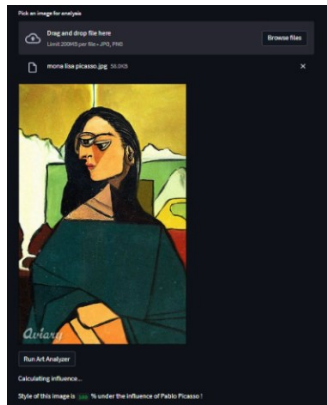


Leonardo da Vinci – Mona Lisa

Stil: **Leonardo da Vinci**

Izvor:

https://commons.wikimedia.org/wiki/File:Leonardo_da_Vinci_-_Mona_Lisa.jpg

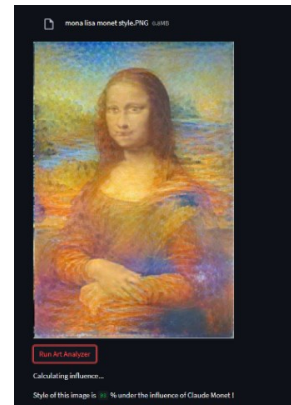


La Giacoda – Mona Lisa

Stil: **Pablo Picasso**

Izvor:

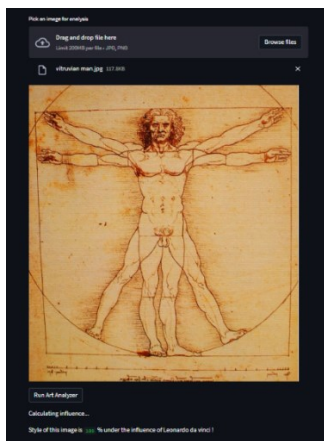
<https://sk.pinterest.com/pin/196751077447190882/>



Experiments with style transfer [2015] – Mona Lisa

Stil: **Claude Monet**

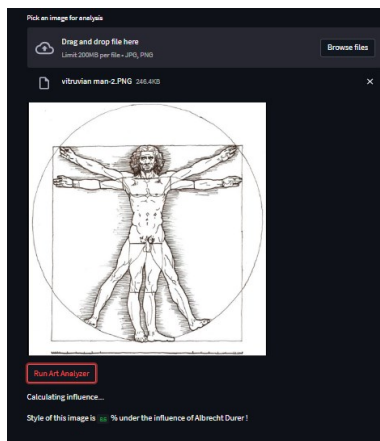
Izvor: <https://genekogan.com/works/style-transfer/>



Leonardo da Vinci – Vitruvian Man

Stil: **Leonardo da Vinci**

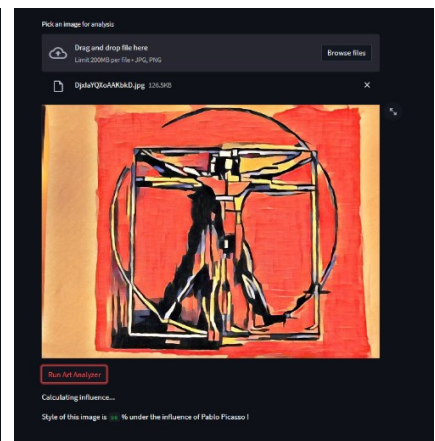
Izvor: <https://www.smithsonianmag.com/smart-news/vitruvian-man-may-have-had-hernia-180949806/>



Style transfer

Stil: **Albrecht Durer**

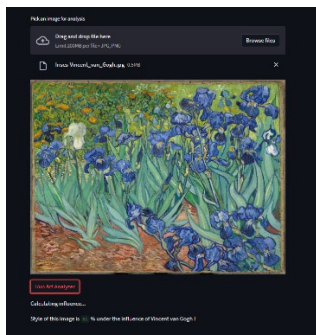
Izvor: <https://twitter.com/>



Amateur adaptation – Vitruvian Man

Stil: **Pablo Picasso**

Izvor: <https://twitter.com/jvmenon/status/1025774447578505221>

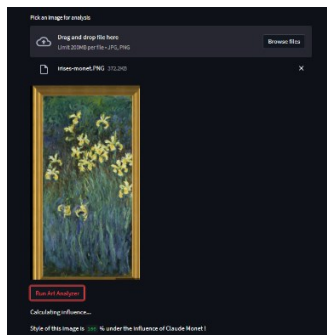


Vincent van Gogh - Irises

Stil: **Vincent van Gogh**

Izvor:

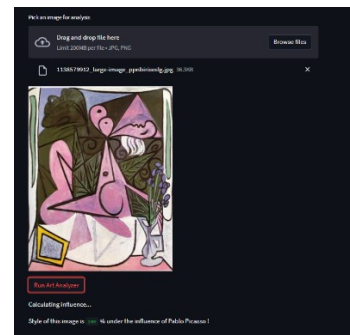
[https://en.wikipedia.org/wiki/Iris_\(van_Gogh_painting\)](https://en.wikipedia.org/wiki/Iris_(van_Gogh_painting))



Claude Monet – Irises

Stil: **Claude Monet**

Izvor: <https://www.encore-editions.com/yellow-irises-by-claude-monet/>

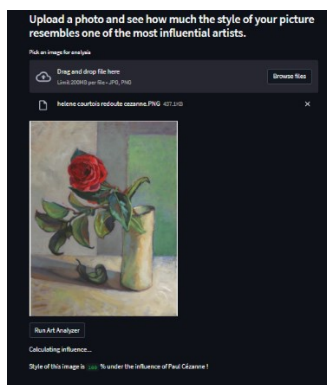
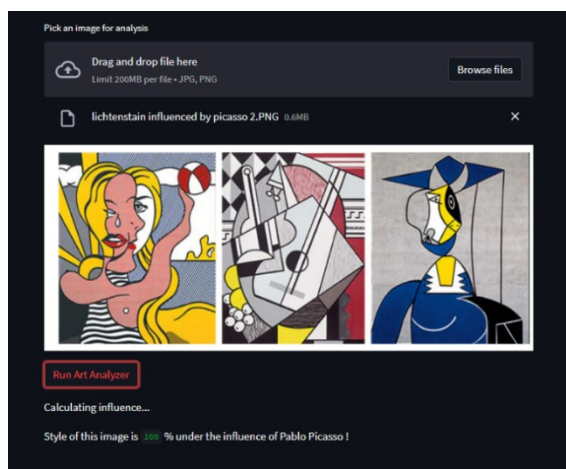


Picasso – nude with a bouquet of irises and a mirror

Stil: **Pablo Picasso**

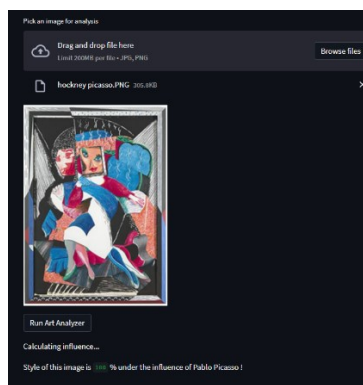
Izvor: <https://www.royal-painting.com/Pablo-Picasso/nude-with-a-bouquet-of-irises-and-a-mirror.html>

Tablica 5.3 Treća skupina: Umjetnička djela stvorena pod utjecajem autora iz uzorka



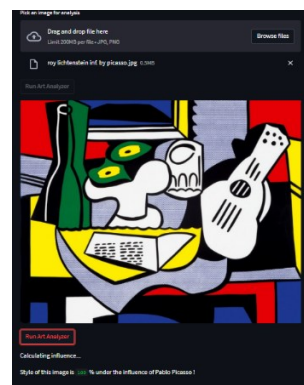
Umjetnica: Hélène Courtois-Redouté
Utjecaj: **Paul Cezanne**

Izvor: <https://www.kazoart.com/en/design-trends/6-artists-inspired-by-cezanne>



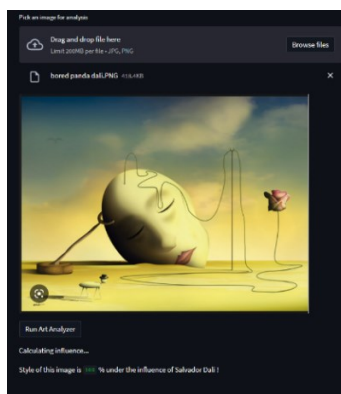
Umjetnik: David Hockney – An Image of Celia
Utjecaj: **Pablo Picasso**

Izvor: <https://artrepublic.com/blogs/news/323-picassos-influence-on-other-artists.html>



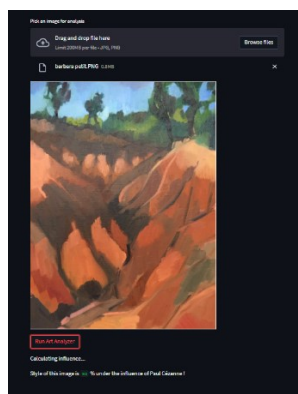
Umjetnik: Roy Lichtenstein – Still life by Picasso
Utjecaj: **Pablo Picasso**

Izvor: https://arthive.com/roylichtenstein/works/482884-Still_life_by_Picasso



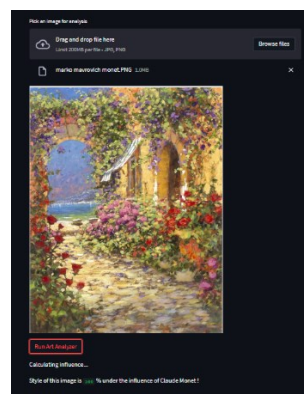
Umjetnik: amateur art, Bored panda
Utjecaj: **Salvador Dali**

Izvor: <https://www.boredpanda.com/surreal-digital-paintings-marcel-caram/>



Umjetnica: Barbara Petit
Utjecaj: **Paul Cezanne**

Izvor: <https://www.kazoart.com/en/design-trends/6-artists-inspired-by-cezanne>



Umjetnik: Marko Mavrovich
Utjecaj: **Claude Monet**

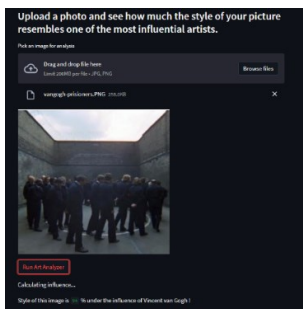
Izvor: <https://www.parkwestgallery.com/park-west-artists-tribute-claude-monet/>

Tablica 5.4 Četvrta skupina: Fotografije koje ne prikazuju umjetnička djela, ali je vidljiv stil

7. *A Clockwork Orange* (1971) and Vincent Van Gogh's *Prisoners Exercising*

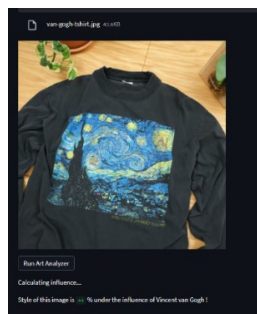


Right: *A Clockwork Orange* (1971) / Left: Vincent Van Gogh's *Prisoners Exercising*



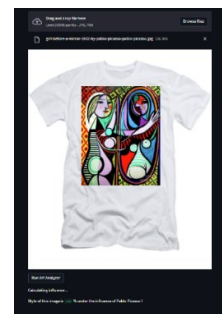
Utjecaj: Vincent Van Gogh

Izvor: <https://www.pinterest.com/pin/22426-5256430901678/>



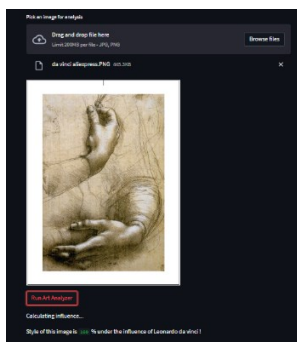
Utjecaj: Vincent Van Gogh

Izvor: <https://www.ebay.com/itm/313938134516>



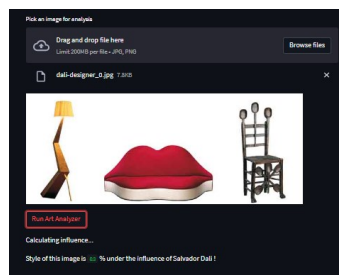
Utjecaj: Pablo Picasso

Izvor: <https://fineartamerica.com/featured/girl-before-a-mirror-1932-by-pablo-picasso-pablo-picasso.html?product=adult-tshirt>



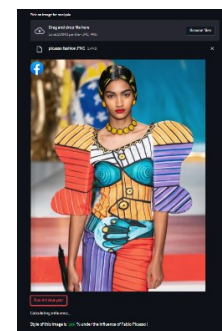
Utjecaj: Leonardo da Vinci

Izvor: <https://www.aliexpress.com/item/2030055461.html>



Utjecaj: Salvador Dalí

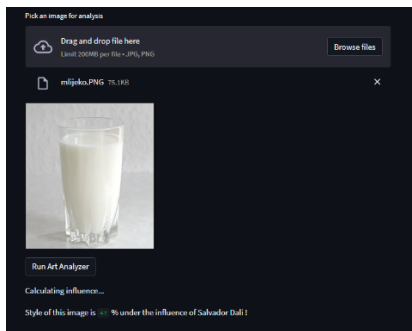
Izvor: <https://www.daliparis.com/en/salvador-dali-en/dali-design/>



Utjecaj: Pablo Picasso

Izvor: <https://www.facebook.com/photo/?fbid=1028135147643263&set=pcb.1028135774309867>

Tablica 5.5 Peta skupina: Nasumično testiranje objekata



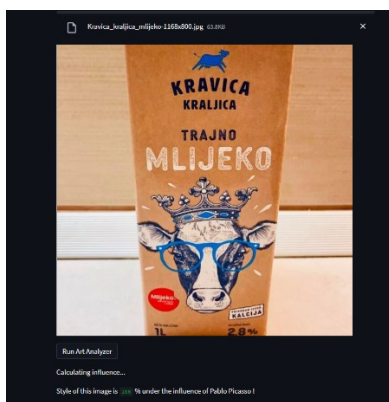
Izvor: <https://en.wiktionary.org/wiki/milk>

A. Salvador Dali – *Lobster Telephone* - <https://www.nationalgalleries.org/art-and-artists/166050>

B. Salvador Dali - *Apparition of Face and Fruit Dish on a Beach* -



© Salvador Dali, Fundació Gala-Salvador Dali, DACS, London, 2022

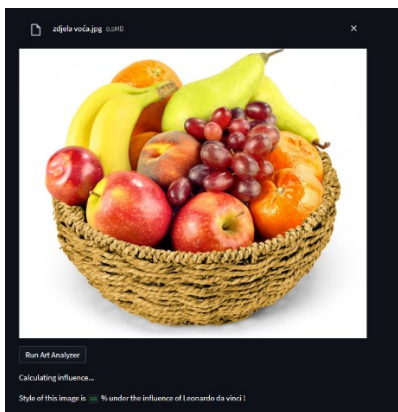


Izvor: <https://danica.hr/belje-spasilu-tvornicu-i-na-trziste-izbacilo-novo-domace-mlijeko/>

A. El Towle – *Picasso Cow IV*
<https://www.miltonartistsguildstore.com/product/el-towle-picasso-cow-iv/4988>

B. Pablo Picasso – *Picasso 85 gravures 1966*
<https://www.mchampetier.com/Poster-Pablo-Picasso-111519-work.html>





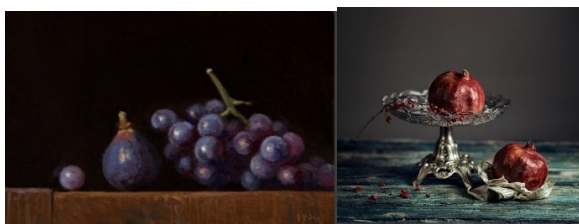
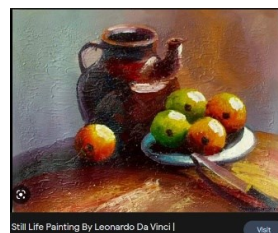
Izvor: <https://www.regencyhampers.com/classic-gift-hampers/classic-fresh-fruit-bowl>

A. Still life painting – Leonardo Da Vinci

B. Abbey Ryan – Still Life with Fig and Concord grapes (utjecaj: Leonardo da Vinci)

<http://new.abbeyryan.com/still-life/still-life-with-fig-and-concord-grapes-no-2-leonardo-da-vinci/>

C. <https://www.pinterest.com/pinterest/leonardo-da-vinci/>



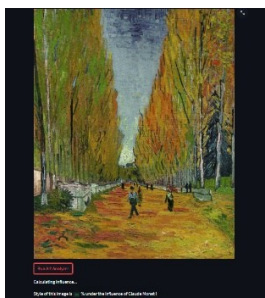
Tablica 5.6 Šesta skupina: Primjer pogrešnih klasifikacija modela



Točno: Titian

Predviđeno: Paul Cezanne

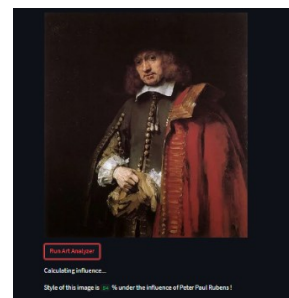
Izvor: <https://www.smithsonianmag.com/>



Točno: Van Gogh

Predviđeno: Claude Monet

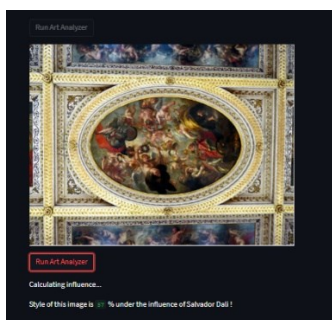
Izvor: <https://www.nytimes.com/2015/05/06/arts/design/van-gogh-painting-is-star-during-sothebys-auction.html>



Točno: Rembrandt

Predviđeno: Rubens

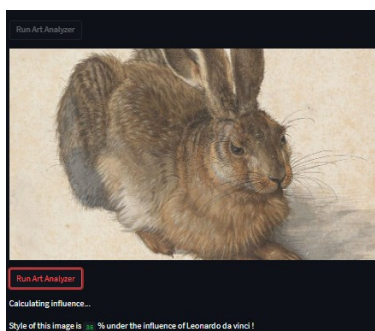
Izvor: <https://learnodo-newtonic.com/rembrandt-famous-paintings>



Točno: Rubens

Predviđeno: Dali

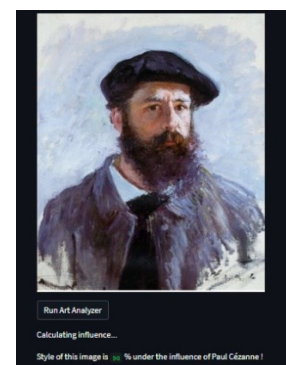
Izvor: <https://theculturetrip.com/europe/belgium/articles/10-artworks-by-rubens-you-should-know/>



Točno: Albrecht Dürer

Predviđeno: Leonardo da Vinci

Izvor: <https://www.bbc.com/culture/article/20190917-albrecht-drer-the-painter-with-a-magical-touch>



Točno: Monet

Predviđeno: Paul Cezanne

Izvor: <https://artsology.com/monet-van-gogh-still-life-paintings.php>

U posljednjoj tablici prikazano je nekoliko pogrešnih klasifikacija modela (Tablica 5.6 Tablica 5.6). Prema metrici točnosti, najbolje rezultate pokazuje model razvijen na temelju prilagođene ResNet50 arhitekture. Model je u stanju točno klasificirati 85,5% fotografija koje mu nisu poznate, međutim i dalje su preuzete sa internetske stranice *WikiArt*. Prilikom korištenja aplikacije, korisnik će nerijetko pokušati analizirati fotografije koje dolaze iz drugih izvora te sadrže veću količinu šuma ili distrakcija. To mogu biti različiti objekti koji se nalaze u okolini umjetničkog djela, bogati ukrasni okviri, varijacije u kontrastu ili intenzitetu boja, kao i u položaju umjetničkog djela na postavljenoj fotografiji. Jedan od primjera iz tablice pokazuje pogrešno klasificiranu Rubensovu sliku, u slučaju kad okvir zauzima uvjerljivu većinu postavljene fotografije. Ovaj tip testa također pokazuje da je model sklon pogreškama prilikom klasifikacije individualnog umjetničkog stila autora koji pripadaju istom ili srodnom umjetničkom pokretu. Na primjer, pogreške se češće događaju u sljedećim grupama:

- Van Gogh – Cezanne – Monet
(postimpresionizam – postimpresionizam – impresionizam)
- Da Vinci – Titian – Durer
(visoka renesansa- visoka renesansa – sjevernoeuropska renesansa)
- Rembrandt – Rubens
(barok – barok)

Smjernice za buduća istraživanja

6. S obzirom na rezultate provedenog istraživanja, vidljivo je da umjetna inteligencija može pomoći prilikom prepoznavanja stupnja sličnosti individualnog umjetničkog stila. Ovaj rad bavi se relativno novim područjem te postoje različite mogućnosti za daljnja istraživanja. Prethodno poglavlje pokazalo je više mogućih primjena testiranog modela, ali i nekoliko ograničenja. U narednim istraživanjima, preporuča se proširiti broj autora koji su uključeni u razvoj modela za prepoznavanje individualnog umjetničkog stila. Time bi se na jednostavan način povećala širina i primjenjivost rada. U praksi, sposobnost generalizacije modela najčešće se povećava treniranjem na većoj količini podataka. S obzirom da je broj umjetničkih djela po autoru ograničen, nova umjetnička djela moguće je kreirati korištenjem modela za prijenos stila ili generativnih suparničkih mreža. Ukoliko želimo širu primjenu modela koji bi bio u stanju prepoznati utjecaj individualnog umjetničkog stila na objekte koji nisu fotografije umjetničkih djela (majice, namještaj) potrebno je dodatno trenirati model na fotografijama koje sadrže što sličnije objekte željenima. Prisutnost prekomjernog prilagođavanja moguće je ublažiti primjenom kros-validacije te povećanim korištenjem tehnika regularizacije koje su detaljno opisane u poglavlju 2.3. Nekoliko istraživanja pokazuje bolje rezultate prilikom inicijalnog treniranja plitkih slojeva, dok istovremeno zamrzavaju dublje slojeve dobivene na *ImageNet* skupu podataka za detekciju objekata. S obzirom da je za izradu ovog rada korišten GPU procesor s ograničenim kapacitetima (besplatna verzija), isto nije bilo moguće primijeniti. Drugi mogući smjer za daljnja istraživanja je korištenje težina koje su fino podešene za prepoznavanje scene i predviđanje osjećaja, u odnosu na težine za detekciju objekata.

Iako je ovo istraživanje prikazalo individualni umjetnički stil kao matematičku varijablu, bilo bi korisno istražiti koje karakteristike umjetničkog djela ili konkretno, koje piksele na fotografiji, model smatra najbitnijima za detekciju individualnog umjetničkog stila. Navedeno bi omogućilo bolje razumijevanje modela i samim time bolje fino podešavanje arhitekture neuronske mreže. Također, to bi objektivno odgovorilo na pitanje što pojedinog autora čini jedinstvenim u odnosu na druge autore. Na primjer, zašto je Picasso Picasso?

Zaključak

Ovaj rad istražuje u kojoj mjeri algoritam za duboko učenje može prepoznati karakteristike i stupanj sličnosti individualnog umjetničkog stila. U teorijskom dijelu rada opisani su temeljni pojmovi i važnost umjetne inteligencije te su predstavljena četiri istraživačka smjera koja spajaju područje računalnih znanosti s vizualnom umjetnosti. Ovaj rad pripada u skupinu istraživanja koja se bave automatskom klasifikacijom slika. Također, opisana je tipična arhitektura konvolucijskih neuronskih mreža te su izdvojene one karakteristike mreže za koje je prepoznato da imaju najveći značaj u kontekstu detekcije individualnog umjetničkog stila.

U praktičnom dijelu rada prikazani su koraci potrebni za izradu web rješenja za prepoznavanje stupnja sličnosti individualnog umjetničkog stila. Kôd je napisan u programskom jeziku *Python*, dok su modeli vezani uz aplikacijski okvir *PyTorch*. Za potrebe ovog rada, odabrana su umjetnička djela deset najutjecajnijih slikara prema globalnoj rang listi El Zanta et al.: Leonardo da Vinci, Pablo Picasso, Vincent Van Gogh, Rembrandt Van Rijn, Peter Paul Rubens, Albrecht Durer, Titian, Claude Monet, Salvador Dali i Paul Cézanne (2018). Ukupno 8.159 fotografija umjetničkih djela preuzeto je s internetske stranice *WikiArt*, od čega je 1.400 uključeno u izradu modela za detekciju individualnog umjetničkog stila. Za svakog autora nasumično je odabrano 100 fotografija za trening modela, 20 za testiranje te 20 za validaciju. U stvarnom životu, umjetnici se jako razlikuju po broju stvorenih umjetničkih djela. Iako su neki hiperproduktivni, kao na primjer Van Gogh ili Monet, opus većine umjetnika čini stotinjak umjetničkih djela. Uzorak je sveden na 140 umjetničkih djela po autoru kako bi se rješenje moglo generalizirati na što više likovnih umjetnika. Kako bi rezultati bili što reprezentativniji na malenom uzorku, korišteno je transferno učenje. Težine dobivene na *ImageNet*-u pokazale su se relevantne za detekciju individualnog umjetničkog stila. Kao rezultat prvog istraživačkog cilja, razvijen je model temeljen na ResNet50 arhitekturi koji postiže točnost od 85,5% na testnom skupu podataka. Točnost je prilično stabilna, međutim prilikom treniranja vidljivo je da se funkcija gubitka ne smanjuje te još uvijek pokazuje prisutnost problema prekomjernog prilagođavanja. Ovaj fenomen bi se mogao ublažiti primjenom kros-validacije i dodatnih tehnika regularizacije, naročito tehnikama koje umjetno povećavaju skup podataka za učenje, što je navedeno i u smjernicama za buduća istraživanja. Prijašnja istraživanja pokazuju da je umjetnički stil izrazito teško definirati budući da se radi o kombinaciji boja i geometrijskih oblika (plitki

slojevi) te semantičkog sadržaja i motiva (duboki slojevi). Moguće je i da je umjetnički stil više pod utjecajem ugođaja i emocija u odnosu na detekciju objekata. Također, mreža na ulazu prihvaća kombinaciju crno-bijelih fotografija, kao i fotografija u boji što može dodatno utjecati na rezultate. Budući da su modeli trenirani na fotografijama umjetničkog djela, elementi fotografije kao što su osvjetljenje, udaljenost od umjetničkog djela u trenutku fotografiranja i kut slikanja također utječu na prepoznavanje – njihov utjecaj bi se mogao smanjiti odgovarajućim korištenjem metoda augmentacije. Unatoč uočenim nedostacima, dobiveni rezultat je zadovoljavajući za ovaj tip istraživanja te predstavlja gornju granicu točnosti i u prethodnim istraživanjima. Na temelju ovih rezultata vidljivo je da umjetna inteligencija može pomoći prilikom prepoznavanja individualnog umjetničkog stila te je navedeni koncept moguće promatrati kao matematički problem.

U sklopu drugog istraživačkog cilja, prilagođene su tri moderne arhitekture (ResNet, ResNeXt i DenseNet) koje su prethodno trenirane na tisuću ImageNet kategorija. Suprotno inicijalnoj pretpostavci, nije dobivena značajnija razlika u točnosti prilikom korištenja različitih arhitektura za duboko učenje. Iz navedenog se može zaključiti da tip korištene arhitekture konvolucijske neuronske mreže sam po sebi nije presudan prilikom detekcije individualnog umjetničkog stila. Zapravo, suprotno očekivanjima, najstabilnijom pokazala se je najstarija arhitektura ResNet50, iako je razlika u točnosti minimalna, odnosno iznosi 0,5%. S obzirom da pokazuje najveću točnost, model temeljen na spomenutoj arhitekturi odabran je za razvoj web rješenja za detekciju stupnja sličnosti individualnog umjetničkog stila. Minimalna razlika u točnosti nije toliko iznenađujuća s obzirom da se radi o tri renomirane arhitekture koje su primijenjene na malenom uzorku. Značajnije razlike u točnosti možda bi se pojavile prilikom korištenja većeg uzorka.

Prema trećem cilju rada, dizajniran je prototip aplikacije koja korisniku omogućava dodavanje vlastite fotografije u *.jpg* ili *.png* formatu. Nakon što model analizira dodanu fotografiju, izračunava stupanj sličnosti te taj podatak korisniku prezentira u obliku postotka. Izmjereni postotak predstavlja najveći identificirani stupanj sličnosti dodane fotografije u odnosu na individualni umjetnički stil autora iz uzorka. Važno je primijetiti da će program uvijek ponuditi jednu, najveću vrijednost stupnja sličnosti. Problem prepoznavanja umjetničkog stila promatran je iz perspektive uravnoteženih klasa budući da nova fotografija koju korisnik postavi ima jednaku vjerojatnost pripadanja svakoj kategoriji. Također, utjecaj je izražen u postotku i prikazan kroz metriku točnosti, koja nije uspješna u slučaju neuravnoteženih kategorija.

Konačni model pokazuje zadovoljavajući stupanj točnosti uz ograničenja, koja su očekivana za modele koji su temeljni na umjetnoj inteligenciji. Može se zaključiti da je individualni umjetnički stil moguće promatrati kao kombinaciju matematičkih vrijednosti, što omogućuje objektivnu analizu međusobnog utjecaja autora. Također, testiranjem modela potvrđeno je da se primjenom dubokog učenja može detektirati stupanj sličnosti individualnog umjetničkog stila različitih autora.

Model ovog tipa može imati više primjena u stvarnom životu. Na primjer, model omogućuje korisnicima da analiziraju umjetnička djela nekadašnjih, današnjih i budućih umjetnika te da istraže pod čijim utjecajem djeluju i stvaraju, bilo svjesno ili nesvjesno. Ovo bi moglo biti posebno korisno studentima i nastavnicima. Model omogućuje i širu primjenu: moguće je analizirati utjecaj autora i na druge oblike kreativnog stvaralaštva kao što su filmska umjetnost, promotivni materijali, dizajn interijera ili moda. U prilog tome, model je uspio prepoznati da je scena iz filma inspirirana jednim autorom, a komad namještaja drugim autorom. Daljnji razvoj modela temeljenog na prepoznavanju individualnog umjetničkog stila, u kombinaciji s drugim AI alatima omogućio bi dizajn prostora, namještaja ili odjeće u stilu željenog autora.

Popis kratica

AI	<i>Artificial Intelligence</i>	umjetna inteligencija
API	<i>Application Programming Interface</i>	aplikacijsko programsko sučelje
CNN	<i>Convolutional Neural Network</i>	konvolucijska neuronska mreža
DenseNet	<i>Dense Convolutional Network</i>	gusta konvolucijska mreža
HOG	<i>Histograms of oriented gradients</i>	histogrami usmjerenih gradijenata
k-NN	<i>k-nearest neighbours</i>	lijeni klasifikatori
R-CNN	Region-based Convolutional Neural Network	konvolucijske neuronske mreže bazirane na regijama
ReLU	<i>Rectified linear unit</i>	ReLU prijenosna funkcija
ResNet	<i>Residual Neural Network</i>	rezidualna neuronska mreža
ResNeXt	<i>Next Generation Residual Neural Network</i>	rezidualna neuronska mreža nove generacije
SIFT	<i>Scale-invariant feature tranforms</i>	transformacije invarijantnih značajki mjerila
SVM	<i>Support Vector Machines</i>	metoda potpornih vektora
YOLO	You Only Look Once	„YOLO“ model

Popis slika

Slika 1.1 Usporedba klasifikacije i detekcije objekta.....	4
Slika 2.1 Tipična arhitektura konvolucijske neuronske mreže (Goyal et al., 2021).....	8
Slika 2.2 Model s primjenom <i>dropout</i> sloja (Srivastava et al., 2014).....	12
Slika 3.1 Primjer korištenog http zahtjeva za prikupljanje podataka	14
Slika 3.2 Primjer nekoliko fotografija umjetničkih djela preuzetih s <i>WikiArt-a</i>	15
Slika 3.3 Primjer nekoliko fotografija koje su isključene iz analize	16
Slika 3.4 Struktura pripremljenog skupa podataka.....	18
Slika 3.5 Arhitekture dubokih rezidualnih mreža predložene za ImageNet (He et al., 2016)	20
Slika 3.6 Prikaz točnosti i funkcije gubitka za model ResNet50.	22
Slika 4.1 Usporedba ResNet-50 modela s ResNeXt50_32x4d arhitekturom (Xie et al., 2017).	23
Slika 4.2 Prikaz točnosti i funkcije gubitka za model resnext50_32x4d.....	24
Slika 4.3 Prikaz slojeva u DenseNet mreži (Huang et al., 2017)	25
Slika 4.4 Predložene DenseNet arhitekture za ImageNet (Huang et al., 2017).....	25
Slika 4.5 Prikaz točnosti i funkcije gubitka za model densenet121.	26
Slika 5.1 Sučelje aplikacije.....	29
Slika 5.2 Primjer rezultata klasifikacije pomoću aplikacije <i>Art Analyzer</i>	29
Slika 5.3 Izgled koda za autentifikaciju.	30
Slika 5.4 Izgled veze za pokretanje aplikacije.....	33

Popis tablica

Tablica 2.1 Primjer transformacija slika koje su dostupne u PyTorch biblioteci (PyTorch, 2022).....	11
Tablica 3.1 Pregled dostupnog istraživačkog uzorka	17
Tablica 4.1 Usporedba rezultata treniranja različitih modela.....	27
Tablica 5.1 Prva skupina: umjetnička djela autora iz uzorka, iz drugih izvora.....	36
Tablica 5.2 Druga skupina: isti motivi u različitim umjetničkim stilovima	37
Tablica 5.3 Treća skupina: Umjetnička djela stvorena pod utjecajem autora iz uzorka	38
Tablica 5.4 Četvrta skupina: Fotografije koje ne prikazuju umjetnička djela, ali je vidljiv stil	39
Tablica 5.5 Peta skupina: Nasumično testiranje objekata	40
Tablica 5.6 Šesta skupina: Primjer pogrešnih klasifikacija modela	41

Popis kodova

Kôd 3.1 Potrebne <i>Python</i> biblioteke za reprodukciju transfernog učenja	13
Kôd 3.2 <i>Python</i> kôd za preuzimanje slika u .jpg i .png formatu	15
Kôd 3.3 Podjela podataka na skup za učenje, testiranje i validaciju.	18
Kôd 3.4 Kôd koji prikazuje ulazne parametre za sva tri modela.....	19
Kôd 3.5 Pozivanje Resnet50 modela.....	21
Kôd 4.1 Pozivanje ResNeXt50_32x4d modela.	24
Kôd 4.2 Pozivanje <i>DenseNet</i> modela.....	26
Kôd 5.1 Instalacija biblioteka.....	30
Kôd 5.2 Naredbe za autentifikaciju	30
Kôd 5.3 Kôd u pozadini prototipa web aplikacije.....	32

Literatura

- [1] Adedokun, G. (2019). Deep Learning - An Overview. *Journal of Current Engineering and Technology*, 1(1), Article ISSN: 2582-1210.
- [2] *Art movement*. Wikipedia. https://en.wikipedia.org/wiki/Art_movement (Pristupljeno: 14.2.2023.).
- [3] *Art Styles Explained - Explore Popular Styles of Art! — Art is Fun. (2023)*. Art Is Fun. <https://www.art-is-fun.com/art-styles> (Pristupljeno: 14.2.2023.).
- [4] Castellano, G., & Vessio, G. (2021a, veljača). A Brief Overview of Deep Learning Approaches to Pattern Extraction and Recognition in Paintings and Drawings. *Lecture Notes in Computer Science Book Series, LNIP, volume 12663*. https://doi.org/10.1007/978-3-030-68796-0_35 (Pristupljeno: 14.2.2023.).
- [5] Castellano, G., & Vessio, G. (2021b, listopad). Deep learning approaches to pattern extraction and recognition in paintings and drawings: an overview. *Neural Computing and Applications*, 33(19), 12263–12282. <https://doi.org/10.1007/s00521-021-05893-z> (Pristupljeno: 14.2.2023.).
- [6] Cetinic, E., Lipic, T., & Grgic, S. (2018). Fine-tuning Convolutional Neural Networks for fine art classification. *Expert Systems With Applications*, 114, 107–118. <https://doi.org/10.1016/j.eswa.2018.07.026> (Pristupljeno: 14.2.2023.).
- [7] Chollet, F. (2021). *Deep Learning with Python, Second Edition* (2nd ed.). Manning.
- [8] *CS231n: Deep Learning for Computer Vision*. (2022). Stanford University. <https://cs231n.stanford.edu/> (Pristupljeno: 14.2.2023.).
- [9] Dalbelo Bašić, B., & Šnajder, J. (n.d.). *Vrednovanje klasifikatora* [Slide show]. <https://www.fer.unizg.hr/>. https://www.fer.unizg.hr/_download/repository/SU-12-VrednovanjeKlasifikatora.pdf (Pristupljeno: 14.2.2023.).
- [10] El Zant, S., Jaffres-Runser, K., Frahm, K. M., & Shepelyansky, D. L. (2018). Interactions and Influence of World Painters From the Reduced Google Matrix of Wikipedia Networks. *IEEE Access*, 6, 47735–47750. <https://doi.org/10.1109/access.2018.2867327> (Pristupljeno: 14.2.2023.).
- [11] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning (Adaptive Computation and Machine Learning series)* (Illustrated). The MIT Press.
- [12] Goyal, V., Maheshwari, A., & Kathuria, M. (2021). Image Recognition using Deep Learning Techniques. *Journal of Xidian University*, 15(6). <https://doi.org/10.37896/jxu15.6/005> (Pristupljeno: 14.2.2023.).
- [13] Haldar, S. (2019, March 2). *DeepArtist : Identify Artist from Art - Best Artworks of All Time*. Kaggle. <https://www.kaggle.com/datasets/ikarus777/best-artworks-of-all-time> (Pristupljeno: 14.2.2023.).
- [14] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. <https://doi.org/10.1109/CVPR.2016.90> (Pristupljeno: 14.2.2023.).
- [15] Hong, Y., & Kim, J. (2017). Art Painting Identification using Convolutional Neural Network. *International Journal of Applied Engineering Research*, 12(4), 532–539.

- [16] Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017, July). Densely Connected Convolutional Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.48550/arXiv.1608.06993> (Pristupljeno: 14.2.2023.).
- [17] *Introduction to PyTorch | Deep Learning | Udacity Free Courses*. (2022). Udacity. <https://www.udacity.com/course/deep-learning-pytorch--ud188> (Pristupljeno: 14.2.2023.).
- [18] Li, G., Wang, X., & Li, M. (Eds.). (2021). *A Review of Recent Trends and Industry Prospects for Artificial Intelligence Technologies Publisher: IEEE Institute of Electrical and Electronics Engineers* (2021 8th International Conference on Behavioral and Social Computing (BESC)).
- [19] Mikolajczyk, A., & Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. *2018 International Interdisciplinary PhD Workshop (IIPhDW)*. <https://doi.org/10.1109/iiphdw.2018.8388338> (Pristupljeno: 14.2.2023.).
- [20] Mohammad, S. M., & Kiritchenko, S. (2018, May). WikiArt Emotions: An Annotated Dataset of Emotions Evoked by Art. *LREC 1*, <https://aclanthology.org/L18-1197.pdf> (Pristupljeno: 14.2.2023.).
- [21] Mondal, K., & Anita, H. (2021). Categorization of artwork images based on painters using CNN. *Journal of Physics: Conference Series, 1818(1)*, 012223. <https://doi.org/10.1088/1742-6596/1818/1/012223> (Pristupljeno: 14.2.2023.).
- [22] Nneji, I. (2022, June 3). How to create and launch a Streamlit app directly from Google Colab. *Medium*. Retrieved September 16, 2022, from <https://faun.pub/how-to-launch-a-streamlit-app-directly-from-google-colab-8014165e0267> (Pristupljeno: 14.2.2023.).
- [23] *PyTorch*. (2022). https://pytorch.org/hub/pytorch_vision_resnet/ (Pristupljeno: 14.2.2023.).
- [24] Qu, K., Xiao, D., & Zhu, H. (n.d.). DeepArt: Identify Artist from Painting. *Department of Computer Science, University of Virginia, Charlottesville, VA 22904*.
- [25] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research, 15*.
- [26] Stevens, E., Antiga, L., & Viehmann, T. (2020). *Deep Learning with PyTorch: Build, train, and tune neural networks using Python tools* (1st ed.). Manning.
- [27] *Streamlit Docs*. (2022). Streamlit Inc. <https://docs.streamlit.io/> (Pristupljeno: 14.2.2023.).
- [28] Viswanathan, N. (2017). *Artist Identification with Convolutional Neural Networks* (No. 406). Stanford University. <http://cs231n.stanford.edu/reports/2017/pdfs/406.pdf> (Pristupljeno: 14.2.2023.).
- [29] *Wikiart Api - WikiArt.org*. (n.d.). www.wikiart.org, <https://www.wikiart.org/en/App/GetApi> (Pristupljeno: 14.2.2023.).
- [30] *WikiArt.org - Visual Art Encyclopedia*. (n.d.). www.wikiart.org. Retrieved September 13, 2022, <https://www.wikiart.org/> (Pristupljeno: 14.2.2023.).

- [31] Wikipedia contributors. (2013). *Likovni motivi.*, https://hr.wikipedia.org/wiki/Likovni_motivi (Pristupljeno: 14.2.2023.).
- [32] Wikipedia contributors. (2023). *Style (visual arts).* Wikipedia. [https://en.wikipedia.org/wiki/Style_\(visual_arts\)](https://en.wikipedia.org/wiki/Style_(visual_arts)) (Pristupljeno: 14.2.2023.).
- [33] Wikipedia contributors. (2022). *Visual Studio Code.* Wikipedia. https://en.wikipedia.org/wiki/Visual_Studio_Code (Pristupljeno: 14.2.2023.).
- [34] Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017). Aggregated Residual Transformations for Deep Neural Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Authors:* <https://doi.org/10.1109/CVPR.2017.634> (Pristupljeno: 14.2.2023.).
- [35] Zhao, Z., Zheng, P., Xu, S., & Wu, X. (2019). Object Detection With Deep Learning: A Review. *IEEE Trans Neural Netw Learn Syst.*, 30(11). <https://doi.org/10.1109/TNNLS.2018.2876865> (Pristupljeno: 14.2.2023.).