|  |  |
| --- | --- |
|  | Zbornik radova Fakulteta tehničkih nauka, Novi Sad |

**UDK:**

**DOI:**

**SISTEM ZA REVERZNU PRETRAGU SLIKA**

**REVERSE IMAGE SEARCH SYSTEM**

Ivan Mršulja, *Fakultet tehničkih nauka, Novi Sad*

**Oblast – PRIMENJENE RAČUNARSKE NAUKE I INFORMATIKA**

**Kratak sadržaj –** *Ovaj rad predlaže alternativni pristup reverznoj pretrazi slika korišćenjem modela mašinskog učenja za izvlačenje tagova i svojstava iz slike. Svođenjem vektorske pretrage na pretragu teksta, predloženi metod povećava efikasnost i tačnost pretrage slika.*

**Ključne reči:** *Reverse pretraga slika, mašinsko učenje i računarska vizija su ključni koncepti u ovom istraživanju.*

**Abstract** – *This paper proposes an alternative approach to reverse image search by leveraging machine learning models to extract tags and features from an image. By reducing vector image search to a full-text search, the proposed method improves the efficiency and accuracy of image retrieval..*

**Keywords:** *Reverse image search, machine learning, computer vision.*

**1. UVOD**

Reverzna pretraga slika (reverse image search) je tehnika pretrage koja omogućava korisnicima da pronađu slične ili identične slike na internetu koristeći već postojeću sliku kao ulazni parametar [1]. Umesto da korisnici unose tekstualne upite, mogu otpremiti sliku ili uneti URL slike kako bi pronašli druge slike sa sličnim vizuelnim karakteristikama [2]. Algoritmi za reverznu pretragu slika analiziraju ključne elemente slike, kao što su boje, oblici, teksture ili drugi vizuelni detalji, kako bi pronašli slike sa sličnim karakteristikama iz baza podataka ili veb stranica širom interneta. Ova tehnika je korisna u mnogim situacijama, kao što su pronalaženje izvora slike, identifikacija objekata, pronalaženje sličnih vizuelnih sadržaja, otkrivanje plagijata ili pronalaženje više informacija o određenom objektu na slici. Reverzna pretraga slika može biti od pomoći u istraživačkim, kreativnim i bezbednosnim kontekstima, pružajući korisnicima mogućnost da pronađu relevantne slike na osnovu već postojeće vizuelne reference [2].

У овом раду, рјешаван је проблем имплементације једног оваквог система, гдје корисник може унијети произвољну слику као узорачки упит а систем му за исти враћа слике које најбоље одговарају на њега, поред тога, омогућено је и индексирање слика. Такође, адресиран је и проблем великог броја корисника који би потенцијално користили овај систем, те се развој водио праксама сервисно-орјентисане архитектуре гдје је сваки сервис могуће независно скалирати у произвољном обиму.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**NAPOMENA:**

**Ovaj rad proistekao je iz master rada čiji mentor je bio dr Dragan Ivanović, red. prof.**

**1.1. Svođenje problema pretrage vektora na pretragu teksta**

Jedinstvenost ovog rješenja leži u formatu svojstava koja se koriste prilikom indeksiranja i pretrage. Većina današnjih sistema se oslanja na vektorske reprezentacije dokumenata kako bi implementirala pretragu. Iako je ovo veoma efikasan način pretrage vizuelno sličnih slika, veliki je problem što se eksplicitno u obzir ne uzimaju svi objekti sa slike već se čitava slika reprezentuje kao neki "black-box" skup svojstava. Sa druge strane, moje rješenje svodi pretragu slika na pretragu teksta, modelom mašinskog učenja (YOLOv5) se prepoznaju objekti, koji se posmatraju kao ključne riječi prilikom indeksiranja u Elasticsearch system ya pronala\enje informacija, pomoću kojih možemo prepoznati sve kontekstualno slične slike sa uzorkom dok se prostor boja koristi kao opcioni filter kako bismo mogli dobiti vizuelno najsljičnije slike.

**1.2. Elasticsearch**

Elasticsearch je visoko skalirajući, distribuirani sistem za pretragu i analizu podataka. On je osmišljen da brzo i efikasno indeksira, čuva i pretražuje velike količine strukturisanih i nestrukturisanih podataka. Elasticsearch je baziran na Apache Lucene biblioteci, koja pruža moćne algoritme za pretragu. Takođe podržava brojne napredne funkcionalnosti kao što su pretraga punim tekstom (full-text search), filtriranje, agregacija, geolokacija, analiza teksta, vektorsko indeksiranje i pretraga i još mnogo toga.

Osim pretrage, Elasticsearch takođe omogućava analizu podataka i vizualizaciju rezultata. Kroz Kibana interfejs, korisnici mogu pristupiti bogatim vizualizacijama, izveštajima i nadzoru nad podacima. Elasticsearch je dizajniran da bude otporan na kvarove i pruža mogućnost replikacije podataka i distribucije preko više čvorova. Takođe podržava horizontalno i vertikalno skaliranje, omogućavajući dodavanje novih čvorova kako bi se postigla bolja brzina i otpornost na visoka opterećenja, kao i unapređenu dostupnost usluge. Sa svojom fleksibilnošću, skaliranjem i moćnim alatima za pretragu i analizu, Elasticsearch je postao popularan izbor za razne primene kao što su pretraga veb stranica, analiza logova i sistemi za praćenje događaja (event tracking systems).

**1.3. YOLOv5**

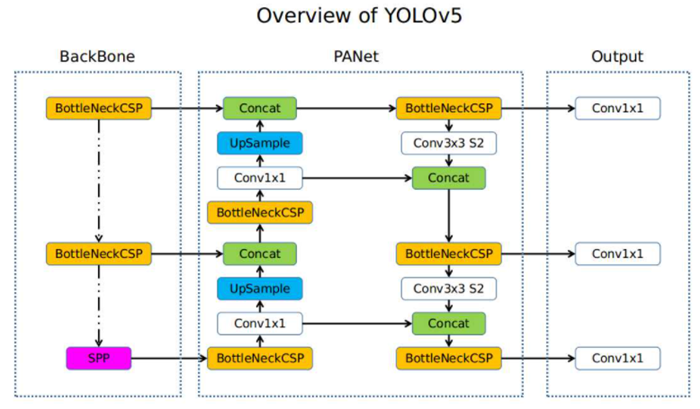
YOLOv5 model za detekciju objekata baziran je na end-to-end (e2e) arhitekturi i sastoji se od tri glavne komponente:

1. Vlasničke mreže (backbone network)
2. Glavne mreže (neck network)
3. Mreže izlaza (output network)

Vlasnička mreža (backbone) u YOLOv5 modelu je CSPDarknet53 mreža, koja se bazira na konvolucionim slojevima. Ova mreža je modifikovana verzija originalne Darknet53 mreže i pruža bolju reprezentaciju svojstava i razumevanje konteksta objekata. Ona je odgovorna za identifikaciju i izdvajanje bitnih svojstava iz ulaznih slika.

Glavna mreža (neck network) je dodatni sloj između vlasničke mreže i mreže izlaza. Ona ima za cilj da usavrši svojstva iz vlasničke mreže i podesi ih tako da bi lakše detektovala objekte različitih veličina i proporcija. U specifikaciji YOLOv5 modela, koristi se PANet mreža koja je korišćena i u starijem YOLOv4 modelu.

Mreža izlaza (output network) predstavlja poslednju komponentu modela i sadrži konvolucione slojeve koji izvode konačne predikcije o klasama i lokacijama objekata. Ona preobrađuje izlazne "feature" iz glavne mreže u anchor-box-ove i vraća koordinate, sigurnost i klase detektovanih objekata za svaki bounding-box. U pojedinoj literaturi, ovaj segment u "pipeline"-u modela naziva se i "head network". Na slici 3.9.1 prikazana je gore opisana arhitektura.



Slika 1. *Архитектура YOLOv5 модела*

**1.4. Pregled relevantne literature**

U radu [8] prikazana je *Similar Looks* funkcionalnost унутар Pinterest апликације која омогућава корисницима да изаберу предмет унутар слике, стављајући оквир око њега, а затим враћа визуелно сличне слике/предмете. Главни циљ ове функције је помоћи корисницима да пронађу ствари које не могу именовати [2].

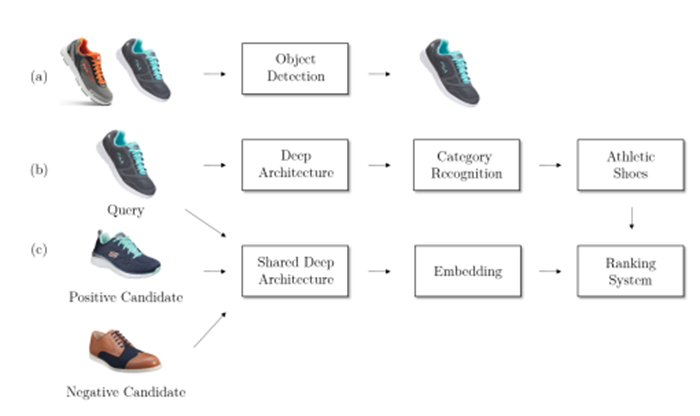
Индексирање слика се у овом систему врши по incremental fingerprinting service (IFS) систему. Овај систем функционише по принципу да свака слика има свој отисак који се састоји од свих одвојених објеката који се налазе унутар те слике (нпр. чипеле, торбе, чаше итд.). Сваки објекат има ознаку која специфицира име класе којој објекат припада [8].

Приликом претраге, систем користи технологију препознавања објеката да локализује и класификује предмете (објекте) на сликама. Из овако детектованих објеката се извлаче карактеристике и користе се за даљу претрагу. Другим ријечима, скраћује се задатак детекције објеката више класа у класификацију категорија, тако да уместо да се тражи подударање у свим сликама на Pinterest-u, прво се добављају слике које су у истој категорији. Ово је одличан приступ јер се упитна слика упоређује са сликама које имају високу вероватноћу да су сличне. Претходни корак филтрирања повећава стопу тачних позитива. Систем се заснива на концептима дубоког учења, као што су конволутивне неронске мреже (CNN) и детектори објеката засновани на дубоком учењу (Single Shot Detectors, SSD). Систем је имплементиран коришћењем радног оквира Caffe (Java технологија). Подаци система се скаладиште на Аmazon S3-у [2][8].

У раду [1] ток рада визуалне претраге је следећи: корисник upload-ује слику или је снима камером, а резултат ће бити сличне слике и предмети; корисник може изабрати да ли жели да купи (слично „Shopping“ секцији код Google претраживача) или даље истражује предмете [2]. Овај систем обухвата три главне фазе:

1. Обука модела: У систему се користи неколико модела дубоких неронских мрежа како би се побољшала релевантност резултата, као што су AlexNet, ZFSPPNet, GoogleNet, ResNet, а такође се користи и заједнички алгоритам кmeans како би се изградио инверзни индекс у нивоу-0 подударања, при чему ниво 0 означава најмању сличност са упитном сликом [1].
2. Разумевање упита (слике): Из упитне слике извлаче се различите карактеристике које описују њен садржај, укључујући карактеристике препознавања категорије, препознавања лица, боје и откривања дупликата, енкодере дубоких неронских мрежа (DNN) и детекцију објеката (слика 2.1.1).
3. Претрага слика: Проналази визуално сличне слике на основу извучених карактеристика и намера корисника.

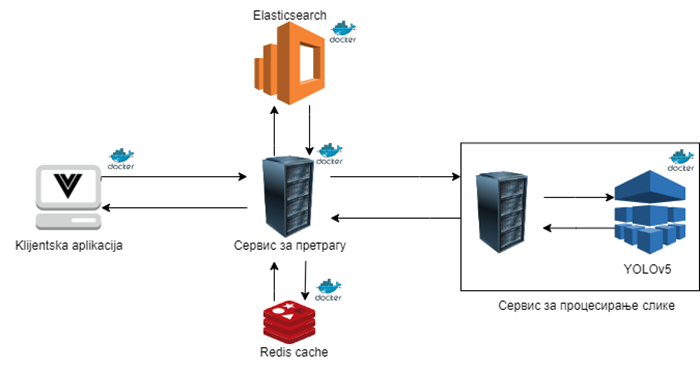
Сви модели се тренирају на обучавајућем скупу података који су прикупљени за одређене домене нпр. куповина. За надзор обуке DNN користе се више функција губитка, као што су Softmax, Pairwise-loss и Triplet-loss функције [1].



Slika 1. *Примjене DNN modela koji se koriste u Bing-овом систему*

**2. SPECIFIKACIJA SISTEMA ZA REVERZNU PRETRAGU SLIKA**

U ovom poglavlju prikazana je arhitektura sistema sa aspekta dijagrama razmeštaja. Dijagram razmeštaja (slika 1.) prikazuje softverske komponente i veze uz pomoć kojih iste komuniciraju.



Slika 1. *Dijagram razmeštaja*

Ово софтверско рјешење се састоји из два засебна сервиса:

* сервиса за претрагу и индексирање имплементираног у Java Spring Boot радном оквиру
* сервиса за процесирање слике имплементираног у Python FastAPI радном оквиру уз ослонац на ОpenCV и NumPy библиотеке за имплементацију и интеграцију неопходних алгоритама машинског учења који се користе приликом процесирања слике

Поред ова два сервиса битно је напоменути и да постоји одвојени frontend слој имплементиран у Vue3 радном оквиру користећи JavaScript програмски језик. Сервис за претрагу и индексирање комуницира са сервисом за процесирање слике како би извукао својства (features) са слике која је задата као узорак за упит. Поред функције проналажења информација и индексирања овај сервис врши и кеширање својстава слике користећи Redis key-value базу података како би убрзао вријеме претраге за више сукцесивних упита са истом сликом (приликом листања пагинираних резултата) што представља веома битан фактор у раду овог система како би био погодан за коришћење. Индексирање као и претрага је имплементирана користећи Elasticsearch систем за проналажење информација.

**3. REZULTATI I TUMAČENJA**

**4. ZAKLJUČAK**

U ovom radu smo vidjeli jedan način implementacije sistema za reverzno pretraživanje slika. Ovakav sistem omogućava korisnicima da umesto da se oslanjaju na ključne reči ili opise za pretragu slika, na efikasan način pronađu željene informacije na osnovu same slike. Ovo može da ima raznoliku primenu, od pomoći u identifikaciji različitih predmeta i lokacija, do potpunog unapređenja iskustva kupovine preko interneta.

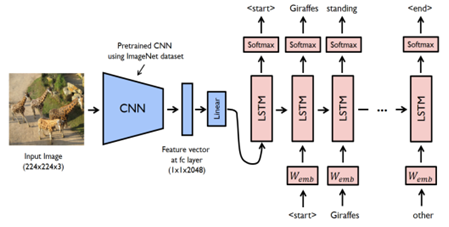
Sistem se sastoji iz dva odvojena servisa: servisa za pretragu koji vrši indeksiranje i pretragu koristeći Elasticsearch sistem za pronalaženje informacija, kao i keširanje koristeći Redis in-memory bazu, i servisa za procesiranje slike koji koristi moćne modele mašinskog učenja kako bi izvukao svojstva iz slike i predstavio je u formatu koji je lako pretraživ.

Za razliku od većine sistema koji vrše pretragu koristeći vektorski način pretrage (npr. Web-Scale Responsive Visual Search), ovaj sistem svodi problem pretrage slike na problem pretrage teksta gde se svojstva izvlače automatski, nasuprot tradicionalnom pristupu gde su se unosila ručno. Ovaj inovativan pristup već dugo je zastupljen u Pinterest-ovom Similar Search servisu i veoma lako može biti nadopunjen ručnim unosom opisa slike, ključnih reči itd.

**4.1. Dalji razvoj sistema**

S obzirom na to da je implementiran samo podskup različitih mogućnosti procesiranja slika, dodavanjem bilo kakvih dodatnih načina procesiranja moguće je postići poboljšanje sistema. Jedna stvar koja najviše ima smisla je da se postojeći način pretrage ukombinuje sa vektorskim načinom pretrage, tako da postojeći način isfiltrira kontekstno relevantne rezultate dok vektorska pretraga pronalazi vizuelno najsličniju sliku. Olakšavajuća okolnost kod ovog pristupa je ta što Elasticsearch podržava vektorsko indeksiranje i pretragu out-of-the-box.

Takođe, dodavanje modula za automatsko generisanje opisa ili ključnih reči iz slike korišćenjem mašinskog učenja moguće je poboljšati kvalitet izvučenih svojstava i proširiti domen primene gde bi se pored reverzne pretrage mogla implementirati i standardna tekst bazirana pretraga. Jedan od načina na koji bi ovo moglo biti uvedeno je korišćenjem CNN-a za ekstrakciju obeležja i rekurentnu neuronsku mrežu (poput LSTM-a) ili transformer za generisanje teksta.



Slika 1. *Приједлог архитектуре pipeline-а за генерисање описа слике*

Još jedan zanimljiv način implementacije je integracija ovog sistema sa nekim LLM-om koji može generisati opis/ključne reči slike (na primer, ChatGPT4).

**5. LITERATURA**

[1] Azeroual, O., & Schöpfel, J. (2019). Quality issues of CRIS data: An exploratory investigation with universities from twelve countries. Publications, 7(1), 14.

[2] <https://blog.scopus.com/posts/scopus-roadmap-whats-coming-up-in-2020-2021> (pristupljeno u septembru 2021.).

[3] Brito, G., Mombach, T., & Valente, M. T. (2019). Migrating to GraphQL: A practical assessment. In 2019 IEEE 26th International Conference on Software Analysis, Evolution and Reengineering (SANER) (pp. 140-150). IEEE.

[4] <https://arxiv.org/> (pregledano u septembru 2021.).

[5] Vogel, M., Weber, S., & Zirpins, C. (2017). Experiences on migrating RESTful web services to GraphQL. In International Conference on Service-Oriented Computing (pp. 283-295). Springer, Cham.

[6] Brito, G., & Valente, M. T. (2020). Rest vs graphql: A controlled experiment. In 2020 IEEE International Conference on Software Architecture (ICSA) (pp. 81-91). IEEE.

**Kratka biografija:**

|  |  |
| --- | --- |
|  | **Ivan Mršulja** rođen je 31.01.2000. godine u Kotoru. Smjer softversko inženjerstvo i informacione tehnologije na Fakultetu tehničkih nauka u Novom Sadu upisao je 2018 godine. Osnovne studije završio je u septembru 2022. godine. Od oktobra 2022. godine upisuje master studije i angažovan je kao saradnik u nastavi. |