

Anotacija skupa podataka za različite probleme dubokog učenja na primjeru ai.Shuttle robota

Brodarič, Luka

Undergraduate thesis / Završni rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Pula / Sveučilište Jurja Dobrile u Puli**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:137:857791>

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

Download date / Datum preuzimanja: **2024-12-20**



Repository / Repozitorij:

[Digital Repository Juraj Dobrila University of Pula](#)



Sveučilište Jurja Dobrile u Puli

Fakultet informatike u Puli

Luka Brodarič

Anotacija skupa podataka za različite probleme dubokog učenja na
primjeru ai.Shuttle robota

završni rad

Pula, rujan, 2023.

Sveučilište Jurja Dobrile u Puli

Fakultet informatike u Puli

Luka Brodarič

Anotacija skupa podataka za različite probleme dubokog učenja na
primjeru ai.Shuttle robota

završni rad

JMBAG: 0303094657, redovan student

Studijski smjer: Informatika

Znanstveno područje: Društvene znanosti

Znanstveno polje: Informacijske i komunikacijske znanosti

Znanstvena grana: Informacijski sustavi i informatologija

Kolegij: Sustavi poslovne inteligencije

Mentor: doc. dr. sc. Goran Oreški



IZJAVA O AKADEMSKOJ ČESTITOSTI

Ja, dolje potpisani Luka Brodarič, kandidat za prvostupnika Informatike ovime izjavljujem da je ovaj Završni rad rezultat isključivo mogega vlastitog rada, da se temelji na mojim istraživanjima te da se oslanja na objavljenu literaturu kao što to pokazuju korištene bilješke i bibliografija. Izjavljujem da niti jedan dio Završnog rada nije napisan na nedozvoljen način, odnosno da je prepisan iz kojega necitiranog rada, te da ikoji dio rada krši bilo čija autorska prava. Izjavljujem, također, da nijedan dio rada nije iskorišten za koji drugi rad pri bilo kojoj drugoj visokoškolskoj, znanstvenoj ili radnoj ustanovi.

Student _____

U Puli, _____, _____ godine



IZJAVA o korištenju autorskog djela

Ja, Luka Brodarič dajem odobrenje Sveučilištu Jurja Dobrile u Puli, kao nositelju prava iskorištavanja, da moj završni rad pod nazivom „Anotacija skupa podataka za različite probleme dubokog učenja na primjeru ai.Shuttle robota“ koristi na način da gore navedeno autorsko djelo, kao cjeloviti tekst trajno objavi u javnoj internetskoj bazi Sveučilišne knjižnice Sveučilišta Jurja Dobrile u Puli te kopira u javnu internetsku bazu završnih radova Nacionalne i sveučilišne knjižnice (stavljanje na raspolaganje javnosti), sve u skladu s Zakonom o autorskom pravu i drugim srodnim pravima i dobrom akademskom praksom, a radi promicanja otvorenoga, slobodnoga pristupa znanstvenim informacijama. Za korištenje autorskog djela na gore navedeni način ne potražujem naknadu

U Puli, _____ (datum)

Potpis

Sažetak:

Projekt ai.Shuttle predstavlja automobil robot koji pokreće Nvidia Jetson Nano. Dizajniran kao idealna platforma za istraživanje algoritama dubokog učenja u robotici, ai.Shuttle je opremljen nizom senzora i kamera koje omogućuju percepciju okoline i interakciju. Putem označavanja podataka i tehnika dubokog učenja, ai.Shuttle je osposobljen za navigaciju stazom, svladavanje prepreka, prepoznavanje objekata i izvođenje zadataka kao što su semantička segmentacija i otkrivanje objekata. Projekt obuhvaća različite faze, počevši od prikupljanja podataka iz stvarnog svijeta tijekom upravljanja robotom na stazi. Prikupljanje podataka uključivalo je snimanje fotografija, povezivanje koordinata s njima i utvrđivanje različitih klasa na fotografijama. Tehnike regresije isprva su korištene za predviđanje koordinata za navigaciju. Međutim, potreba za većom brzinom dovela je do uključivanja semantičke segmentacije. Cilj je bio stvoriti skup podataka koji omogućuje preciznu identifikaciju objekata i označavanje na fotografijama. Fotografije predstavljaju različite scenarije, uključujući klase ceste i prepreka. Najveći izazov s kojim smo se susreli bilo je različito osvjetljenje staze prilikom premještanja među prostorijama. Povećanje podataka poboljšalo je raznolikost i robusnost, dok su provjere kvalitete smanjile nedosljednosti u testiranju vožnje. Uvježbani model zatim je testiran za autonomnu navigaciju i izbjegavanje prepreka na stazi. Kao rezultat rada na projektu ai.Shuttle, ovaj završni rad istražuje te naglašava važnost preciznog prikupljanja podataka, bavi se uobičajenim izazovima strojnog učenja te ukazuje na etičke aspekte razvoja umjetne inteligencije. Također prikazuje put razvoja sustava za autonomnu vožnju, od prikupljanja podataka do izrade modela koji se testira na ai.Shuttle robotu.

Ključne riječi: ai.Shuttle, Nvidia Jetbot, duboko učenje, anotacija skupa podataka

Abstract:

The ai.Shuttle project presents a robot car powered by Nvidia Jetson Nano. Designed as an ideal platform for exploring deep learning algorithms in robotics, ai.Shuttle is equipped with a variety of sensors and cameras that enable environment perception and interaction. Through data annotation and deep learning techniques, ai.Shuttle is trained to navigate a path, overcome obstacles, recognize objects, and perform tasks such as semantic segmentation and object detection. The project encompasses different phases, starting with collecting real-world data while controlling the robot on the track. Data collection included taking photos, linking coordinates to them, and determining different classes in the photos. Regression techniques were initially used to predict coordinates for navigation. However, the need for greater speed led to the inclusion of semantic segmentation. The goal was to create a data set that enables precise object identification and labeling in photographs. The photos represent different scenarios, including road and obstacle classes. The biggest challenge we faced was the different lighting of the track when moving between rooms. Data augmentation has improved variety and robustness, while quality checks have reduced inconsistencies in test driving. The trained model was then tested for autonomous navigation and obstacle avoidance on the track. As a result of work on the ai.Shuttle project, this final paper explores and emphasizes the importance of accurate data collection, deals with common machine learning challenges, and points to the ethical aspects of artificial intelligence development. It also shows the development path of an autonomous driving system, from data collection to the creation of a model that is tested on the ai.Shuttle robot.

Keywords: ai.Shuttle, Nvidia Jetbot, deep learning, dataset annotation

Sadržaj

1. Uvod.....	1
1.1. Svrha završnog rada.....	1
1.2. Pregled ai.Shuttle (Jetbot) robota.....	2
1.3. Opis ciljeva završnog rada	2
2. Pregled literature	3
2.1. Pregled dostupne literature vezane za anotaciju podataka za duboko učenje.....	3
2.2. Pregled postojećih istraživanja vezanih za treniranje modela dubokog učenja na Jetbot robotima	3
3. Pregled dubokog učenja.....	4
3.1. Uvod u umjetnu inteligenciju (AI).....	5
3.2. Razumijevanje strojnog učenja	6
3.3. Nadzirano učenje.....	8
3.4. Učenje bez nadzora	8
3.5. Polu-nadzirano učenje	10
3.6. Duboko učenje - potpodručje strojnog učenja.....	11
3.7. Važnost i primjena dubokog učenja	12
4. Problemi dubokog učenja	13
4.1. Pretreniranost i podtreniranost	13
4.2. Nedostatak dovoljno označenih podataka.....	14
4.3. Neuravnoteženost podataka.....	15
4.4. Računalni resursi i vrijeme obuke	15
4.5. Etičko razmatranje privatnosti i sigurnosti.....	16
5. Anotacija skupa podataka	17
5.1. Definicija i važnost anotacije skupa podataka	17

5.2. Vrste anotacija skupa podataka.....	18
5.3. Klasifikacija fotografija.....	19
5.4. Anotacija otkrivanja objekata.....	20
5.5. Semantička segmentacija	22
5.6. Alati i tehnike za označavanje skupa podataka	26
6. Uvod u ai.Shuttle robot.....	27
6.1. Uvod i slučaj upotrebe.....	28
6.2. NVidia Jetson Nano hard ver	29
6.3. Svrha ai.Shuttle robota	30
7. Stvaranje skupa podataka za ai.Shuttle robot	31
7.1. Proces prikupljanja skupa podataka	31
7.2. Tehnike označavanja podataka.....	34
7.3. Izazovi i ograničenja u anotaciji skupa podataka.....	39
7.4. Testiranje modela autonomne vožnje i izbjegavanje prepreka	40
8. Zaključak.....	41
8.1. Sažetak završnog rada i njegovih doprinosa	41
8.2. Budućnost projekta.....	41
8.3. Predstavljanje ai.Shuttle projekta na danima maturanata.....	42
9. Reference	43

1. Uvod

Anotacija skupa podataka ključan je korak u obuci modela dubokog učenja. Uključuje označavanje podataka relevantnim informacijama, metapodacima, kako bi se modelu pomoglo u učenju i izradi točnih predviđanja. Duboko učenje podskup je strojnog učenja koje koristi neuronske mreže za učenje iz podataka. Cilj ovog završnog rada je prezentirati proces označavanja podataka za razne probleme dubokog učenja koristeći ai.Shuttle robot kao primjer.

ai.Shuttle robot automobil idealna je platforma za istraživanje upotrebe algoritama dubokog učenja u robotici jer može prikupljati podatke s raznih senzora ili kamera, te se ti podaci koriste za treniranje modela dubokog učenja za zadatke kao što su otkrivanje i zaobilazanje objekata, segmentacija i praćenje.

1.1. Svrha završnog rada

Primarna svrha ovog završnog rada je istražiti proces označavanja skupa podataka i njegov značaj u obuci modela dubokog učenja. Bilježenje skupa podataka uključuje zadatak označavanja podataka relevantnim informacijama i metapodacima kako bi se pomoglo modelu u učenju i izradi točnih predviđanja. Usredotočujući se na ai.Shuttle robota kao glavni primjer, ovaj završni rad ima za cilj pokazati praktičnu primjenu označavanja skupa podataka za različite probleme dubokog učenja. ai.Shuttle robot, kojeg pokreće Nvidia Jetson Nano, služi kao idealna platforma za istraživanje upotrebe algoritama dubokog učenja u robotici. Ima mogućnost prikupljanja podataka s različitih senzora i kamera, koji se zatim koriste za treniranje modela dubokog učenja za zadatke kao što su otkrivanje i izbjegavanje objekata, segmentacija i praćenje. Ovim završnim radom nastojim pokazati vrijednost označavanja skupa podataka u omogućavanju ai.Shuttleu da učinkovito upravlja svojim okruženjem i izvršava složene robotske zadatke.

1.2. Pregled ai.Shuttle (Jetbot) robota

ai.Shuttle robot, izrađen je na Nvidia Jetson Nano platformi koja predstavlja značajan napredak u autonomnoj robotici. Nvidia Jetson Nano platforma uključuje Nvidia Jetson Nano developer kit, koji služi kao mozak robota ai.Shuttle. Jetson Nano opremljen je Nvidia GPU-om, četverojezgrenim ARM Cortex-A57 CPU-om i 4 GB RAM-a, pružajući značajne računalne mogućnosti za zadatke dubokog učenja. Što se tiče ostalog hardvera, robot ai.Shuttle ima niz senzora i kamera koje mu omogućuju opažanje i interakciju s okolinom. Softverska arhitektura robota ai.Shuttle temelji se na ROS-u (engl. Robot Operating System), široko prihvaćenom okviru za razvoj robotskih aplikacija. ROS olakšava učinkovitu komunikaciju između različitih komponenti robota, omogućujući besprijekornu integraciju percepcije, kontrole i modula dubokog učenja.

1.3. Opis ciljeva završnog rada

Završni rad kojem je temelj ai.Shuttle je postavljen s nekoliko jasnih ciljeva. Prvo, cilj je demonstrirati proces označavanja skupa podataka za različite probleme dubokog učenja koristeći robota ai.Shuttle kao praktičan primjer. Označavanjem prikupljenih podataka relevantnim informacijama cilj je stvoriti visokokvalitetni označeni skup podataka koji će služiti kao temelj za obuku modela dubokog učenja. Drugo, ima za cilj istražiti i prikazati mogućnosti algoritama dubokog učenja u omogućavanju ai.Shuttle robotu da se kreće svojim okruženjem, otkriva i izbjegava prepreke, izvodi semantičku segmentaciju i prati objekte. Iskorištavanjem snage dubokog učenja, nastojimo poboljšati autonomiju ai.Shuttle robota i poboljšati njegovu ukupnu izvedbu. U konačnici, završni rad ima za cilj pridonijeti širem razumijevanju označavanja skupa podataka i njegovog značaja u području dubokog učenja i autonomne vožnje.

2. Pregled literature

2.1. Pregled dostupne literature vezane za anotaciju podataka za duboko učenje

Proces označavanja podataka služi kao temeljni korak u obuci robusnih i učinkovitih modela dubokog učenja. Opsežna literatura posvećena je istraživanju različitih metoda, izazova i napretka u području označavanja podataka za aplikacije dubokog učenja. Alati za označavanje igraju ključnu ulogu u pojednostavljivanju procesa označavanja podataka. Labelbox, popularna platforma za označavanje, nudi sveobuhvatan uvid u raznolike alate za označavanje podataka i njihov utjecaj na tijekove rada strojnog učenja. (Labelbox, 2022). Dovoljno označeni podaci ključni su u modelima obuke koji se dobro generaliziraju u različitim scenarijima. Goodfellow, Bengio i Courville (2016) govore o središnjoj ulozi označenih podataka u obuci neuronskih mreža i naglašavaju važnost reprezentativnih skupova podataka za postizanje visokih performansi modela. U kontekstu zadataka računalnog vida, gdje vizualni podaci čine osnovu učenja, literatura Szeljski (2010) te Gonzalez i Woods (2007) pruža temeljno znanje o obradi slika te ekstrakciji značajki iz njih. Ovi resursi pomažu u razumijevanju složenosti zadataka dubokog učenja te rješavanja problema pomoću skupa podataka koji se temelji na slikama. Napredak u robotici i autonomnim sustavima također je pridonio evoluciji metodologija označavanja podataka. Autori Correll, Hayes, i Grocholsky (2016) naglašavaju važnost označenih podataka u modelima obuke za autonomnu navigaciju i kontrolu.

2.2. Pregled postojećih istraživanja vezanih za treniranje modela dubokog učenja na Jetbot robotima

Obuka modela dubokog učenja na Jetbot robotima privukla je značajnu pozornost istraživača zbog svog potencijala za poboljšanje sposobnosti autonomnih sustava. Istraživanje koje su proveli Kawakura i Shibasaki (2020) fokusira se na razvoj sustava autonomnog automobila koji se temelji na principima dubokog učenja za dostavu predmeta poljoprivrednim radnicima unutar poljoprivrednog okruženja. Fokus istraživanja je na robotu u obliku automobila, koji pokreće

Nvidia Jetson Nano, koji je uključen u pokuse provedene u stvarnom skladištu gdje su bili razbacani predmeti poput rukavica, čizama, škara i motika koji su bili označeni kao prepreke. Cilj je bio razvoj sustava posebno dizajniranog za precizno otkrivanje prepreka, korištenjem tehnika dubokog učenja, uključujući također automatiziranu isporuku malih predmeta poljoprivrednim radnicima. U istraživanju, robot je okarakteriziran kao vrlo sposoban za manevriranje, vješt u prepoznavanju prepreka te učinkovit u učenju iz okoline.

U drugome istraživanju autora Chang, Juan, Huang i Chen (2020) predstavlja se analiza scenarija praćenja ceste korištenjem Nvidia Jetbot platforme. U radu se istražuje primjena ResNet-a (engl. residual neural network) za praćenje ceste, pri čemu autori treniraju model za vođenje bespilotnih kopnenih vozila (engl. UGV - unmanned ground vehicle) na temelju ulaznih slika i izlaznih vrijednosti. Proces uključuje korištenje PyTorch-a za obuku modela arhitekture neuronske mreže ResNet18. Rad ilustrira postupke obrade i predviđanja za praćenje ceste, ocrtavajući korake od prikupljanja podataka do instalacije modela na robotu. Rezultati sugeriraju da specifični uvjeti poput kutova fotografiranja fotografija utječu na odabir točaka za kretanje unaprijed te održavanje točnosti putanje.

U sljedećem istraživanju autora Ciolacu, Haderer, Berl i Svasta (2021), autori se bave analizom i procjenom jedinica za grafičku obradu (engl. GPU - graphics processing unit) koje čine sastavne komponente sustava za treniranje modela dubokog učenja. Istraženi kriterij obuhvaćaju lakoću korištenja, modularnost, prilagodljivost te isplativost. Temelj ovog istraživanja jest interakcija jedinica za grafičku obradu u interakciji sa sustavima umjetne inteligencije, fokusirajući se na autonomne robote poput Nvidia Jetbot-a.

3. Pregled dubokog učenja

Duboko učenje je specijalizirano potpolje strojnog učenja koje se fokusira na obuku umjetnih neuronskih mreža s više slojeva za učenje iz složenih podataka. Privuklo je značajnu pozornost i

postiglo izuzetan uspjeh u raznim domenama, uključujući autonomnu vožnju, prepoznavanje govora i slično.

Srž dubokog učenja su umjetne neuronske mreže, koje su inspirirane strukturom i funkcioniranjem ljudskog mozga. Te se mreže sastoje od međusobno povezanih čvorova, zvanih neuroni, organiziranih u slojeve. Svaki neuron prima ulazne signale, izvodi proračune i proizvodi izlazni signal koji se prosljeđuje sljedećim neuronima. Modeli dubokog učenja obično imaju više skrivenih slojeva, što im omogućuje učenje zamršenih uzoraka i prikaza u podacima.

Duboke neuronske mreže mogu otkriti složene obrasce i prikaze na način od početka do kraja, što im omogućuje da briljiraju u zadacima kao što su prepoznavanje slika, otkrivanje objekata i semantička segmentacija. Iskorištavanjem snage dubokog učenja, istraživači su postigli revolucionarne rezultate u područjima kao što su autonomna vožnja, analiza medicinskih slika i razumijevanje prirodnog jezika.

3.1. Uvod u umjetnu inteligenciju (AI)

Umjetna inteligencija (engl. AI - artificial intelligence) je područje koje se fokusira na razvoj inteligentnih sustava sposobnih za obavljanje zadataka koji obično zahtijevaju ljudsku inteligenciju. Obuhvaća širok raspon tehnologija, algoritama i metodologija čiji je cilj omogućiti računalima da simuliraju kognitivne sposobnosti nalik ljudskim. Cilj umjetne inteligencije je stvoriti strojeve koji mogu percipirati okolinu, razmišljati o informacijama, donositi informirane odluke i učiti iz iskustva. Oponašanjem ljudske inteligencije, AI ima za cilj povećati produktivnost, automatizirati zadatke koji se ponavljaju, poboljšati procese donošenja odluka i omogućiti napredak u različitim domenama. (Russel & Norvig, 2016).

Korijeni umjetne inteligencije mogu se pratiti od sredine 20. stoljeća kada su istraživači počeli istraživati koncept strojeva koji mogu pokazati inteligentno ponašanje. Rani pristupi umjetne inteligencije bili su usredotočeni na simboličko razmišljanje, korištenje logike i sustava temeljenih na pravilima za obradu i manipuliranje znanjem. Međutim, značajan napredak u računalnoj snazi, dostupnosti podataka i algoritamskim inovacijama u posljednjim desetljećima doveo je AI do novih visina.

Umjetna inteligencija obuhvaća različita potpodručja, uključujući strojno učenje, obradu prirodnog jezika, računalni vid, robotiku i reprezentaciju znanja. Posebno se strojno učenje pojavilo kao istaknuto područje unutar umjetne inteligencije, omogućujući računalima da uče iz podataka bez eksplicitnog programiranja. Ovo potpolje revolucioniralo je način na koji se razvijaju AI sustavi, omogućujući strojevima da izdvajaju uzorke, daju predviđanja i prilagode svoje ponašanje na temelju podataka s kojima se susreću.

Umjetna inteligencija pronalazi primjenu u brojnim domenama, uključujući zdravstvo, financije, prijevoz, obrazovanje i zabavu. U zdravstvu, sustavi umjetne inteligencije mogu pomoći u dijagnosticiranju bolesti, analizi medicinskih slika i personalizaciji planova liječenja. U financijama se algoritmi umjetne inteligencije koriste za otkrivanje prijevara, algoritamsko trgovanje i procjenu rizika. U prijevozu umjetna inteligencija pokreće autonomna vozila, optimizirajući rute i povećavajući sigurnost. AI također prevladava u industriji zabave, gdje se koristi za sustave preporuka, virtualne pomoćnike i generiranje sadržaja. „Mnoge ideje koje su kreirala istraživanja umjetne inteligencije asimilirane su u računalnu znanost općenito za primjene u ekspertnim sustavima, obradi prirodnog jezika, interakciji čovjeka i računala, obradi grafike i slike, rudarenju podataka i robotici.“ (Nilsson, 2014).

Vrlo bitna stvar oko korištenja umjetne inteligencije jesu etička razmatranja. Kako sustavi umjetne inteligencije postaju sve sofisticiraniji i utjecajniji, pitanja privatnosti, pristranosti, transparentnosti i odgovornosti izbijaju u prvi plan.

3.2. Razumijevanje strojnog učenja

Strojno učenje je grana umjetne inteligencije koja se bazira na stvaranju algoritama i modela sposobnih za učenje iz podataka bez potrebe za programiranjem. Takav proces možemo promatrati kao sposobnost računala (stroja) da uči i poboljšava se iz iskustva, na sličan način kao što ljudi uče iz vlastitih opažanja i interakcija s okolinom.

U svojoj srži, strojno učenje omogućuje računalima da analiziraju i razumiju složene obrasce i odnose unutar podataka, omogućujući im da donose predviđanja ili odluke. Uključuje

uvježbavanje modela korištenjem ogromne količine podataka koji pomažu modelu shvatiti temeljne obrasce i opća načela prisutna u podacima.

Postoje različite vrste algoritama strojnog učenja, svaki sa svojim jedinstvenim pristupom i karakteristikama. Strojno učenje može se općenito kategorizirati u nekoliko vrsta, od kojih svaka ima svoj jedinstveni pristup i svrhu. Jedan uobičajeni tip je nadzirano učenje (engl. supervised learning), gdje se model uvježbava na označenim primjerima, što znači da su podaci već kategorizirani ili klasificirani. Učeći iz ovih označenih primjera, model tada može predvidjeti oznake novih, nevidenih podataka. Ovo je osobito korisno za zadatke poput prepoznavanja slika ili otkrivanja neželjene e-pošte.

Druga vrsta strojnog učenja je učenje bez nadzora (engl. unsupervised learning), koje uključuje modele obuke na neoznačenim podacima. Ovdje je cilj da model identificira smislene obrasce ili strukture unutar podataka bez ikakvih unaprijed definiranih oznaka. Ovaj se pristup često koristi za zadatke kao što je klasteriranje sličnih točaka podataka ili pronalaženje skrivenih uzoraka u velikim skupovima podataka.

Učenje s pojačanjem (engl. reinforcement learning) još je jedna grana strojnog učenja, inspirirana konceptom kako ljudi uče kroz nagrade i kazne. U ovom pristupu, model uči interakcijom s okolinom i primanjem povratnih informacija u obliku nagrada ili kazni. Kroz pokušaje i pogreške, model otkriva optimalne akcije ili strategije za maksimiziranje ukupne nagrade.

Strojno učenje je polje koje se iznimno brzo razvija, a istraživači neprestano razvijaju nove algoritme i tehnike učenja. Sa sve većom dostupnošću velikih skupova podataka i napretkom računalne snage, strojno učenje nastavlja pomicati granice onoga što računala mogu postići. Razumijevanjem principa i tehnika strojnog učenja, možemo otključati potencijal ovog snažnog alata, omogućujući rješavanje složenih problema, pravljenje točnih predviđanja i dobivanje vrijednih uvida iz podataka. Također razumijevanjem karakteristika i primjenom svake vrste, istraživači mogu odabrati najprikladniji pristup za svoje projekte strojnog učenja.

3.3. Nadzirano učenje

Nadzirano učenje je istaknuti pristup u strojnom učenju gdje modeli uče iz označenih podataka kako bi napravili točna predviđanja ili klasificirali nove, neviđene primjere. U ovoj vrsti učenja skup podataka sastoji se od ulaznih podataka zajedno s odgovarajućim ciljnim oznakama ili željenim izlazima. Cilj je uvježbati model koji može razumjeti obrasce i odnose između ulaznih značajki i njihovih odgovarajućih oznaka. (Bishop, 2006).

Tijekom procesa obuke, model analizira označene primjere kako bi identificirao temeljne obrasce i naučio kako preslikati ulazne značajke na njihove odgovarajuće izlaze. Generalizira to znanje kako bi dao predviđanja na temelju novih, neviđenih podataka, čak i ako se prije nije susreo s tim konkretnim primjerima.

Nadzirano učenje obično se koristi u raznim primjerima, kao što je klasifikacija slika, gdje je cilj klasificirati slike u unaprijed definirane kategorije. Na primjer, nadzirani model učenja obučen na velikom skupu podataka označenih slika može naučiti prepoznavati i razlikovati različite objekte, kao što su automobili, bicikli i pješaci, u stvarnom vremenu. U našem slučaju iznimno je važno da model može prepoznati oznake ceste, znakove kao i prepreke na putu.

Drugi primjer je otkrivanje neželjene e-pošte. Uvježbavanjem modela nadziranog učenja na skupu podataka označenih e-poruka (neželjena pošta ili ne), model može naučiti identificirati obrasce u sadržaju e-pošte i odrediti je li dolazna e-pošta vjerojatno neželjena ili ne.

Nadzirano učenje omogućuje strojevima generalizaciju iz poznatih primjera i primjenu njihovog razumijevanja na nevidljive slučajeve, što ga čini moćnim pristupom za širok raspon zadataka koji zahtijevaju točna predviđanja ili klasifikaciju.

3.4. Učenje bez nadzora

Učenje bez nadzora je pristup u strojnom učenju gdje modeli pregledavaju neoznačene podatke kako bi otkrili temeljne obrasce, strukture ili odnose bez vodstva unaprijed definiranih ciljnih

oznaka. Nudi jedinstvenu priliku za izvlačenje značajnih uvida iz sirovih podataka i stjecanje dubljeg razumijevanja karakteristika podataka.

Za razliku od nadziranog učenja, gdje se modeli oslanjaju na označene primjere kako bi naučili obrasce i dali predviđanja, nenadzirano učenje radi na neoznačenim podacima, dopuštajući modelu da autonomno istražuje i uči iz intrinzične strukture podataka. „Primarni zadatak učenja bez nadzora se razmatra kao modeliranje strukture razvodnika podataka.“ (Bengio, Courville, i Vincet, 2021). Ova vrsta učenja često se koristi kada dostupnim podacima nedostaju eksplicitne oznake ili kada je cilj otkriti skrivene obrasce ili odnose koji možda nisu odmah vidljivi. U nenadziranom učenju, model funkcionira ispitivanjem statističkih svojstava podataka i identificiranjem sličnosti ili klastera unutar primjera. Iskorištavanjem različitih algoritama i tehnika, modeli učenja bez nadzora mogu otkriti obrasce koji bi inače mogli ostati neprimijećeni, otkrivajući vrijedne uvide o temeljnim podacima.

Kako bismo bolje razumjeli učenje bez nadzora, razmotrimo primjer korištenja velikog skupa podataka o transakcijama kupaca s internetskog tržišta. Bez ikakvih unaprijed definiranih oznaka, model učenja bez nadzora može istraživati skup podataka i identificirati različite skupine kupaca na temelju njihovog ponašanja pri kupnji. Grupiranjem kupaca koji pokazuju slične obrasce kupnje, model može pružiti vrijedan uvid u segmentaciju kupaca, omogućujući tvrtkama da prilagode svoje marketinške strategije ili identificiraju ciljne segmente kupaca za određene kampanje. Tehnike učenja bez nadzora često uključuju algoritme klasteriranja koji grupiraju slične podatkovne točke zajedno na temelju njihovih intrinzičnih sličnosti. Ovi klasteri mogu otkriti prirodne skupine unutar podataka, omogućujući istraživačima da steknu dublje razumijevanje složenih fenomena ili identificiraju skrivene subpopulacije.

Druga uobičajena tehnika u učenju bez nadzora je smanjenje dimenzionalnosti. Mnogi skupovi podataka iz stvarnog svijeta su visokodimenzionalni, sadrže brojne značajke ili atribute. Metode smanjenja dimenzionalnosti pomažu pojednostaviti podatke izvlačenjem manjeg skupa reprezentativnih značajki uz zadržavanje bitnih informacija. Ovo smanjenje dimenzionalnosti može pomoći u vizualizaciji, kao i poboljšanju učinkovitosti i performansi naknadnih zadataka strojnog učenja. Dopuštajući modelima da istražuju neoznačene podatke i otkrivaju skrivene obrasce, učenje bez nadzora nudi snažan pristup za dobivanje uvida i razumijevanje složenih skupova podataka.

3.5. Polu-nadzirano učenje

Polu-nadzirano učenje (engl. semi-supervised learning) je strojno učenje koje kombinira elemente nadziranog i nenadziranog učenja. Ono se bavi izradom modela kada je samo mali dio dostupnih podataka označen, dok značajan dio ostaje neoznačen. Ova jedinstvena paradigma omogućuje modelima da iskoriste ograničene označene podatke zajedno s obiljem neoznačenih podataka za postizanje poboljšane izvedbe i generalizacije. U polu-nadziranom učenju, mali označeni skup podataka služi kao vodeći signal za model, pružajući primjere s poznatim oznakama iz kojih se može učiti. Ovi označeni primjeri pomažu modelu da uspostavi početne obrasce i napravi predviđanja na temelju dostupnih informacija. Međutim, prava snaga polu-nadziranog učenja leži u njegovoj sposobnosti da iskoristi golemu količinu neoznačenih podataka za daljnje usavršavanje i poboljšanje razumijevanja osnovnih obrazaca unutar podataka. Korištenjem neoznačenih podataka, koji često čine značajan udio ukupnog skupa podataka, polu-nadzirani modeli učenja mogu učinkovitije uhvatiti temeljnu distribuciju i strukturu podataka. Model se može generalizirati s označenih podataka na neoznačene podatkovne točke koje pokazuju slične uzorke ili svojstva, čime se poboljšava njegova sposobnost pravljenja točnih predviđanja na novim, neviđenim primjerima.

Jedna od ključnih prednosti polu-nadziranog učenja je njegova sposobnost da iskoristi neoznačene podatke za rješavanje problema ograničene dostupnosti označenih podataka. Prikupljanje označenih podataka može biti skupo, dugotrajno ili čak neizvedivo u određenim scenarijima. Uključivanjem neoznačenih podataka, polu-nadzirano učenje može maksimalno iskoristiti dostupne resurse, omogućujući modelima da uče iz šireg i reprezentativnijeg uzorka temeljne populacije. Uobičajeni pristup u polu-nadziranom učenju je kombiniranje tehnika nadziranog učenja, kao što su algoritmi klasifikacije, s tehnikama nenadziranog učenja kao što su grupiranje ili smanjenje dimenzionalnosti. Označeni podaci koriste se za obuku modela korištenjem nadziranih metoda, dok neoznačeni podaci pomažu modelu da otkrije dodatne obrasce ili strukture koje mogu dodatno poboljšati njegovu izvedbu.

Polu-nadzirano učenje pokazalo se vrlo korisnim u raznim domenama. U obradi prirodnog jezika, može se koristiti za poboljšanje zadataka klasifikacije teksta, gdje su dostupni ograničeni označeni podaci, iskorištavanjem velikih količina neoznačenih tekstualnih podataka. U računalnom vidu,

polu-nadzirano učenje uspješno je primijenjeno na zadatke kao što su klasifikacija slika i prepoznavanje objekata, omogućujući modelima da uče i iz označenih slika i iz ogromne količine neoznačenih slika dostupnih na internetu.

3.6. Duboko učenje - potpodručje strojnog učenja

Duboko učenje je potpolje strojnog učenja koje je posljednjih godina privuklo značajnu pozornost. Usredotočeno je na uvježbavanje umjetnih neuronskih mreža s više slojeva za izdvajanje zamršenih uzoraka i prikaza iz složenih podataka. U srcu dubokog učenja su umjetne neuronske mreže, inspirirane međusobno povezanim neuronima u našem mozgu. „Duboko učenje omogućuje računalnim modelima koji se sastoje od višestrukih slojeva obrade da nauče prikaze podataka s više razina apstrakcije.“ (LeCun, Bengio, i Hinton, 2015).

Obuka modela dubokog učenja uključuje proces u dva koraka: širenje naprijed i širenje unatrag (engl. propagation). Tijekom širenja prema naprijed, podaci teku mrežom, a proračuni se izvode na svakom sloju za generiranje predviđanja. Ta se predviđanja zatim uspoređuju sa željenim rezultatom, a izvedba modela se procjenjuje pomoću funkcije gubitka.

Širenje unazad ključ je za obuku modela dubokog učenja. Uključuje izračunavanje gradijenata funkcije gubitka s obzirom na parametre modela i prilagođavanje težine i pristranosti mreže u skladu s tim. Ovaj iterativni proces pomaže modelu naučiti i poboljšati svoja predviđanja minimiziranjem razlike između predviđenih i željenih rezultata.

Jedna od ključnih prednosti dubokog učenja je njegova sposobnost automatskog učenja značajki iz neobrađenih podataka, čime se eliminira potreba za ručnom obradom. Duboke neuronske mreže mogu otkriti složene obrasce i prikaze na način od kraja do kraja, što ih čini vrlo učinkovitim u zadacima kao što su prepoznavanje slika i govora, pa čak i igranje igara poput šaha. „Poteškoće s kojima se suočavaju sustavi koji se oslanjaju na tvrdo kodirano znanje sugeriraju da sustavi umjetne inteligencije trebaju sposobnost stjecanja vlastitog znanja izvlačenjem uzoraka iz neobrađenih podataka.“ (Goodfellow, Bengio i Courville, 2016).

Područje dubokog učenja svjedočilo je značajnom napretku posljednjih godina, potaknutom otkrićima u mrežnim arhitekturama, algoritmima optimizacije i dostupnošću skupova podataka velikih razmjera. Tehnike poput konvolucijskih neuronskih mreža (engl. CNN - Convolutional Neural Network) revolucionirale su zadatke računalnog tipa, dok su rekurentne neuronske mreže (engl. RNN - Recurrent Neural Networks) omogućile napredak u sekvencijalnoj obradi podataka.

3.7. Važnost i primjena dubokog učenja

Duboko učenje dobilo je golemu važnost i pronašlo je široku primjenu u raznim domenama, revolucionirajući način na koji pristupamo složenim problemima. Njegov značaj leži u njegovoj sposobnosti obrade i razumijevanja velikih količina podataka, omogućujući strojevima da naprave točna predviđanja, prepoznaju obrasce i generiraju vrijedne uvide.

Jedno od ključnih područja gdje je duboko učenje imalo značajan utjecaj je računalni vid. Modeli dubokog učenja izvrsni su u zadacima kao što su klasifikacija slika, detekcija objekata i semantička segmentacija. Korištenjem konvolucijskih neuronskih mreža, strojevi mogu precizno analizirati vizualne podatke, otvarajući vrata aplikacijama kao što su autonomna vožnja, sustavi nadzora i analiza medicinskih slika.

U obradi prirodnog jezika (engl. natural language processing), duboko učenje poboljšalo je razumijevanje i generiranje jezika. Duboke neuronske mreže, u kombinaciji s tehnikama kao što su rekurentne neuronske mreže i transformatori, omogućuju strojevima da razumiju i generiraju jezik sličan ljudskom. To je dovelo do napretka u strojnom prevođenju, analizi osjećaja, chatbotovima i sažimanju teksta, poboljšanju interakcije između čovjeka i stroja i olakšavanju zadataka temeljenih na jeziku. Duboko učenje također ima značajan utjecaj na prepoznavanje govora i obradu zvuka. Korištenjem dubokih neuronskih mreža, strojevi mogu transkribirati govorni jezik, izvršiti prepoznavanje govornika, pa čak i generirati realističan govor. To ima posebnu ulogu tijekom razvoja glasovnih asistenata, usluge prijepisa i poboljšanje pristupačnosti za osobe s oštećenjem sluha.

U području sustava preporuka, modeli dubokog učenja pokazali su se vrlo učinkovitima. Analizirajući korisničke preferencije i povijesne podatke, algoritmi dubokog učenja mogu pružiti

personalizirane preporuke za proizvode, filmove, glazbu i još mnogo toga. To je promijenilo način na koji tvrtke komuniciraju sa svojim klijentima, što je dovelo do poboljšanja korisničkog iskustva i povećanog zadovoljstva kupaca.

Još jedna značajna primjena dubokog učenja je u autonomnim sustavima, posebno u autonomnim automobilima. Algoritmi dubokog učenja mogu obraditi podatke u stvarnom vremenu iz raznih senzora i donositi odluke na temelju složenih scenarija vožnje. Ova tehnologija ima potencijal za revoluciju u prijevozu, čineći ga sigurnijim i učinkovitijim, dok također smanjuje nesreće uzrokovane ljudskom greškom.

Dok nastavljamo istraživati njegove mogućnosti, usavršavati algoritme i razvijati nove arhitekture, duboko učenje obećava otključavanje novih prilika i poticanje inovacija u svim industrijama.

4. Problemi dubokog učenja

Duboko učenje nedvojbeno je revolucioniralo razna područja svojim izvanrednim mogućnostima. Međutim, kao i svaka napredna tehnologija, dolazi sa svojim nizom izazova i ograničenja. Razumijevanje problema kojima se možemo suočiti kao i raznih izazova ključno je za razvoj učinkovitih rješenja i iskorištavanje punog potencijala dubokog učenja u stvarnim aplikacijama.

4.1. Pretreniranost i podtreniranost

Pretreniranost (engl. overfitting) i podtreniranost (engl. underfitting) nastaju tijekom obuke modela te mogu značajno utjecati na performanse modela te njegove izvedbe. Pronalaženje prave ravnoteže između pretreniranosti i podtreniranosti ključno je za obuku uspješnih modela dubokog učenja. U idealnom slučaju, cilj je postići model koji se dobro generalizira na nevidljive podatke dok učinkovito hvata obrasce prisutne u podacima o obuci. (Hastie, Tibshirani, i Friedman, 2009).

Pretrreniranost se događa kada model dubokog učenja postane previše specijaliziran za učenje iz podataka do te mjere da se ne uspijeva dobro prilagoditi novim, neviđenim podacima. Drugim riječima, model pamti primjere obuke umjesto da uči nove temeljne obrasce. Kao rezultat toga, kada se prezentiraju novi podaci, pretrrenirani model može proizvesti netočna predviđanja jer se ne može prilagoditi podacima koje prije nije vidio. Kako bi se riješilo problem pretrreniranosti, koriste se različite tehnike regularizacije, kao što su “dropout” te smanjivanje težine. “Dropout” uključuje nasumično deaktiviranje neurona tijekom treninga, sprječavajući model da se previše oslanja na specifične neurone i potiče robusnije učenje. Smanjenje težine, s druge strane, kažnjava velike težine u modelu, sprječavajući da model postane pretjerano složen. S druge strane, podtrreniranost se događa kada model dubokog učenja ne uspije uhvatiti temeljne obrasce u podacima o obuci. U ovom scenariju, kapacitet modela može biti nedovoljan da bi se saznala složenost podataka. Posljedično, loše radi ne samo na podacima o obuci, već i na novim podacima. Podtrreniranost karakterizira nedostatak fleksibilnosti u modelu, što dovodi do ograničenih mogućnosti učenja i predviđanja. Problem podtrreniranosti može se riješiti na nekoliko načina, a neki od njih su dodavanje veće količine trening podataka, te dodavanje više epoha prilikom treniranja modela.

4.2 Nedostatak dovoljno označenih podataka

Jedan od glavnih izazova u dubokom učenju je nedostatak dovoljno označenih podataka. Modeli dubokog učenja zahtijevaju znatnu količinu označenih primjera za učenje i učinkovito generaliziranje. Međutim, dobivanje označenih podataka može biti dugotrajan, skup i radno intenzivan proces, posebno u domenama gdje je stručnost u domeni neophodna za točne bilješke. Za mnoge složene zadatke, prikupljanje velikih označenih skupova podataka može biti nepraktično ili čak nemoguće. „Dostupnost označenih podataka o vježbanju iznimno je ograničena zbog prirode podataka koji sadrže zaštićene zdravstvene informacije.“ (Droz, Mukherjee i Gevaert, 2022). Ovo ograničenje može značajno ometati izvedbu modela dubokog učenja. Nedostatak označenih podataka može dovesti do pretjeranog prilagođavanja, budući da se model bori s generalizacijom izvan ograničenih primjera obuke i umjesto toga pamti specifične slučajeve.

Kako bi riješili problem nedovoljno označenih podataka, možemo pribjeći tehnikama povećanja podataka. Povećanje podataka uključuje generiranje dodatnih primjera obuke primjenom

transformacija na postojeće označene podatke. Na primjer u istraživanju autora Wang, Gao, Long, i Wang, (2021)., u zadacima klasifikacije slika, povećanje podataka može uključivati okretanje, rotiranje ili skaliranje slika kako bi se stvorile varijacije izvornih podataka. Ovi prošireni primjeri služe kao vrijedni dodaci, pomažući u povećanju raznolikosti skupa za obuku i poboljšanje generalizacije modela.

4.3. Neuravnoteženost podataka

Neuravnoteženost podataka nastaje kada je distribucija klasa ili kategorija u podacima o obuci jako iskrivljena. Drugim riječima, određene klase imaju znatno više primjera od drugih, što dovodi do nejednake zastupljenosti podatkovnih točaka u različitim kategorijama. To može nepovoljno utjecati na izvedbu modela dubokog učenja, posebno u zadacima klasifikacije. Kada je klasa nedovoljno zastupljena u trening podacima model može imati problema s učenjem obrazaca povezanih s tom klasom jer ima manje primjera iz kojih može učiti. Kao rezultat toga, model može dati prioritet većinskoj klasi, što dovodi do pristranih predviđanja i smanjene točnosti manjinskih klasa. U scenarijima s velikom neuravnoteženosti podataka, model može čak pribjeći jednostavnom predviđanju većinske klase za sve instance, što rezultira visokom točnošću za dominantnu klasu, ali lošom izvedbom u cjelini. Uspostavljanje prave ravnoteže i pronalaženje najprikladnijeg pristupa rješavanju neravnoteže podataka ključno je za izgradnju točnih i robusnih modela dubokog učenja, osobito u scenarijima u kojima prevladavaju neuravnoteženi podaci.

4.4. Računalni resursi i vrijeme obuke

Modeli dubokog učenja često su veliki i složeni te zahtijevaju znatnu računalnu snagu i memoriju za učinkovito treniranje. Proces obuke uključuje višestruke iteracije ili epohe, gdje model usavršava svoje parametre kako bi smanjio pogrešku u podacima obuke. Svaka epoha može biti računalno intenzivna, posebno za duboke arhitekture s mnogo slojeva. „Ovo vrijeme obuke može se promatrati kao umnožak vremena obuke po epohi i broja epoha koje je potrebno izvesti da bi se postigla željena razina točnosti.“ (Justus, Brennan, Bonner i McGough, 2017).

Potražnja za računalnim resursima postaje još izraženija s povećanjem veličine skupova podataka i složenosti modela. Ukoliko imamo ograničeni pristup visokoučinkovitim računalnom hardveru, obuka modela dubokog učenja može postati dugotrajna i skupa.

4.5. Etičko razmatranje privatnosti i sigurnosti

Etička razmatranja povezana s privatnošću i sigurnošću ključni su aspekti koji se moraju razmatrati u dubokom učenju. Kako modeli dubokog učenja postaju sve moćniji i sposobniji za obradu golemih količina podataka, zabrinutost oko toga kako se ti podaci prikupljaju, koriste i štite izbila je u prvi plan.

Jedno od primarnih etičkih pitanja je privatnost pojedinaca čiji se podaci koriste za treniranje modela dubokog učenja. U mnogim slučajevima veliki skupovi podataka sastoje se od osjetljivih informacija, poput osobnih podataka, medicinske dokumentacije ili financijskih transakcija. Osiguravanje da su ti podaci anonimizirani i pravilno zaštićeni ključno je za zaštitu prava pojedinaca na privatnost. Prilikom treniranja modela moramo primijeniti snažne tehnike anonimizacije podataka i pridržavati se propisa o zaštiti podataka kako bi spriječili bilo kakav neovlašteni pristup ili otkrivanje osobnih podataka. Ispitanici bi trebali imati pravo kontrolirati kako se njihovi podaci koriste i biti u potpunosti informirani o svrhama u koje će se koristiti. (Ashurst, Barocas, Campbell i Raji, 2022).

Također ako podaci o obuci odražavaju društvene predrasude ili stereotipe, model može ovjekovječiti te predrasude u svojim predviđanjima. To može imati dalekosežne posljedice, dovesti do diskriminirajućih ishoda u procesima donošenja odluka. Zato moramo nastojati identificirati i ublažiti pristranosti u podacima i osigurati da su modeli poštene i pravični.

Još jedno etičko razmatranje potencijalni su sigurnosni rizici povezani s modelima dubokog učenja. Suparnički napadi, gdje zlonamjerni akteri namjerno manipuliraju ulaznim podacima kako bi prevarili model, predstavljaju značajne sigurnosne probleme. Ovi napadi mogu imati ozbiljne implikacije u kritičnim aplikacijama, kao što su autonomna vozila ili zdravstveni sustavi. Istraživači moraju razviti robusne modele koji su otporni na kontradiktorne napade i uzeti u obzir potencijalne sigurnosne ranjivosti prilikom implementacije rješenja dubokog učenja.

5. Anotacija skupa podataka

Anotacija skupa podataka uključuje dodavanje relevantnih informacija i oznaka neobrađenim podacima kako bi se modelima omogućilo učenje i točna predviđanja. (Everingham, Van Gool, Williams, Winn i Zisserman, 2010). Proces uključuje zadatke kao što su otkrivanje objekata, segmentacija slike, označavanje teksta i slično, ovisno o specifičnom problemu dubokog učenja. Bilješke skupa podataka igraju ključnu ulogu u obučavanju ai.Shuttle robota da učinkovito upravlja svojim okruženjem, osiguravajući da može prepoznati i reagirati na različite objekte i prepreke na stazi. U ovom odjeljku istražiti ćemo vrste korištenih metoda označavanja skupa podataka, njihove izazove i njihov značaj u pripremi visokokvalitetnih podataka za obuku modela dubokog učenja.

5.1. Definicija i važnost anotacije skupa podataka

Označavanje skupa podataka temeljni je proces dubokog učenja koji uključuje obogaćivanje neobrađenih podataka dodatnim informacijama kako bi bili smisleni i korisni za modele obuke. (Russakovsky i ostali, 2015). Putem naredbi dajemo oznake ili metapodatke, omogućujući algoritmima dubokog učenja da razumiju i prepoznaju obrasce, objekte ili odnose unutar podataka.

Anotirani podaci služe kao temelj za obuku modela dubokog učenja, omogućujući im da uče iz primjera i generaliziraju na nove, neviđene podatke. Bez odgovarajućih napomena, modelu bi nedostajale potrebne smjernice za točna predviđanja ili klasifikacije.

Na primjer, u klasifikaciji slika, označavanje slika odgovarajućim oznakama (npr. "mačka" ili "pas") omogućuje modelu da nauči razlikovne značajke svake klase, što mu omogućuje da ispravno klasificira nove slike. Slično tome, u detekciji objekata, napomene daju granične okvire oko objekata od interesa, omogućujući modelu da locira i prepozna određene objekte unutar slika.

Kvaliteta i točnost označavanja skupa podataka izravno utječu na izvedbu i pouzdanost modela dubokog učenja. Precizne i dosljedne primjedbe ključne su kako bi se osigurala sposobnost modela da generalizira i napravi robusna predviđanja u scenarijima stvarnog svijeta.

Dok smo razvijali skup podataka za obuku robota ai.Shuttle, pedantna i promišljena anotacija podataka bila je presudna u omogućavanju robotu da se pouzdano kreće stazom, izbjegava prepreke, prepoznaje prometne znakove i donosi informirane odluke na temelju svog okruženja.

5.2. Vrste anotacija skupa podataka

Anotacija skupa podataka nije ograničena na slike; proširuje se na različite vrste podataka, uključujući tekst, audio i video. U obradi prirodnog jezika, tekstualne bilješke uključuju označavanje teksta za analizu osjećaja, prepoznavanje imenovanih entiteta ili zadatke prijevoda jezika. Audio bilješke mogu uključivati identifikaciju govornih segmenata ili klasificiranje zvukova u audio snimkama.

U nastavku ću se osvrnuti na tri najčešća zadatka u računalnom vidu: klasifikaciju, detekciju i segmentaciju te postupak anotacije podataka za svaki navedeni zadatak.

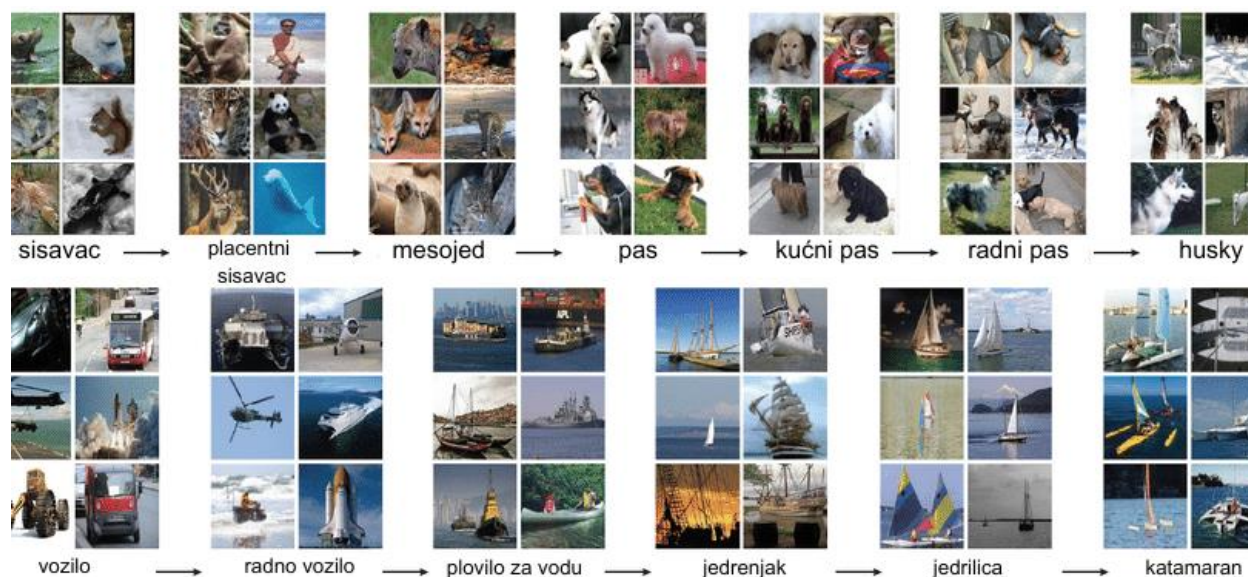


Slika 1. – Prikaz vrsta anotacija skupa podataka (LMB)

5.3. Klasifikacija fotografija

Klasifikacija fotografija temeljna je i naširoko korištena vrsta označavanja skupa podataka u dubokom učenju. To uključuje dodjeljivanje jedne oznake ili kategorije svakoj slici u skupu podataka, omogućujući modelu da točno prepozna i klasificira objekte unutar slika. Ovaj je zadatak posebno vrijedan u različitim aplikacijama iz stvarnog svijeta, kao što je prepoznavanje objekata, prepoznavanje lica i dohvaćanje slika temeljenih na sadržaju.

Kako bi izvršili klasifikaciju slika, moramo pažljivo ispitati svaku fotografiju i dodijeliti joj oznaku koja najbolje predstavlja prikazani primarni objekt ili scenu. Na primjer, u skupu podataka slika životinja, svaka slika može biti označena oznakama kao što su "mačka", "pas", "zec" ili "slon", na temelju primarne životinje prisutne na slici. Označeni skupovi podataka postaju trening podaci za model dubokog učenja. Tijekom treninga, model uči identificirati i izdvojiti bitne značajke iz slika povezanih sa svakom oznakom. Zatim koristi te naučene značajke za točno razlikovanje i klasificiranje novih, neviđenih slika.



Slika 2. – Prikaz klasifikacije fotografija

Modeli dubokog učenja za klasifikaciju slika često koriste konvolucijske neuronske mreže (CNN), koje su prikladne za učenje prostornih hijerarhija značajki na fotografijama. CNN-ovi koriste

slojeve filtara za otkrivanje uzoraka i značajki u različitim razmjerima, omogućujući modelu da uhvati zamršene detalje i složene odnose unutar slika. „U usporedbi sa standardnim neuronskim mrežama sa slojevima slične veličine, CNN-ovi imaju mnogo manje veza i parametara i stoga ih je lakše trenirati, dok će njihova teoretski najbolja izvedba vjerojatno biti samo malo lošija.“ (Krizhevsky, Sutskever i Hinton, 2012). Izvedba modela klasifikacije slika uvelike ovisi o kvaliteti i raznolikosti označenog skupa podataka. Skup podataka mora predstavljati širok raspon varijacija, kao što su različite poze, uvjeti osvjetljenja i pozadine, kako bi se osigurala robusnost modela i sposobnost dobre generalizacije na nevidljive podatke. Primjene klasifikacije slika viđamo u raznim domenama, od autonomnih vozila koja prepoznaju prometne znakove do medicinskih slikovnih sustava koji dijagnosticiraju bolesti, klasifikacija slika igra ključnu ulogu u omogućavanju strojeva da točno razumiju i interpretiraju vizualne informacije.

5.4. Anotacija otkrivanja objekata

Anotacija otkrivanja objekata posebno se koristi za aplikacije u kojima je ključno precizno lociranje i identificiranje objekata unutar slika. Za razliku od klasifikacije slika, gdje se jedna oznaka dodjeljuje cijeloj slici, označavanje detekcije objekta uključuje precizno definiranje granica svakog objekta od interesa unutar slike. Na primjer, u scenariju autonomne vožnje, bilješke za otkrivanje objekata mogu uključivati crtanje graničnih okvira oko pješaka, vozila, prometnih znakova i drugih relevantnih objekata na cesti.

Označeni skup podataka služi za obuku za modele detekcije objekata, omogućujući im da nauče otkrivati i točno locirati objekte na novim, dosad neviđenim slikama. Modeli dubokog učenja za otkrivanje objekata često koriste sofisticirane arhitekture, kao što su regionalne konvolucijske neuronske mreže (R-CNN) ili detektori pojedinačnih kutija (engl. SSD - single-shot detector), koji su posebno dizajnirani za otkrivanje objekata u složenim scenama.

Osiguravanje visokokvalitetnih zabilješki u detekciji objekata ključno je za postizanje točne i pouzdane izvedbe modela. Izazovi u anotaciji otkrivanja objekata uključuju rukovanje objektima različitih oblika i veličina ili preklapajućih objekata te točno identificiranje granica objekta.

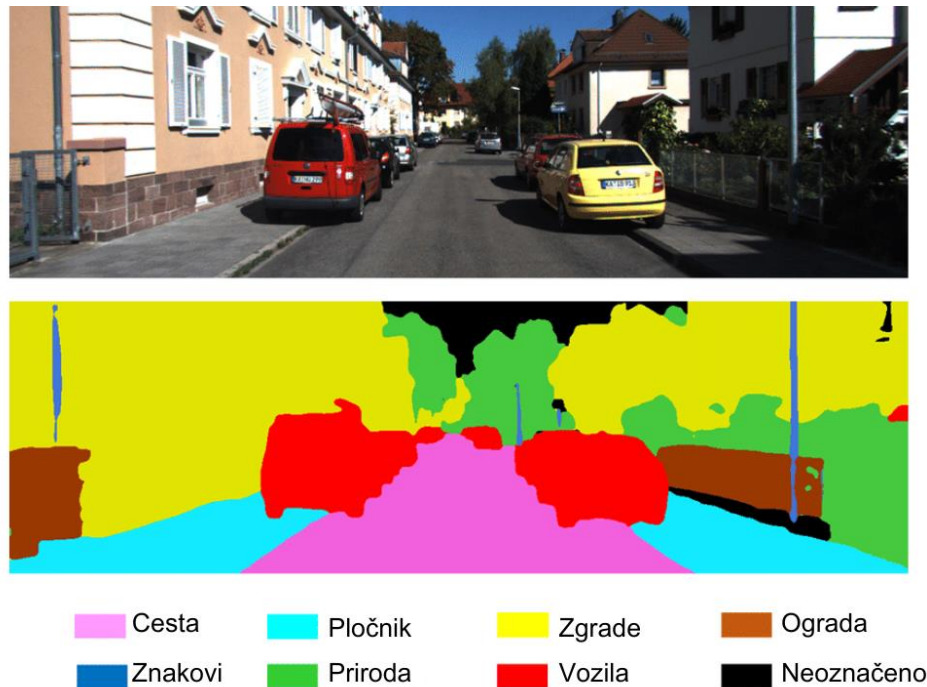


Slika 3. – Prikaz detekcije prometnih znakova na cesti

5.5. Semantička segmentacija

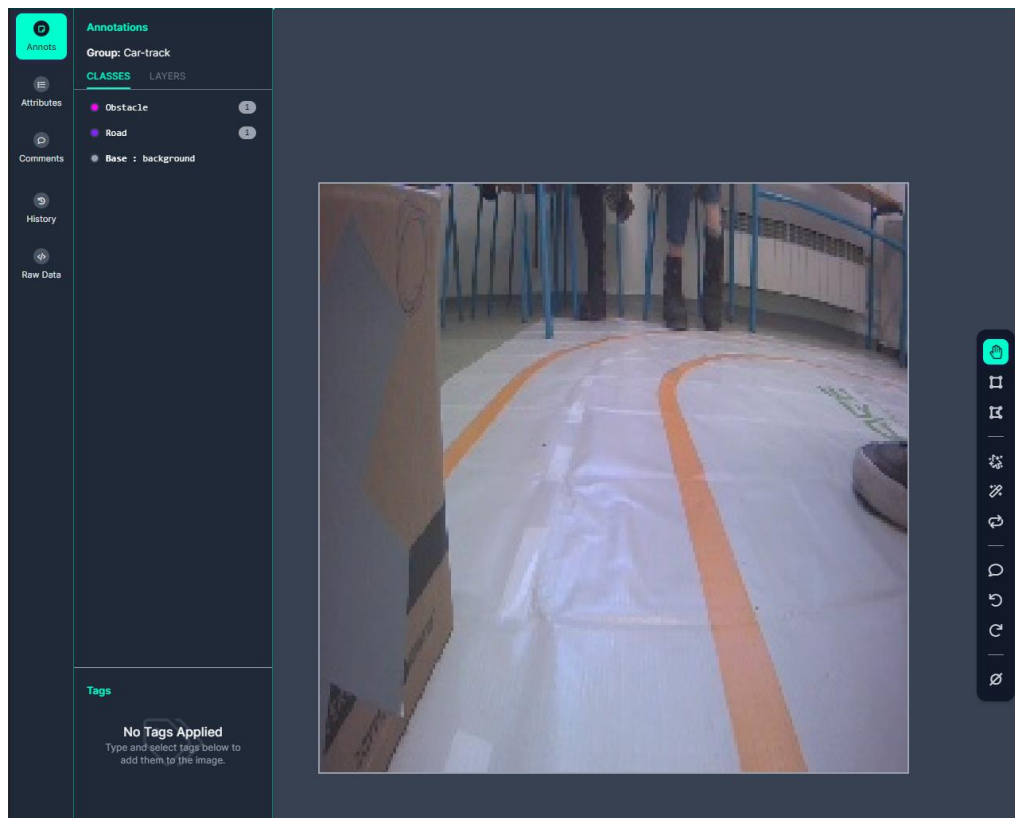
Semantička segmentacija je sofisticirana vrsta označavanja skupa podataka u dubokom učenju, fokusirana na razumijevanje slika na razini piksela. Za razliku od klasifikacije slika ili detekcije objekata, gdje je cilj dodijeliti jednu oznaku cijeloj slici ili lokalizirati određene objekte, semantička segmentacija ima za cilj segmentirati svaki piksel slike i dodijeliti mu oznaku klase koja predstavlja semantičku kategoriju kojoj pripada.

U anotaciji semantičke segmentacije, svaki piksel pažljivo je označen, stvarajući gustu oznaku slike po pikselu. Rezultirajuće bilješke tvore detaljnu i preciznu kartu različitih regija ili objekata prisutnih na slici. Na primjer, u uličnoj sceni, anotacija semantičke segmentacije može uključivati označavanje svakog piksela kao "cesta", "zgrada", "pješač", "auto" ili druge relevantne klase, čime se pruža bogato razumijevanje sadržaja scene. (Guo, Liu, Georgiou i Lew, 2017).

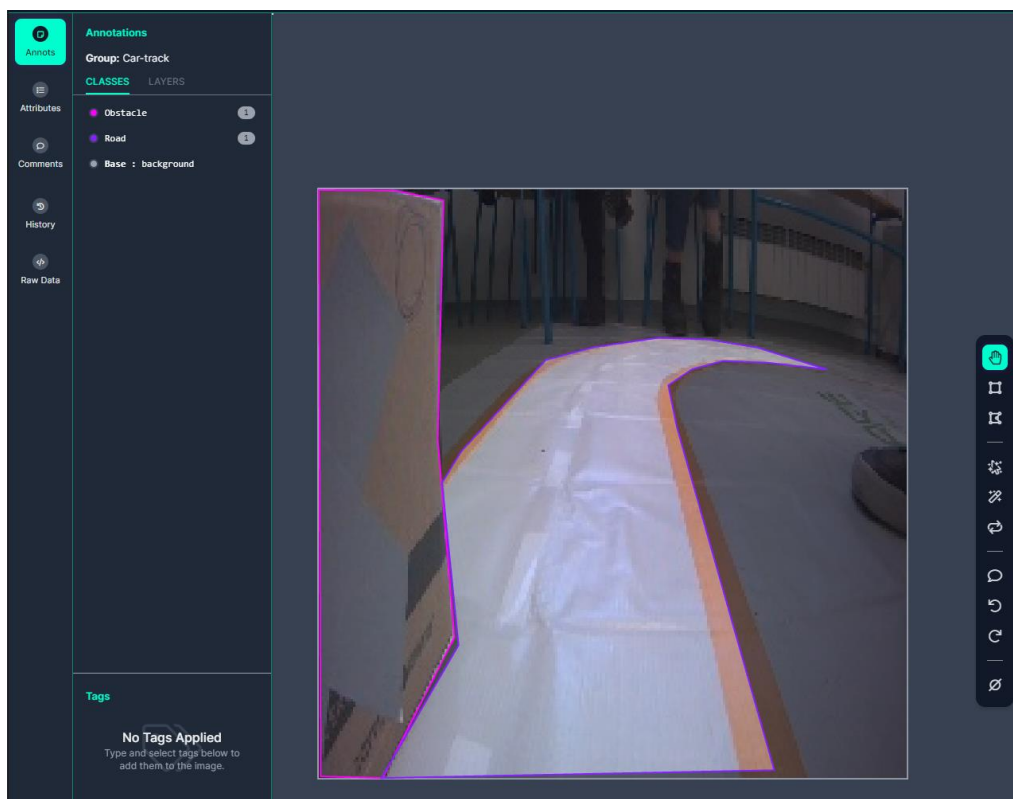


Slika 4. – Prikaz semantičke segmentacije u uličnoj sceni (Towards AI)

Pristup ručnog označavanja omogućuje visoku preciznost jer su lako razumljive složene konture objekata na slici. Na platformi Roboflow označavanje slika može se obaviti koristeći njihovo sučelje za označavanje. Unutar sučelja priložimo slike koje želimo označiti, zatim na slici označimo željene objekte, te na te objekte dodamo oznake odnosno klase. Nakon što smo označili objekte i dodijelili im klase, spremimo promjene te sustav pohrani oznake na sliku, te možemo preuzeti slike u odabranom formatu.



Slika 5. – Prikaz neoznačene slike unutar sučelja platforme Roboflow



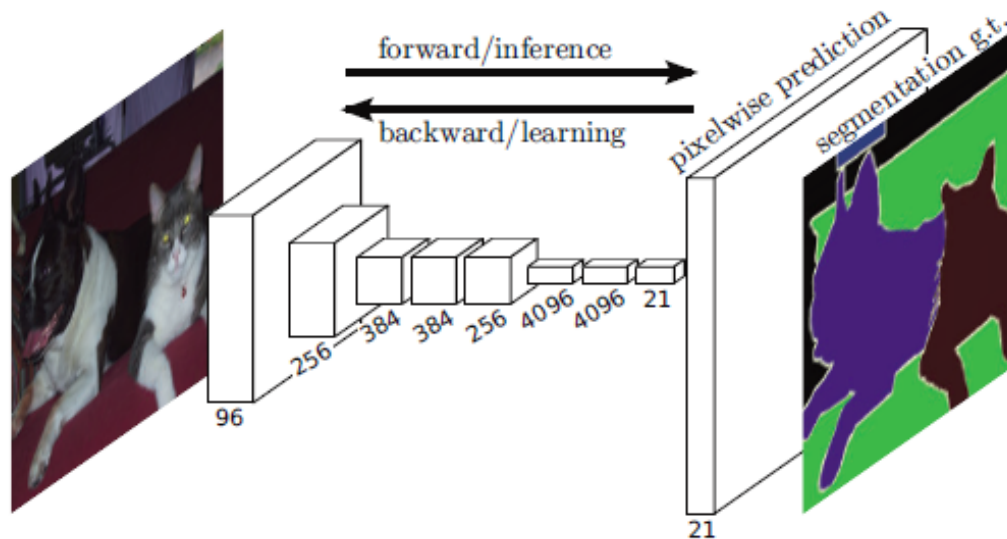
Slika 6. – Prikaz označenih klasa „obstacle“ i „road“ na slici unutar sučelja platforme Roboflow

Alternativno, za velike skupove podataka, možemo koristiti alate za automatsko označavanje koji koriste tehnike strojnog učenja kako bi brže označili slike. Ovi alati koriste prethodno označene podatke kao referencu za označavanje novih slika. Iako su brži, automatski alati za označavanje mogu biti manje precizni od ručnog označavanja i zahtijevati dodatnu provjeru.

Jedan od izazova u semantičkoj segmentaciji je raznolikost osvjetljenja na slikama. Promjene u osvjetljenju mogu značajno utjecati na točnost segmentacije, jer objekti mogu izgledati različito pod različitim uvjetima osvjetljenja. Na primjer, isti objekt može izgledati svijetlo i jasno na jednoj slici, dok će pod drugim uvjetima osvjetljenja biti zamagljen ili tamniji.

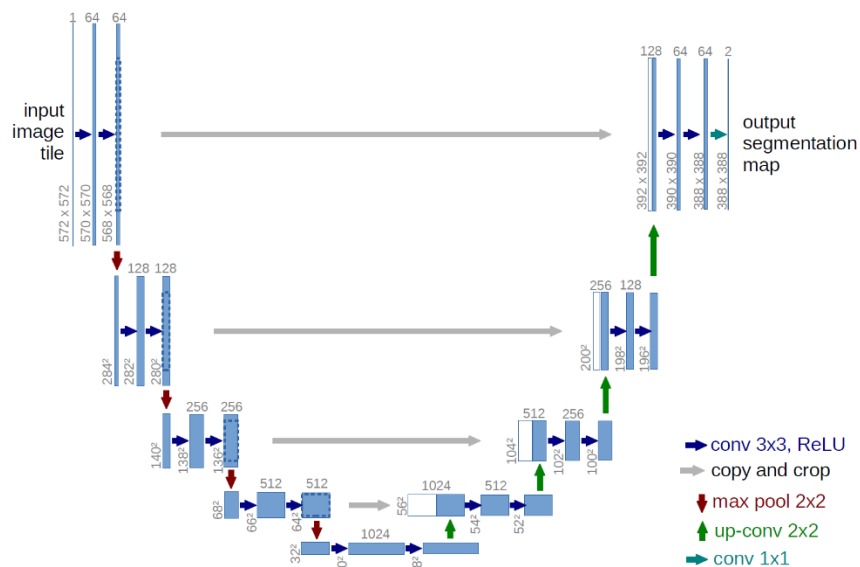
Modeli dubokog učenja za semantičku segmentaciju često koriste potpuno konvolucijske neuronske mreže (engl. FCN - fully convolution network) ili U-Net arhitekture, koje su posebno dizajnirane za zadatke predviđanja po pikselima. Ovi modeli koriste konvolucijske slojeve za hvatanje prostornih informacija i učenje točnog segmentiranja objekata. FCN-ovi često koriste slojeve za povećanje rezolucije (engl. upsample) kako bi izlaz modela bio iste dimenzije kao i

ulazna slika. Ovi slojevi za povećanje rezolucije omogućuju modelu precizno označavanje svakog piksela slike sa semantičkom klasom kojoj pripada.



Slika 7. – Prikaz potpuno konvolucijske neuronske mreže (Towards data science)

U-Net arhitektura ima karakterističan oblik slova "U" i sastoji se od konvolucijskih slojeva za ekstrakciju značajki i slojeva za povećanje rezolucije kako bi se dobila visoka razlučivost slike.



Slika 8. – Prikaz U-NET arhitekture (LMB)

Osim FCN-a i U-Net arhitektura, postoje i drugi modeli i arhitekture koji se koriste u semantičkoj segmentaciji. Na primjer, Mask R-CNN kombinira semantičku segmentaciju s detekcijom objekata kako bi pružio precizne maske za svaki objekt na slici.

Primjene semantičke segmentacije su raznolike i obuhvaćaju različite domene. U autonomnoj vožnji semantička segmentacija omogućuje vozilima razumijevanje prizora na cesti, prepoznavanje granica traka i prepoznavanje prometnih znakova i pješaka. U medicinskom snimanju, pomaže u segmentiranju organa ili tumora za točnu dijagnozu i planiranje liječenja.

5.6. Alati i tehnike za označavanje skupa podataka

Alati i tehnike označavanja skupa podataka igraju ključnu ulogu u pojednostavljivanju procesa označavanja i osiguravanju kvalitete označenih podataka za obuku modela dubokog učenja. Ručno označavanje velikih i složenih skupova podataka može biti dugotrajno i podložno pogreškama, zbog čega su automatizirani ili poluautomatizirani alati ključni za učinkovite i točne anotacije. Postoji širok raspon dostupnih alata za anotacije koji služe za procese kao što su klasifikacija slika, otkrivanje objekata i semantička segmentacija. Ovi alati pružaju korisničko sučelje koje

omogućuje korisnicima da s lakoćom crtaju granične okvire, segmentiraju objekte ili označavaju piksele. Neki popularni alati za označavanje jesu Labelbox, VGG Image Annotator (VIA) i RectLabel.

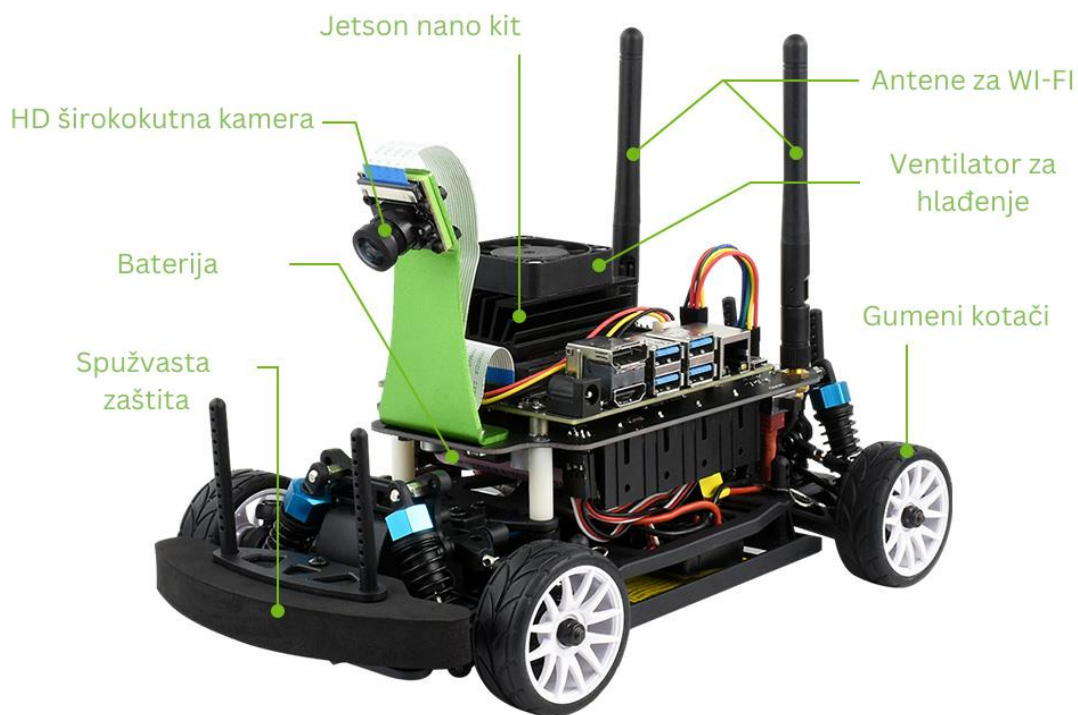
Tehnike povećanja podataka mogu povećati označeni skup podataka generiranjem dodatnih primjera obuke kroz transformacije poput okretanja, rotiranja ili skaliranja slika. Povećanje podataka ne samo da povećava raznolikost skupa za obuku, već također pomaže u poboljšanju generalizacije i robusnosti modela.

Aktivno učenje (engl. active learning) je tehnika u strojnom učenju koja se koristi za poboljšanje performansi modela dubokog učenja s manje označenih podataka. Osnovna ideja aktivnog učenja je da model sam odabire koje primjere treba dodatno označiti iz skupa nepoznatih podataka kako bi poboljšao svoje predviđanje.

U situacijama u kojima ručno označavanje velikih skupova podataka nije izvedivo, mogu se upotrijebiti platforme za masovno prikupljanje za distribuciju zadatka označavanja između više anotatora. Crowd-sourcing omogućuje brže bilješke i može pružiti različite perspektive, poboljšavajući ukupnu kvalitetu skupa podataka.

6. Uvod u ai.Shuttle robot

ai.Shuttle robot je inovativan i inteligentan mali robot automobil kojeg pokreće Nvidia Jetson Nano. Dizajniran kao idealna platforma za istraživanje algoritama dubokog učenja u robotici, ai.Shuttle je opremljen raznim sensorima i kamerama koje mu omogućuju percepciju okoline i interakciju s njom. Putem označavanja podataka i dubokog učenja, ai.Shuttle se obučava za navigaciju na maloj stazi, svladavanje prepreka, prepoznavanje objekata i izvođenje zadataka kao što su semantička segmentacija i otkrivanje objekata.



Slika 9. – Prikaz komponenti Nvidia Jetson Nano robota (Nvidia Jetson Nano Developer Kit)

6.1. Uvod i slučaj upotrebe

Jedna od primarnih upotreba ai.Shuttle robota je autonomna navigacija na unaprijed definiranoj stazi. Putem označavanja podataka i tehnika dubokog učenja, robot je osposobljen da prepozna i reagira na različite elemente na stazi, kao što su granice ceste, prepreke i prometni znakovi. Obradom podataka senzora i iskorištavanjem računalnih mogućnosti Nvidia Jetson Nano, ai.Shuttle može donositi inteligentne odluke, prilagođavajući svoju putanju i brzinu za sigurno i učinkovito manevriranje duž staze.

ai.Shuttle služi kao izvrstan primjer kako se duboko učenje može primijeniti u robotskim scenarijima stvarnog svijeta. Iskorištavanjem snage Nvidia Jetson Nano, robot dobiva mogućnost obrade podataka u stvarnom vremenu i izvršavanja složenih algoritama dubokog učenja. To otvara

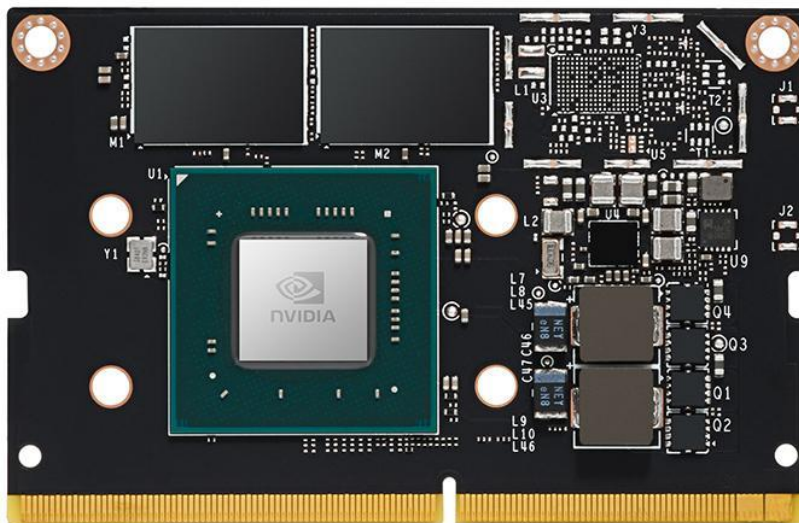
mnoštvo mogućnosti za istraživanje i eksperimentiranje u robotici, omogućujući razvoj sofisticiranih autonomnih sustava s praktičnom primjenom.

Štoviše, autonomne navigacijske sposobnosti ai.Shuttlea imaju vrijedne primjene izvan slučaja upotrebe na stazi. Tehnologija i uvidi dobiveni ovim završnim radom mogu se proširiti na autonomna vozila, skladišne robote, sustave nadzora itd. Spoj dubokog učenja i robotike omogućuje inteligentno donošenje odluka, razumijevanje okoliša i sigurnu interakciju s okolnim svijetom.

6.2. NVidia Jetson Nano hardver

NVidia Jetson Nano je bitna komponenta koja pokreće mogućnosti robota ai.Shuttle, omogućujući mu da obrađuje podatke, izvršava složene algoritme i donosi inteligentne odluke u stvarnom vremenu.

Srce Jetson Nano robota je GPU Nvidia Maxwell arhitekture sa 128 CUDA jezgri. Ovaj GPU pruža procesorsku snagu potrebnu za ubrzanje izračuna dubokog učenja, što ga čini prikladnim za pokretanje neuronskih mreža i složenih algoritama. Uz to, Jetson Nano ima četverojezgreni ARM Cortex-A57 procesor (engl. CPU - Central Processing Unit), koji se bavi zadacima obrade opće namjene i pruža ravnotežu između performansi i energetske učinkovitosti.



Slika 10. – GPU Nvidia Maxwell arhitekture (Nvidia Jetson Nano Developer Kit)

Jetson Nano opremljen je raznim sučeljima, uključujući višestruke USB priključke, HDMI izlaz i GPIO (ulaz/izlaz opće namjene) pinove. Ova sučelja omogućuju ai.Shuttle da se poveže sa širokim spektrom senzora, kamera i drugih perifernih uređaja, obogaćujući njegovu sposobnost učinkovite interakcije sa svojim okruženjem. Jetson Nano dolazi s ugrađenim opcijama povezivanja kao što su Gigabit Ethernet i Wi-Fi, omogućujući ai.Shuttleu da komunicira s vanjskim sustavima i uslugama u oblaku, ako je potrebno. Ova povezanost osigurava besprijekoran prijenos podataka i ažuriranja, što olakšava upravljanje i optimizira performanse robota.

6.3. Svrha ai.Shuttle robota

ai.Shuttle robot služi višestrukoj svrsi, obuhvaćajući obrazovne i praktične aspekte. Kao obrazovna platforma, ai.Shuttle nudi idealno okruženje za studente, istraživače i entuzijaste robotike da zarone u svijet umjetne inteligencije i dubokog učenja. Služi kao praktični alat za učenje, omogućujući pojedincima da istražuju i eksperimentiraju s vrhunskim tehnologijama, kao što su Nvidia Jetson Nano i algoritmi dubokog učenja. Kroz proces označavanja skupa podataka i obuke

modela, ai.Shuttle olakšava dublje razumijevanje principa dubokog učenja i njegove primjene u scenarijima stvarnog svijeta.

Osim obrazovne svrhe, ai.Shuttle je projektiran za demonstraciju praktične primjene AI u robotici. Obučavajući robota da samostalno upravlja stazom, izbjegava prepreke i prepoznaje objekte, on prikazuje potencijal dubokog učenja u stvaranju autonomnih sustava koji mogu percipirati svoju okolinu i komunicirati s njom.

7. Stvaranje skupa podataka za ai.Shuttle robot

Stvaranje skupa podataka za robota ai.Shuttle temeljni je i ključan korak na njegovom putu prema postojanju inteligentnog i autonomnog robotskog automobila. Skup podataka služi kao temelj za obuku modela dubokog učenja koji pokreću percepciju i sposobnosti donošenja odluka ai.Shuttlea.

7.1. Proces prikupljanja skupa podataka

Prvi korak bio je uspostaviti vezu između robota i njegovog upravljača, što nam je omogućilo da vozimo robota po stazi. Ova nam je postavka omogućila izravnu kontrolu nad kretanjem ai.Shuttlea, omogućujući nam da njime manevriramo kroz različite scenarije i prikupljamo podatke u stvarnim uvjetima.


```

from jetracer.nvidia_racecar import NvidiaRacecar
import traitlets

car = NvidiaRacecar()

car.throttle_gain = 0.3

car.steering_offset=0
car.steering = 0

print(controller.axes)

(Axis(value=-0.003921568393707275), Axis(value=-0.003921568393707275), Axis(value=-0.003921568393707275), Axis(value=0.0), Axis
(value=0.0), Axis(value=-0.003921568393707275), Axis(value=0.0), Axis(value=0.0), Axis(value=0.0), Axis(value=0.0))

# Na kontroleru trebaju svijetliti dvije crvene lampice, ako ne svijetle pretisnuti gumb "Home"
left_link = traitlets.dlink((controller.axes[0], 'value'), (car, 'steering'), transform=lambda x: -x)
right_link = traitlets.dlink((controller.axes[5], 'value'), (car, 'throttle'), transform=lambda x: -x)

Awesome! Our robot should now respond to our gamepad controller movements. Now we want to view the live video feed from the camera!

print(controller.axes)

(Axis(value=-0.003921568393707275), Axis(value=-0.003921568393707275), Axis(value=-0.003921568393707275), Axis(value=0.0), Axis
(value=0.0), Axis(value=-0.003921568393707275), Axis(value=0.0), Axis(value=0.0), Axis(value=0.0), Axis(value=0.0))

car.throttle_gain = 0.12
car.steering_gain = 0.25

```

Slika 11. – Programski kod za uspostavu veze između ai.Shuttle robota te upravljača

U prvim redovima se uvoze potrebne biblioteke. *NVidiaRacecar* je klasa iz *jetracer* biblioteke, a predstavlja ai.Shuttle robota. *Traitlets* je biblioteka koja se koristi za upravljanje ažuriranjima uživo i povezivanje različitih Python objekata.

Zatim se kreira instanca klase *NVidiaRacecar* i dodjeljuje varijabli *car*. Ova instanca predstavlja robota ai.Shuttle i omogućuje kontrolu njegovih pokreta i parametara.

Parametar *throttle_gain* postavlja osjetljivost robotove kontrole gasa, dok parametri *steering_offset* i *steering* konfiguriraju kut upravljanja robota (0 stupnjeva u ovom slučaju).

Funkcija *dlink* povezuje određene osi kontrolera s odgovarajućim atributima *NVidiaRacecar* instance (*car*). U ovom slučaju, lijevi gumb (*controller.axes[0]*) povezan je s atributom upravljanja, dok je desni gumb (*controller.axes[5]*) povezan s atributom gasa. Lambda funkcije koriste se za transformaciju vrijednosti iz regulatora u odgovarajući raspon za upravljanje i gas.



Videozapis 1. – Testiranje upravljanja ai.Shuttle robota

Ovaj videozapis prikazuje prve trenutke ai.Shuttle robota u vožnji. Na snimci možemo vidjeti trenutke navigacije, ubrzanja te čak i zanošenja kotača robota. Ovime smo prikupljali iskustvo vožnje prije samog prikupljanja podataka na stazi, kako bi stekli što bolji i precizniji osjećaj za vožnju na stazi.

Kako bismo uhvatili vizualni unos, spojili smo kameru na ai.Shuttle. Ova kamera služila je kao "oči" robota, pružajući nam visokokvalitetne slike okoline dok je ai.Shuttle istraživao stazu. Sposobnost kamere da snima slike u stvarnom vremenu bila je ključna jer nam je omogućila prikupljanje podataka za različite uvjete okoline i različite kutove gledanja. Rezoluciju kamere postavili smo na 224x224 piksela.

```
import nanocamera as nano
camera = nano.Camera(flip=0, width=224, height=224, fps=30)
```

Slika 12. – Programski kod za uspostavu veze s kamerom na ai.Shuttle robotu

Dok smo vozili robota po stazi, aktivno smo prikupljali podatke s njegovih senzora i kamere. Proces prikupljanja podataka uključivao je prikupljanje slika staze i njezine okoline, pružajući raznolik raspon vizualnih podataka za učinkovito treniranje modela dubokog učenja. Također uključili smo ključni aspekt koji se zove "ground truthing". To je uključivalo dobivanje predviđenih koordinata položaja ai.Shuttlea na slikama dok smo ih prikupljali. Povezivanjem koordinata s nazivima slika, stvorili smo mapiranje koje je povežalo stvarni položaj robota s odgovarajućim vizualnim podacima. Ukupno smo kreirali više od 2000 fotografija s koordinatama u njihovim nazivima. Sustavnom vožnjom ai.Shuttlea po stazi i simultanim prikupljanjem slika s pripadajućim koordinatama, izgradili smo opsežan skup podataka za obuku. Skup podataka sadržavao je različite scenarije, uključujući različite uvjete osvjetljenja, rasporede staza i prepreke. Ovaj skup podataka postavio je temelje za sljedeće faze projekta ai.Shuttle, gdje bi se prikupljeni podaci komentirali i koristili za obuku modela dubokog učenja. Kako idemo naprijed, prepoznavamo važnost visokokvalitetnih i raznolikih podataka u osiguravanju točnosti modela i sposobnosti generalizacije tijekom njegove autonomne navigacije na stazi.

7.2. Tehnike označavanja podataka

Nakon što smo prikupili znatnu količinu podataka kroz istraživanje staze, započeli smo proces izrade softvera za treniranje modela dubokog učenja. Kako bismo postigli što bolje rezultate, odlučili smo se za kombinaciju tehnika dubokog učenja i regresije. Ovaj hibridni pristup omogućio nam je da iskoristimo snagu algoritama dubokog učenja dok smo uključili regresijske metode za predviđanje kontinuiranih vrijednosti, što se pokazalo vrlo učinkovitim za robota ai.Shuttle. Jedna od najzahtjevnijih komponenti projekta bila je navigacija robota oko prepreka. Ovdje smo naišli na mnoštvo izazova i različitih scenarija koji su od naših modela zahtijevali visok stupanj autonomije. Unatoč složenosti, ostali smo odlučni i usredotočeni, čineći značajne korake u prevladavanju izazova otkrivanja i izbjegavanja prepreka. Kroz obuku i fino podešavanje, naš model je postigao vještine u zaobilazanju prepreka i održavanju glatke navigacije na stazi.

```

import torchvision.transforms as transforms
from xy_dataset import XYDataset

TASK = 'road_following'

CATEGORIES = ['apex']

DATASETS = ['A', 'B']

TRANSFORMS = transforms.Compose([
    transforms.ColorJitter(0.2, 0.2, 0.2, 0.2),
    transforms.Resize((224, 224)),
    transforms.ToTensor(),
    transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
])

datasets = {}
for name in DATASETS:
    datasets[name] = XYDataset(TASK + '_' + name, CATEGORIES, TRANSFORMS, random_hflip=True)
dataset = datasets[DATASETS[0]]

```

Slika 13. – Programski kod kreiranje skupa podataka

Kod počinje uvozom potrebnih biblioteka za transformaciju podataka, upravljanje skupom podataka i vizualizaciju. *torchvision.transforms* pruža različite funkcije transformacije slike, a *XYDataset* je prilagođena klasa skupa podataka za zadatak praćenja ceste. Ostatak uvoza uključuje biblioteku *PyTorch* (*torch*) i povezane module za modele neuronske mreže i interaktivne widgete. Kod stvara skupove podataka za navedeni zadatak praćenja ceste i transformira podatke koristeći definirane transformacije. Stvara instancu *XYDataset* za svaki naziv skupa podataka u *DATASETS*, uključujući nasumično horizontalno okretanje (*random_hflip*) za povećanje podataka. Objekt skupa podataka pohranjuje se u datoteku pod nazivom skupovi podataka, s nazivima skupova podataka kao ključevima i odgovarajućim instancama skupa podataka *XY* kao vrijednostima.

Kod inicijalizira model neuronske mreže koristeći *RESNET-18* arhitekturu iz prethodno obučениh modela *torchvisiona*. Modificira konačni potpuno povezani sloj (*model.fc*) kako bi se prilagodio izlaznim dimenzijama za zadatak praćenja ceste. U ovom slučaju, model predviđa *x* i *y* koordinate za svaku kategoriju, što rezultira *output_dim* 2 puta većim od broja kategorija u skupu podataka.

```

# RESNET 18
model = torchvision.models.resnet18(pretrained=True)
model.fc = torch.nn.Linear(512, output_dim)

```

Slika 14. – Programski kod inicijalizaciju neuronske mreže RESNET-18

Putem koda postavljamo interaktivne widgete kako bi se olakšalo učitavanje i spremanje obučenog modela. Dva gumba, `model_load_button` i `model_save_button`, omogućuju korisnicima učitavanje i spremanje modela. Tekstualni widget, `model_path_widget`, omogućuje korisnicima da unesu putanju na kojoj treba spremiti ili učitati stanje modela.

```

model = model.to(device)

model_save_button = ipywidgets.Button(description='save model')
model_load_button = ipywidgets.Button(description='load model')
model_path_widget = ipywidgets.Text(description='model path', value='road_following_model.pth')

def load_model(c):
    model.load_state_dict(torch.load(model_path_widget.value))
    model_load_button.on_click(load_model)

def save_model(c):
    torch.save(model.state_dict(), model_path_widget.value)
    model_save_button.on_click(save_model)

model_widget = ipywidgets.VBox([
    model_path_widget,
    ipywidgets.HBox([model_load_button, model_save_button])
])

display(model_widget)

```

Slika 15. – Programski kod kreiranje interaktivnih widgeta

U drugom dijelu programskog koda postavljamo broj fotografija u seriji na 8 koje se optimiziraju za kreiranje modela. Dodatno, inicijaliziramo optimizator za model neuronske mreže. U ovom slučaju, *Adam optimizer* se koristi za optimizaciju parametara modela.

```

import time

BATCH_SIZE = 8

optimizer = torch.optim.Adam(model.parameters())
# optimizer = torch.optim.SGD(model.parameters(), lr=1e-3, momentum=0.9)
epochs_widget = ipywidgets.IntText(description='epochs', value=1)
eval_button = ipywidgets.Button(description='evaluate')
train_button = ipywidgets.Button(description='train')
loss_widget = ipywidgets.FloatText(description='loss')
progress_widget = ipywidgets.FloatProgress(min=0.0, max=1.0, description='progress')
def train_eval(is_training):
    global BATCH_SIZE, LEARNING_RATE, MOMENTUM, model, dataset, optimizer, eval_button, train_button, accuracy_widget, loss_widget

```

Slika 16. – Programski kod dijeljenje skupa podataka u skupine veličine 8

Widget *epochs_widget* omogućuje korisnicima određivanje broja epoha obuke. Gumb *eval_button* i gumb *train_button* su gumbi koji pokreću procjenu i obuku modela. *Loss_widget* prikazuje vrijednost gubitka tijekom treninga ili evaluacije. *Progress_widget* pruža vizualni prikaz napretka obuke ili evaluacije. Učitavač podataka dijeli skup podataka u skupine veličine *BATCH_SIZE*, miješajući podatke za bolju izvedbu obuke.



Videzapis 2. – Autonomna vožnja robota ai.Shuttle po stazi bez prepreka

No nažalost korištenjem regresije nismo mogli postići dovoljno veliku brzinu kretanja robota po stazi kojom bi bili zadovoljni, te smo se odlučili posvetiti semantičkoj segmentaciji. Naš je cilj bio stvoriti skup podataka koji je platformi Roboflow omogućio točnu identifikaciju i označavanje objekata na slikama. Kako bismo to postigli, pažljivo smo komponirali fotografije koje su predstavljale različite prizore i scenarije s kojima se susreo tijekom robotovog putovanja. Prikupili smo opsežan niz fotografija i upotrijebili ih kao temelj za daljnju obuku. Označavanje slika zahtijevalo je znatan trud i pozornost na detalje, budući da je svaki piksel trebao biti točno označen kako bi se olakšalo uvježbavanje modela segmentacije. Tijekom označavanja fotografija koristili

smo 2 klase: *road* - koja je označavala stazu kojom vozilo treba ići, te klasu *obstacle* - koja je označavala prepreku koju se treba zaobići.



Slika 17. – Originalna slika snimljena kamerom ai.Shuttle robota

Slika 18. – Klasa „road“ koja označuje stazu

Slika 19. Klasa „obstacle“ koja označuje prepreku

Kod treniranja modela pomoću semantičke segmentacije imali smo puno veću uspješnost te znatno bolje vozne karakteristike iz nekoliko razloga. Semantička segmentacija pruža trening na razini piksela, omogućujući modelu da razlikuje različite klase (npr. cesta, prepreka) na fotografiji. Ove detaljne informacije pomažu robotu ai.Shuttle u preciznom identificiranju granica objekata i razumijevanju njihovog prostornog rasporeda na stazi, povećavajući točnost navigacije. Regresija je ograničena na predviđanje kontinuiranih vrijednosti, kao što su koordinate za određene objekte, što možda nije prikladno za scenarije s više klasa. Semantička segmentacija, s druge strane, omogućuje modelu da rukuje s više klasa istovremeno. Ovo je bitno u zadacima praćenja ceste gdje robot treba prepoznati i razlikovati različite elemente na stazi. Regresijski modeli mogu imati problema s dobrom generalizacijom u okruženjima s različitim uvjetima osvjetljenja, vremenskim promjenama i različitim teksturama cesta. Semantička segmentacija je otpornija na takve varijacije, budući da se usredotočuje na klasifikaciju regija na temelju njihovog vizualnog izgleda, čineći je prilagodljivijom na različitim okruženjima što nam je bilo iznimno važno jer smo onda u mogućnosti našu stazu premjestati na razna okruženja s različitim osvjetljenjem.



Videozapis 3. – Autonomna vožnja robota ai.Shuttle po stazi s izbjegavanjem prepreke

7.3. Izazovi i ograničenja u anotaciji skupa podataka

Tijekom procesa označavanja skupa podataka za zadatak regresije, naišli smo na nekoliko izazova i ograničenja koja su utjecala na izvedbu modela i ukupnu kvalitetu označenih podataka. Jedan značajan izazov s kojim smo se suočili bio je utjecaj različitih uvjeta osvjetljenja na točnost modela. Senzori ai.Shuttle robota, uključujući kameru, bili su osjetljivi na promjene u osvjetljenju, što je često rezultiralo problemom modela da se dobro generalizira u različitim svjetlosnim okruženjima. Na primjer, performanse modela mogu se pogoršati u uvjetima slabog osvjetljenja ili kada je izložen izravnoj sunčevoj svjetlosti.

Još jedno ograničenje proizašlo je iz nedostatka informacija o dubini u označenom skupu podataka. Regresijski model oslanjao se isključivo na 2D koordinate, bez percepcije dubine. Ovo ograničenje robotu je otežavalo točnu procjenu udaljenosti do objekata, što je utjecalo na njegovu sposobnost izbjegavanja sudara ili učinkovite navigacije uskim prostorima.

Nadalje, proces označavanja skupa podataka zahtijevao je značajan ljudski trud i vrijeme. Ručno označavanje podataka o obuci bilo je naporno, posebno za zadatke poput identificiranja granica cesta ili označavanja objekata s visokim stupnjem točnosti. Štoviše, ljudski anotatori mogu uvesti određenu razinu subjektivnosti tijekom procesa anotiranja, što dovodi do potencijalnih nedosljednosti u komentarima. Ove nedosljednosti mogu utjecati na performanse modela i unijeti pogreške u njegova predviđanja.

Unatoč tim izazovima i ograničenjima, nastojali smo ublažiti njihov utjecaj. Koristili smo tehnike povećanja podataka kako bismo poboljšali raznolikost i robusnost skupa podataka, smanjujući osjetljivost modela na varijacije u uvjetima osvjetljenja. Također smo proveli provjere kvalitete anotacija kako bismo smanjili nedosljednosti i poboljšali ukupnu točnost podataka s anotacijama.

7.4. Testiranje modela autonomne vožnje i izbjegavanje prepreka

Nakon što smo dovršili procese označavanja skupa podataka i obuke, sljedeći ključni korak bio je testiranje modela autonomne vožnje na robotu ai.Shuttle. Tijekom ove faze procijenili smo izvedbu modela u scenarijima stvarnog svijeta i ispitali njegovu sposobnost navigacije stazom bez ljudske intervencije. Faza testiranja uključivala je postavljanje uvježbanog modela na ugrađeno računalo robota (NVidia Jetson Nano) i aktiviranje načina autonomne vožnje. Senzori robota, uključujući kameru, prikupljali su podatke iz okoline u stvarnom vremenu dok se kretao stazom.

Jedan od primarnih ciljeva tijekom testiranja bilo je izbjegavanje prepreka. Željeli smo da robot ai.Shuttle identificira prepreke, poput drugih objekata na stazi, i donese odgovarajuće odluke za izbjegavanje sudara. Regresijski model je naučio predviđati koordinate objekata, dok je model semantičke segmentacije omogućio sveobuhvatnije razumijevanje scene, omogućujući robotu da prepozna različite elemente i cestovne strukture. Tijekom testiranja performanse modela procijenjene su na temelju nekoliko ključnih metrika, uključujući točnost, brzinu i sigurnost. Analizirali smo koliko točno je robot pratio cestu i navigirao kroz izazovne scenarije, poput oštih zavoja. Dodatno, procijenili smo odziv modela u otkrivanju i izbjegavanju prepreka u stvarnom vremenu. Sposobnosti izbjegavanja prepreka bile su od najveće važnosti jer su odredile sigurnost i pouzdanost robota ai.Shuttle u autonomnom načinu rada. Kako bismo poboljšali ovaj aspekt,

neprestano smo fino podešavali model i napravili poboljšanja na temelju povratnih informacija dobivenih tijekom testiranja. Ovaj iterativni proces omogućio nam je da riješimo sve potencijalne probleme i učinimo ponašanje robota dosljednijim i inteligentnijim. Iako je model autonomne vožnje pokazao izvanredan napredak u navigaciji stazom i izbjegavanju prepreka, također smo priznali njegova ograničenja. Određene izazovne situacije, kao što su iznenadne i neočekivane prepreke, mogu zahtijevati daljnje usavršavanje u procesu donošenja odluka modela. Osim toga, brzina robota morala je biti prikladno kalibrirana kako bi se osigurala glatka i sigurna navigacija.

8. Zaključak

8.1. Sažetak završnog rada i njegovih doprinosa

Ovaj završni rad predstavlja sveobuhvatan pothvat koji sažima samu bit umjetne inteligencije i dubokog učenja. Središte cijelog završnog rada je ai.Shuttle robot koji ima mogućnosti dubokog učenja te samim time i autonomne vožnje. Pedantno smo se kretali kroz proces prikupljanja visokokvalitetnih podataka za obuku našeg ai.Shuttle robota. Shvatili smo da su podaci temelj na kojem umjetna inteligencija općenito napreduje. Kako smo išli dalje suočili smo se s podtreiranosti te pretreiranosti, koje smo riješili na način da smo pomno izabrali dovoljnu količinu najbitnijih fotografija koje su potrebne za autonomno kretanje ai.Shuttle robota po stazi, te uspješno izbjegavanje prepreka. Uspješno smo trenirali Jetbot za osnovne zadatke kretanja na definiranoj stazi.

8.2. Budućnost projekta

Unutar projekta ai.Shuttle postoje mnogi putevi za daljnji razvoj i usavršavanje. Kako se tehnologija nastavlja razvijati, dosadašnji rad pruža odličan temelj za istraživanje poboljšanja koja mogu podići razine sposobnosti našeg autonomnog robota. To može uključivati poboljšanje brzine

kretanja ai.Shuttle robota po stazi, robusnije i preciznije otkrivanje i izbjegavanje prepreka, kao i detekciju prometnih znakova koje možemo postaviti na stazu. Također spajanje dodatnih senzora koji integriraju podatke iz više izvora kako bi se poboljšala autonomija ai.Shuttlea. Vrlo važno poboljšanje u budućem radu jest sposobnost robota da se prilagodi različitim uvjetima okoline, kao što su različiti uvjeti osvjetljenja, promjene terena ili različite vrste prepreka. To bi uključivalo obuku modela robota na širem rasponu scenarija okoline.

8.3. Predstavljanje ai.Shuttle projekta na danima maturanata

Prvi Dani maturanata trajali su tri dana, a posjetilo ih je više stotina učenika iz više od 15 istarskih škola. Sve sastavnice Sveučilišta, uključujući i Studentski centar imali su svoj punkt za kojim su studenti i profesori zainteresiranim učenicima davali informacije vezane uz studijske programe. U suradnji s profesorima i asistentima prikazali smo ai.Shuttle, te smo zainteresiranim maturantima dali priliku da se sami iskušaju u upravljanju robotom, te ih je to oduševilo.



Slika 20. – Razgovor s maturantima u vezi ai.Shuttle projekta (Unipu Dani maturanata)



Slika 21. – Prikaz oduševljene maturantice tijekom vožnje ai.Shuttle robota

9. Reference

- Ashurst, C., Barocas, S., Campbell, R., i Raji, D. (2022). Disentangling the Components of Ethical Research in Machine Learning. *FACCT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, str. 2057-2068
- Bishop, C., M. (2006). Pattern recognition and machine learning. Springer New York, NY
- Chang, C., Juan, Y., Huang, C. i Chen, H. (2020). Scenario Analysis for Road Following Using JetBot. *2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*
- Ciolacu, M., Haderer, B., Berl, A. i Svasta, P. (2021). Education 4.0: Innovation Learning Lab for AI-Analysis and Concept Proposal. *2021 IEEE 27th International Symposium for Design and Technology in Electronic Packaging (SIITME)*
- Correll, N., Hayes, B., i Grocholsky, B. P. (2016). Introduction to Autonomous Robots.
- Droz, M., Mukherjee, P., i Gevaert, O. (2022). Strategies to Address the Lack of Labeled Data for Supervised Machine Learning Training With Electronic Health Records: Case Study for the Extraction of Symptoms From Clinical Notes. *JMIR Med Inform.*
- Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., i Zisserman, A. (2010). The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision volume 88*
- Goodfellow, I., Bengio, Y., i Courville, A. (2016). Deep Learning. *MIT press*
- Gonzalez, R. C., i Woods, R. E. (2007). Digital Image Processing. PIKS Scientific inside (Vol. 4). Hoboken, New Jersey: Wiley-interscience.
- Guo, Y., Liu, Y., Georgiou, T. i Lew, M. (2017). A review of semantic segmentation using deep neural networks. *International journal of multimedia information retrieval*, 7, 87-93
- Hastie, T., Tibshirani, R., i Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. (Vol. 2, pp. 1-758). New York: springer.
- Justus D., Brennan J., Bonner S. i McGough A. (2017). Predicting the Computational Cost of Deep Learning Models. *In 2018 IEEE international conference on big data (Big Data)*

Kawakura, S., i Shibasaki, R. (2020). Deep Learning-Based Self-Driving Car: JetBot with Nvidia AI Board to Deliver Items at Agricultural Workplace with Object-Finding and Avoidance Functions. *European Journal of Agriculture and Food Sciences*

Krizhevsky, A., Sutskever, I., i Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.

Labelbox, (2022). Data labeling for AI. <https://labelbox.com/guides/data-labeling/>

LeCun, Y., Bengio, Y., i Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444

Nilsson, N. J. (2014). Artificial Intelligence: A New Synthesis. *Morgan Kaufmann Publishers*

Nvidia, (2021). Nvidia Jetson Nano Developer Kit. <https://www.Nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-nano/>

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., i Berg, A. C. (2015). ImageNet large scale visual recognition challenge. *International journal of computer vision*, 115, 211-252.

Russel, S. J., i Norvig, P. (2016). Artificial Intelligence: A Modern Approach. *Upper Saddle River, New Jersey 07458*

Towards AI (2021). Semantic Segmentation: A Complete Guide. <https://towardsai.net/p/l/machine-learning-7>

Towards data science (2018). FCN — Fully Convolutional Network (Semantic Segmentation). <https://towardsdatascience.com/review-fcn-semantic-segmentation-eb8c9b50d2d1>

Unipu, (2022.) Dani maturanata. https://www.unipu.hr/novosti?@=2fcv6#news_107442

University of Freiburg, LMB (2019). U-Net: Convolutional Networks for Biomedical Image Segmentation. <https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/>

Szeliski, R. (2010). Computer Vision: Algorithms and Applications. *Springer Nature*.

Wang, X., Gao J., Long, M. i Wang, J. (2021). Self-Tuning for Data-Efficient Deep Learning. *International Conference on Machine Learning (pp. 10738-10748)*. PMLR.