

Sistem za interaktivno pregledovanje slikarske zbirke

Nik Jenič, Nace Omahen, Maj Udovč, Maša Uhan, Blaž Meden, Narvika Bovcon

Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Večna pot 113, 1000 Ljubljana

E-pošta: nj11108@student.uni-lj.si, narvika.bovcon@fri.uni-lj.si

System for Interactive Exploration of a Painting Collection

Abstract. *This project presents a prototype system for interactive exploration of a painting collection using computer vision techniques. The application enables users to analyze artworks based on visual attributes, such as detected facial emotions, dominant objects, body poses, and color distribution. Using pre-trained models and Python scripts, the system processes each image locally to extract features that serve as filters for visual search.*

Five main modes of interaction were developed: filtering paintings by facial emotions (eg., sadness, happiness), navigating object size (eg., larger or smaller vases), exploring pose-based clusters, sorting images via HSL color space, and examining line structures through Hough transform visualization. The user interface was prototyped with a web-based layout and supports interactive elements like sliders and keyboard navigation. Backend analysis runs locally in Python using libraries such as OpenCV, MediaPipe, and DeepFace, while the web interface serves as a visual tool for displaying precalculated results.

Although not yet public-facing, the system demonstrates the potential for using artificial vision to uncover visual patterns in art and offer novel, engaging ways to explore cultural collections while learning about the workings of each technique. The project lays a solid foundation for further development into an educational and publicly accessible tool.

1 Uvod

Razvoj računalniškega vida in umetne inteligence je omogočil številne nove pristope za analizo in razumevanje umetniških del. Tradicionalno je bila interpretacija slikarskih del prepuščena človeškemu očesu in subjektivnemu zaznavanju, kar pogosto omejuje celovito razumevanje vizualnih elementov, posebej na ravni velikih zbirk. Ta projekt predstavlja interaktivno spletno aplikacijo, ki s pomočjo naprednih metod računalniškega vida omogoča raziskovanje velike zbirke slikarskih del. Uporabili smo zbirko s spletne strani WikiArt, ki vsebuje več deset tisoč del. Aplikacija združuje algoritme za obdelavo slik, kot so barvni histogrami, zaznavanje robov in Houghova transformacija, ter globoko učenje za prepoznavanje obrazov, poz in objektov. Uporabnikom omogoča, da s pomočjo intuitivnih interaktivnih orodij raziskujejo umetnine glede na različne vizualne lastnosti, kar zagotavlja edinstveno in poučno uporabniško izkušnjo.

Skozi uporabo te aplikacije želimo širši javnosti približati delovanje umetne inteligence, predstaviti možnosti računalniškega vida v umetnosti ter spodbuditi globlje razumevanje in cenjenje umetniške kreativnosti. Poleg tega smo želeli, da vsak uporabnik sam določi svojo unikatno pot raziskovanja del, zato smo se odločili za interaktivne načine delovanja. Ker nismo mogli enolično določiti slike za prikaz, smo te izbrali naključno. Tako lahko tudi dva podobna si uporabnika doživita drugo podmnožico zbirke slik.

2 Pregled področja

Računalniški vid je področje umetne inteligence, ki omogoča računalnikom zaznavanje in interpretacijo vizualnih informacij iz digitalnih slik. V kontekstu umetnosti se to uporablja za analizo barvnih kompozicij, detekcijo oblik in prepoznavanje motivov. Tehnike, kot so Cannyjevo zaznavanje robov in Houghova transformacija, omogočajo identifikacijo struktur znotraj slik, medtem ko se globoko učenje uporablja za razpoznavo obrazov, emocij, objektov in človeške drže.

Na spletu smo našli nekaj projektov, ki obravnavajo sorodne koncepte in uporabo umetne inteligence v vizualnem okolju. Eden izmed njih je projekt **Teachable Machine** [1], ki omogoča uporabnikom, da brez programiranja ustvarijo modele strojnega učenja za prepoznavo slik, zvokov ali gibanja. Njegov namen je izobraževalni, z močno poudarjeno enostavnostjo uporabe.

Drug primer je projekt **ARTETIK: From the Art** [2], ki omogoča raziskovanje umetniških del na podlagi barvne palete. Uporabnik lahko izbira barve in prejme prikaz umetnin, kjer te barve prevladujejo, kar spodbuja vizualno razumevanje umetniških slogov.

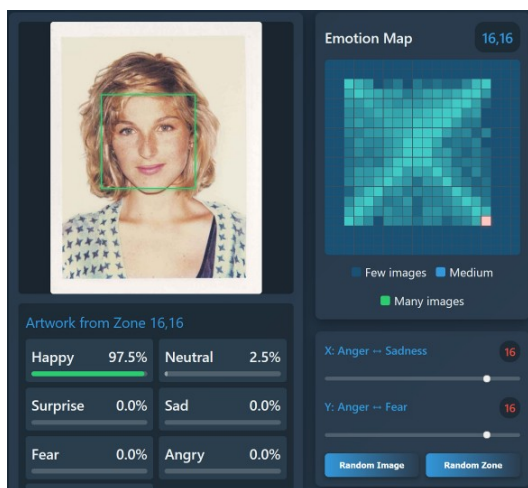
Čeprav omenjena orodja obravnavajo uporabo umetne inteligence in računalniškega vida za dostopno vizualno raziskovanje, menimo, da ne dosega tiste ravni **interaktivnosti, intuitivnosti in poučnosti**, ki jo želimo ponuditi v našem projektu. Naš cilj je uporabniku omogočiti več različnih načinov raziskovanja umetnosti, kjer lahko z lastnim vnosom, navigacijo ali interakcijo neposredno vpliva na potek vizualne analize in pridobi poglobljen vpogled v strukturo umetniškega dela.

3 Metodologija

Za zaznavanje vizualnih značilnosti slik smo uporabili več različnih metod in knjižnic s področja računalniškega vida. Pri zaznavanju obrazov in prepoznavanju emocij smo uporabili predtrenirane modele knjižnice DeepFace. Za zaznavanje telesnih poz

smo implementirali metodo za analizo skeletne strukture posameznika, ki smo jo pridobili s pomočjo knjižnice MediaPipe, kar nam je omogočilo razvrščanje slik glede na podobnost telesnih položajev [4]. Prepoznavanje objektov (npr. vaze, stoli) smo izvedli z detekcijskim modelom YoLo [3], ki omogoča iskanje in primerjavo velikosti izbranih objektov med različnimi slikami. Barvno analizo smo izvedli z uporabo HSL (Hue, Saturation, Lightness) barvnega prostora in barvnih histogramov, ki omogočajo interaktivno primerjavo slik glede na svetlost, nasičenost in barvni ton. Za analizo geometrijskih vzorcev smo uporabili Houghovo transformacijo, s katero lahko zaznamo prisotnost linijskih struktur v slikah. Vse metode so bile izvedene v Pythonu in kasneje v HTML-ju in JavaScript-u s pomočjo knjižnic, kot so OpenCV, NumPy in Torch.

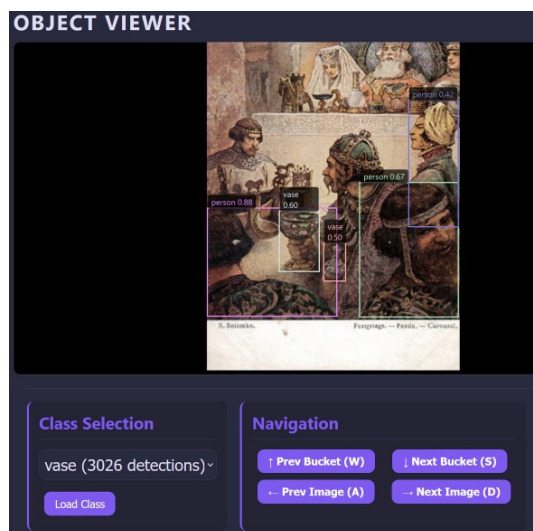
Pregled slik glede na emocije na obrazih: Izbrani model na slikah zazna obraze in čustva, ki jih posamezen obraz izraža. Te emocije so veselje, žalost, strah, presenečenje, jeza, gnus in tudi nevtralnost. Izmed njih smo izbrali štiri (veselje, žalost, jezo in strah), da jih lahko pokažemo v matrični obliki, kjer vsako polje (ang. zone) shrani podzbirko slik. V matriki sredina predstavlja nevtralnost, kot pa večjo prepričanost modela v zaznavo specifičnega čustva (npr. zgoraj levo je jeza, zgoraj desno žalost, itd.). Vmesne vrednosti v matriki prikazujejo mešano prepričanost v sosedna tri čustva (vključno z nevtralnostjo). Ob kliku na polje ali na gumb »Random Image« se iz podzbirke izbere in prikaže naključna slika s temi lastnostmi, ki so pa tudi bolj natančno opisane pod sliko. Matrika deluje na osnovi koordinat od 0,0 do 16,16. Te se določijo ob kliku na polje ali preko dveh drsnikov. Intenziteta barve polja predstavlja število slik v podzbirki.



Slika 1: Sortiranje zbirke slik po emocijah

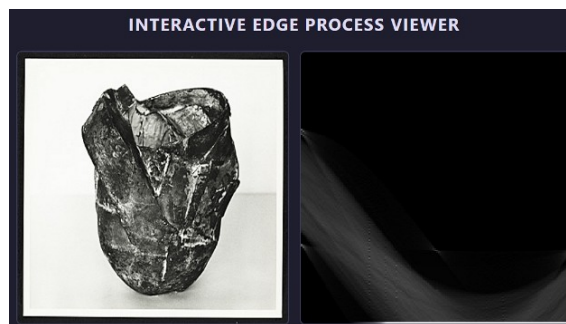
Sortiranje slik glede na prepoznane objekte: Uporabnik izbere zeleni predmet, sistem pa mu prikaže slike, kjer je ta predmet najbolj izrazit. Nato lahko

uporabnik s tipkami na tipkovnici (smerne puščice) prehaja med podobnimi slikami – navzgor za večji primer izbranega predmeta, navzdol za manjšega. Ta način omogoča natančno raziskovanje, kako je določen predmet predstavljen skozi različna umetniška dela.



Slika 2: Zaznava objektov v slikah

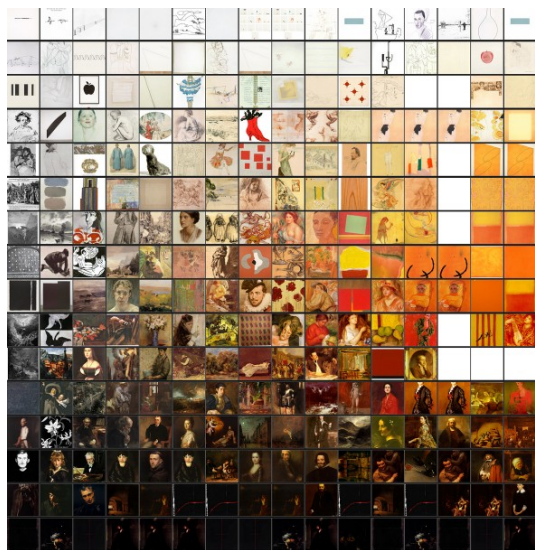
Interaktivna uporaba Houghove transformacije: Houghova transformacija je metoda, ki zaznava geometrijske oblike (npr. črte in kroge) na slikah. Ta način ima izobraževalni namen in uporabniku omogoča, da se v levem oknu s pomočjo drsnika premika skozi korake prirejanja slike, da smo iz nje izluščili rezultat. V desnem oknu je prikazana ustrezna sinusoida, ki matematično opisuje zaznane črte v sliki. Vsaka sinusoida predstavlja en piksel zaznanih robov v sliki. Kjer se sinusoide najbolj prekrivajo, sprejmemo kot zaznavo ravne črte. Uporabnik lahko izbere naključno sliko za obdelavo ali pa sam nariše svojo sinusoido, sistem nato poišče sliko iz zbirke, katere Houghov vzorec se najbolj ujema z narisanim.



Slika 3: Prikaz sinusoid pri Houghovi transformaciji

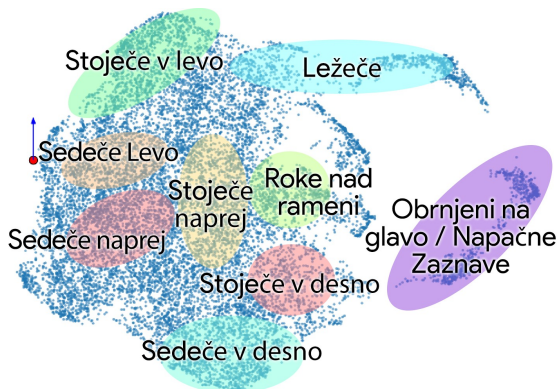
Raziskovanje slik po barvnih značilnostih (HSL analiza): Vmesnik prikazuje mrežo slik 16x16: navpična os predstavlja svetlost (od temne proti svetli),

vodoravna pa nasičenost barv (od manj nasičenih do zelo intenzivnih). Z dodatnim drsnikom lahko uporabnik filtrira slike tudi po izbranem barvnem tonu v razponu od 0° do 360°, kar pokriva celoten barvni spekter. Barvo, ki jo slika predstavlja, smo dobili preko povprečenja barv vseh pikslov na sliki. Ob kliku na posamezno sliko se prikaže histogram njenih barv, uporabnik pa lahko s pomočjo orodja doda nove barve na histogram ali pa poveča intenziteto že prisotnih barv. Nato pa mu sistem poišče slike, ki imajo najbolj podobno barvno sestavo.



Slika 4: Prikaz slik v HSL formatu

Analiza telesnih poz oseb na slikah: Slike so razvrščene na dvodimenzionalni mapi glede na podobnost prikazane poze v primerjavi s pozami na ostalih slikah. Uporabnik se po tej mapi premika s tipkovnico, v levem oknu vidi položaj trenutne slike znotraj celotne skupine (ang. cluster), v desnem pa sliko samo. Tako lahko hitro najde slike, kjer so na primer ljudje z dvignjenimi rokami ali v gibanju, in primerja različne umetniške interpretacije človeškega telesa, saj so si slike s podobnimi pozami na zemljevidu bližje.



Slika 5: Gruče telesnih poz



Slika 6: Več del z zaznanimi pozami

Predstavljeni načini pregledovanja slik tvorijo vsestranski vmesnik, ki združuje umetnost in tehnologijo ter omogoča uporabnikom, da slike ne le gledajo, temveč jih tudi razumejo z vidika računalniškega vida, kot ga spoznajo skozi uporabo orodij ter z raziskovanjem pripetih razlag o delovanju vsake uporabljene tehnike. Pomembno je, da je celotna izkušnja zasnovana tako, da je dostopna tudi popolnim začetnikom brez predhodnega znanja o umetni inteligenci, ter da vsak uporabnik pridobi unikatno izkušnjo, določeno preko njegovih odločitev pri uporabi sistema. Vmesnik je intuitiven, rezultati pa so predstavljeni na vizualno razumljiv in informativen način.

4 Prototipiranje aplikacije

V prvi fazi projekta smo razvili funkcionalen prototip, ki je temeljil na izvedbi vseh obdelav in detekcij znotraj okolja Python. V tej fazi še ni bilo vzpostavljene povezave s spletno aplikacijo, zato so se vsi rezultati prikazovali lokalno na računalniku z uporabo matplotlib grafov in prikazov. Vsaka slika v zbirki je bila obdelana posebej, rezultati pa so vključevali vse prej omenjene načine zaznav. Slike smo obdelovali s pomočjo Python knjižnic, saj so te najbolj razvite na področju in omogočajo najhitrejšo prototipiranje. Poleg tega smo podatke računali in hranili vnaprej, saj bi bilo računanje teh sproti prepočasno in prezahtevno za računalnike posameznih uporabnikov.

Prototip je omogočal tudi razvrščanje slik na podlagi opravljenih analiz. Uporabnik je lahko zagnal posamezne skripte za vsak način zaznave. Kljub temu da so bile funkcionalnosti ločeno implementirane in so uspešno delovale, še niso bile vključene v enotno uporabniško izkušnjo v brskalniku.

Poleg analitičnih funkcij smo pripravili tudi zasnovo spletne aplikacije. Ob vstopu v aplikacijo uporabnik naleti na interaktivno začetno stran z gumbom v obliki očal "računalniškega vida". Klik nanj sproži prikaz možnosti za filtriranje slik po različnih kriterijih, kot so barva, zaznani objekti, poze ali obrazi. Čeprav spletni vmesnik v tej fazi še ni bil interaktiven in ni bil povezan s podatkovnim ozadjem, je služil kot konceptualni prikaz uporabniške izkušnje in je tako predstavljal pomemben korak v razvoju projekta, saj smo z njim potrdili tehnično izvedljivost posameznih komponent, obenem pa opredelili strukturo končne aplikacije.

S preverjenim delovanjem posameznih delov aplikacije smo podatke pretvorili v obliko primerno za spletne aplikacije, ter vsa orodja za prikazovanje le teh pretvorili v spletno obliko in jih povezali z začetno stranjo.

5 Diskusija

Pri razvoju prototipa smo želeli preveriti, kako uspešno lahko računalniški vid in metode globokega učenja pripomorejo k analizi umetniških del na uporabniku prijazen in vizualno razumljiv način, ter kako se lahko preko raziskovanja novih del uporabnik nauči več o svetu računalniškega vida.

Analiza emocij se je v večini primerov izkazala za uspešno, vendar je bila natančnost odvisna od kvalitete posamezne slike, še posebej, če je obraz na sliki majhen, in izraznosti obrazov na slikah. Umetniška dela, ki niso realistična ali ne prikazujejo obrazov jasno, so predstavljala izziv za modele. Podobno so bili pri zaznavanju poz nekateri rezultati netočni pri abstraktnjših slikah, kjer človeška figura ni eksplicitna, je prikazana nenavadno ali je delno prekrita. Rezultate bi lahko nadgradili s povezovanjem več različnih modelov zaznav in preko tega z množičnim preverjanjem veljavnosti vsake zaznave.

Barvna analiza s pomočjo HSL prostora je bila ena izmed najbolj stabilnih in uporabnih funkcij. Omogočila je sistematičen pregled slik po svetlosti, nasičenosti in barvnem tonu. Uporabniki so lahko hitro opazili slogovne podobnosti med deli, ki jih sicer vizualno morda ne bi povezali. Dodatna funkcionalnost interakcije s histogrami barv je pokazala, da je možno barvno analizo še nadgraditi z vnosom lastnih kriterijev. Sistem bi lahko dalje nadgradili s preurejanjem in razlago prikazanih vrednosti v histogramih ter z drugačno klasifikacijo teh znotraj HSL prostora (npr. z uporabo dominantne barve in ne povprečne).

Sortiranje glede na objekte in njihovo velikost je ponudilo zanimiv način raziskovanja ikonografskih elementov v umetnosti. V praksi pa smo opazili, da zaznavanje ni bilo vedno dosledno – na primer predmet, ki ga je model zaznal kot vazo, je včasih bil del ozadja ali simbolična upodobitev, kar odpira vprašanje konteksta in pomena v umetnosti, ki ga trenutni modeli še ne razumejo. Tu bi ponovno lahko uporabili skupek več različnih modelov za boljše rezultate ter dodatno prilagodili parametre modela.

Houghova transformacija in prikaz sinusoid sta predstavljala bolj tehničen in izobraževalen vidik aplikacije. Čeprav je ta način bolj primeren za uporabnike, ki jih zanimajo matematični temelji računalniškega vida, smo ugotovili, da je lahko tudi ta del privlačen – še posebej z možnostjo interaktivnega risanja sinusoid in iskanja podobnih slik. Zaradi načina delovanja Houghove transformacije je potrebno prilagajanje mnogih parametrov pri izračunih, zato je naslednji korak preverjanje veljavnosti vsakega posebej.

Pomembna omejitev projekta je bila ta, da vse funkcionalnosti še niso bile optimizirane za najboljše možne rezultate, ter da podatki niso bili ustvarjeni, hranjeni in prikazani na optimiziran način za spletni prikaz, zaradi česar se rezultati prikazujejo in posodablajo relativno počasi.

Poleg tehničnega razvoja smo osnovno uporabniško izkušnjo testirali znotraj projektne skupine. Zaradi prototipne narave sistema je bilo testiranje izvedeno interno med razvijalci posameznih delov sistema, pri čemer smo beležili, kje je bila potrebna dodatna razlaga in kje so se pojavile težave. Na ta način smo prepoznali in odpravili glavne pomanjkljivosti, kot so počasno nalaganje, neintuitivne interakcije in nedodelana vizualizacija. V prihodnosti načrtujemo širše testiranje z dejanskimi uporabniki, uporabo analitičnih orodij za beleženje interakcij ter vključitev povratnih informacij za izboljšanje uporabniške izkušnje.

6 Zaključek

Razvili smo prototip sistema z računalniškim vidom in umetno inteligenco, ki nudi nove načine interaktivnega raziskovanja in analize slikarskih del. Te metode omogočijo ne le vizualno, temveč tudi analitično komunikacijo z umetniškimi deli.

Prototip je potrdil tehnično izvedljivost posameznih komponent, saj smo s pomočjo javnih orodij uspešno zaznali ključne vizualne značilnosti slik. Čeprav sistem še ni v celoti povezan v delujočo spletno aplikacijo, predstavlja dobro izhodišče za nadaljnji razvoj.

Projekt dokazuje, da lahko umetna inteligenca in računalniški vid prispevata k širšemu razumevanju umetnosti ter omogočita nove oblike izobraževanja in estetske izkušnje, ter odpre vprašanje, ali so ta orodja dovolj uporabljena znotraj umetniških krogov. V prihodnje načrtujemo pospeševanje povezave podatkov z vmesnikom, izboljšanje rezultatov s prilagajanjem in kombinacijo različnih modelov, izboljšanje uporabniške izkušnje in razlago posameznih tehnik. Naš cilj je s pomočjo tehnologije približati umetnost širši javnosti na dostopen, poučen, zabaven in vizualno privlačen način.

Literatura

- [1] Google Creative Lab, *Teachable Machine*. [Online]. Available: <https://teachablemachine.withgoogle.com/> (Accessed: 1 June 2025)
- [2] Google Arts & Culture, *ARTETIK: From the Art*. [Online]. Available: https://artsandculture.google.com/color?project=guggenheim-bilbao&col=RGB_3F9CEF (Accessed: 1 June 2025)
- [3] R. Kundu, "YOLO Algorithm for Object Detection Explained [+Examples]," V7 Labs, 17. januar 2023. [Online]. Dostopno: <https://www.v7labs.com/blog/yolo-object-detection>. [Dostopano: 1. junij 2025].
- [4] Google AI, "MediaPipe Solutions guide," Google AI Edge, marec 2025. Dostopno: <https://ai.google.dev/edge/mediapipe/solutions/guide>. [Dostopano: 1. junij 2025].