

SVEUČILIŠTE U ZAGREBU
FAKULTET ORGANIZACIJE I INFORMATIKE
V A R A Ž D I N

Grgo Penava

David Milanović

**Prepoznavanje lica na slici korištenjem
dubokog učenja**

Varaždin, 2023.

SVEUČILIŠTE U ZAGREBU
FAKULTET ORGANIZACIJE I INFORMATIKE
V A R A Ž D I N

Grgo Penava

David Milanović

Kolegij: Biometrijski sustavi

Prepoznavanje lica na slici korištenjem dubokog učenja

Seminarski rad

Mentorica:

Izv. prof. dr. sc. Petra Grd

Varaždin, prosinac 2023

Sadržaj

1. Uvod	1
2. Povijest dubokog učenja.....	2
2.1. Revolucija dubokog učenja	3
3. Duboko učenje.....	4
3.1. Duboko nadzirano učenje	4
3.2. Duboko polu-nadzirano učenje.....	5
3.3. Duboko nenadzirano učenje	5
3.4. Metode dubokog učenja	8
3.4.1. Deep Neural Networks (DNN).....	8
3.4.2. Convolutional Neural Networks (CNN).....	9
3.4.2.1. Convolution Layer (Konvolucijski sloj)	9
3.4.2.2. Pooling Layer (Agregacijski sloj)	10
3.4.2.3. Fully Connected Layer (Potpuno povezani sloj).....	11
3.4.3. Recurrent Neural Network (RNN).....	12
4. Prepoznavanje lica kao biometrijska karakteristika	14
4.1. Biometrija.....	14
4.2. Povijest biometrijskih sustava	15
4.3. Evaluacija performansi	16
4.4. Biometrijski identifikatori	17
4.4.1. Otisak prsta.....	17
4.4.2. Prepoznavanje šarenice oka.....	18
4.4.3. Prepoznavanje glasa	19
4.4.4. Prepoznavanje lica.....	19
5. Implementacija algoritma za prepoznavanje lica na slici	20
5.1. Odabir i priprema baze	20
5.1.1. Labeled Faces in the Wild	21
5.2. Učitavanje i predprocesuiranje slika	21
5.2.1. Predprocesuiranje slika.....	21
5.3. Implementacija konvolucijske neuronske mreže (CNN)	23
5.4. Evaluacija performansi	24
6. Zaključak	25
7. Popis literature.....	26
8. Popis slika.....	28

1. Uvod

U ovom radu fokusirat ćemo se na jedan od inovativnih aspekata suvremenih tehnoloških trendova - prepoznavanje lica na slici korištenjem dubokog učenja. Duboko učenje, kao ključni segment umjetne inteligencije, doživjelo je značajan napredak, otvarajući vrata raznim mogućnostima u različitim sektorima. Upravo u području prepoznavanja lica, duboko učenje donosi revolucionarne promjene i potiče na razvoj sofisticiranih sustava koji su u stanju prepoznati i interpretirati lica s visokom preciznošću. Proučit ćemo temeljne koncepte dubokog učenja, kako se primjenjuju za analizu i prepoznavanje lica, istražujući različite tehnike i modele koji su ključni za postizanje visoke točnosti u ovom području.

Nakon teorijskih koncepata, prikazat ćemo način na koji smo implementirali programski kod za prepoznavanje lica na slici korištenjem dubokog učenja.

2. Povijest dubokog učenja

Duboko učenje, jedan od najuzbudljivijih domena umjetne inteligencije, ima svoje korijene u razvoju neuronskih mreža, a povijest ovog područja obiluje ključnim trenucima i inovacijama. Dvije vrste neuronskih mreža koje su postale temeljem dubokog učenja su „*feedforward Neural Network*“ (FNN) i „*Recurrent Neural Network*“ (RNN). [1]

U 1920-ima, Wilhelm Lenz i Ernst Ising stvorili su i analizirali Ising model, koji se može smatrati pretečom rekurentnih neuronskih mreža. Shun'ichi Amari je 1972. godine ovu arhitekturu učinio adaptivnom, a popularizirao ju je John Hopfield 1982. godine. Rekurentne neuronske mreže postale su ključne za obradu govora i jezika.[1]

Frank Rosenblatt je 1962. godine razvio i istražio osnovne sastavnice današnjih dubokih sustava učenja, predstavljajući višeslojni perceptron (MLP) s tri sloja. Međutim, samo je izlazni sloj imao poveznice za učenje, što ga nije činilo pravim dubokim učenjem. Prvi pravi algoritam za učenje dubokih, feedforward, višeslojnih perceptrona objavljen je 1967. godine od strane Alexey Ivakhnenko i Lapa. [2]

„Perceptron“ je algoritam u strojnom učenju, također poznat kao McCulloch-Pitts neuron. Koristi se za nadgledano učenje binarnih klasifikatora. Binarni klasifikator je funkcija koja može odlučiti pripada li određeni ulaz, prikazan vektorom brojeva, određenoj klasi ili ne. Ovaj algoritam pripada skupini linearnih klasifikatora, što znači da temelji svoje predikcije na linearnoj funkciji predviđanja koja kombinira skup težina s vektorom značajki. Perceptron je jednostavan, ali važan koncept u strojnom učenju, često korišten za osnovne klasifikacijske zadatke. [3]

Shun'ichi Amari je 1967. godine objavio prvi duboki višeslojni model treniran pomoću stohastičkog gradijentnog spusta. U 1987. godini, Matthew Brand izvijestio je o tome kako se široke 12-slojne nelinearne perceptrone mogu potpuno trenirati do kraja kako bi reproducirali logičke funkcije, ali je zaključio da je tehniku činilo nepraktičnom vrijeme obuke na tadašnjoj opremi. [1]

Seppo Linnainmaa je 1970. godine objavio reverzni način automatskog diferencijacije povezanih mreža, poznat kao „*backpropagation*“, što je učinkovita primjena pravila koje je Gottfried Wilhelm Leibniz derivirao 1673. godine. Termin „*back-propagating errors*“ uveden je 1962. godine od strane Rosenblatta, ali nije znao kako to implementirati. Paul Werbos je 1982. godine primijenio backpropagation na MLP-ove, a 1985. godine Rumelhart je objavio eksperimentalnu analizu tehnike. [4]

Arhitekture dubokog učenja za konvolucijske neuronske mreže (CNNs) s konvolucijskim i downsampling slojevima započele su s „Neocognitronom“ koji je 1980. godine predstavio Kunihiko Fukushima. 1969. godine Fukushima je također predstavio ReLU (eng. *rectified linear unit*) aktivacijsku funkciju koja je postala najpopularnija za CNN-ove i duboko učenje općenito. CNN-ovi postali su ključan alat za računalni vid. [1]

Igor Aizenberg i kolege 2000. godine uvode pojam "Duboko učenje" zajednici strojnog učenja, dok je 1988. godine Wei Zhang primijenio algoritam backpropagation na konvolucijsku neuronsku mrežu za prepoznavanje abecede. Yann LeCun primijenio je backpropagation na CNN s ciljem prepoznavanja ručno pisanih poštanskih brojeva 1989. godine. Transformacija učenja u industriji započela je ranih 2000-ih kada su CNN-ovi već obrađivali između 10% i 20% svih čekova napisanih u SAD-u, prema Yann LeCunu [1]

2.1. Revolucija dubokog učenja

Godine 2006., Geoff Hinton, Ruslan Salakhutdinov, Osindero i Teh pokazuju kako se višeslojna feedforward neuronska mreža može učinkovito prethodno trenirati jedan sloj po jedan, koristeći svaki sloj kao nenadzirani ograničeni Boltzmannov stroj, a zatim ga fino podešavajući korištenjem nadziranog backpropagationa. Ova nova tehnika je dobila naziv učenje za duboke vjerodostojne mreže. [1]

U 2011. godini, danNet postiže nadljudsku učinkovitost u natjecanju prepoznavanja vizualnih uzoraka, premašujući tradicionalne metode tri puta. U rujnu 2012., AlexNet osvaja natjecanje ImageNet, pokazujući značajno poboljšanje u prepoznavanju slika u odnosu na pliću obradu podataka. Od tada, duboko učenje postaje sveprisutno u raznim disciplinama, posebno u računalnom vidu i automatskom prepoznavanju govora. [5]

Napredak u hardveru doveo je do obnove interesa za duboko učenje u kasnim 2000-ima. Nvidia je 2009. godine bila uključena u "veliki prasak" dubokog učenja, jer su neuronske mreže trenirane s njihovim grafičkim procesorskim jedinicama (GPU). Posebice, GPU-ovi su se pokazali pogodnima za matematičke operacije u strojnom učenju, ubrzavajući algoritme obuke redoslijedima veličina, smanjujući vrijeme izvođenja od tjedana na dane.

Duboko učenje počelo je nadmašivati druge metode u natjecanjima strojnog učenja od kraja 2000-ih, a 2019. godine Yoshua Bengio, Geoffrey Hinton i Yann LeCun dobili su Turingovu nagradu za ključne prekretnice koje su duboke neuronske mreže učinile ključnim dijelom računalstva. [1]

3. Duboko učenje

Deep Learning (Duboko učenje) predstavlja fascinantan razvoj unutar područja umjetne inteligencije (UI) i strojnog učenja (SU). Nastao iz područja neuronskih mreža, *Deep Learning* je postao ključna grana strojnog učenja od 2006. godine nadalje, izazivajući značajne promjene u različitim domenama primjene. Osnovna ideja dubokog učenja leži u korištenju dubokih arhitektura učenja ili hijerarhijskih pristupa učenju, što omogućava postojanje više slojeva za obradu informacija na nelinearan način. Ovi slojevi se koriste za učenje značajki i klasifikaciju uzoraka. [6]

Duboko učenje ima različite pristupe, među kojima se ističu nadzirano, polu-nadzirano i nenadzirano učenje. [6]

3.1. Duboko nadzirano učenje

Duboko nadzirano učenje koristi označene podatke, gdje agent uči na skupu ulaza i odgovarajućih izlaza. Ovaj pristup koristi različite arhitekture poput Dubokih neuronskih mreža (DNN), Konvolucijskih neuronskih mreža (CNN), Rekurentnih neuronskih mreža (RNN) s *Long Short Term Memory* (LSTM) i *Gated Recurrent Units* (GRU). [6]

U ovom okviru, sustav uči pridruživati izlazne vrijednosti ulaznim podacima. Ključna karakteristika je označenost ulaznog skupa podataka, gdje svaki podatak ima pripadajuću oznaku koja ga kategorizira. [7]

Primjerice, u prepoznavanju slika životinja, svaka slika psa ili mačke označena je određenom kategorijom. Cilj sustava je naučiti funkciju koja omogućuje automatsko kategoriziranje nepoznatih slika pasa ili mačaka. Drugi primjer obuhvaća prepoznavanje spam poruka, gdje sustav nakon učenja prepoznaje spam poruke u novim ulaznim podacima.

Nadzirano učenje podijeljeno je na klasifikaciju i regresiju. Klasifikacija prepoznaje entitete u ulaznim podacima, dok regresija istražuje veze između varijabli. Duboko nadzirano učenje unapređuje ovaj pristup korištenjem dubokih neuronskih mreža poput DNN-a, CNN-a, RNN-a s LSTM-om i GRU-om, omogućujući sustavu učenje složenih obrazaca i značajki s visokom preciznošću u klasifikaciji i predikciji. [7]

3.2. Duboko polu-nadzirano učenje

Duboko polu-nadzirano učenje koristi podatke koji su djelomično označeni, a primjeri uključuju upotrebu Dubokog učenja pojačanja (DRL) i *Generative Adversarial Networks* (GAN). [6]

Duboko pojačano učenje (DRL) je područje strojnog učenja koje sjedinjuje pojačano učenje (RL) i duboko učenje (DL). RL se bavi time kako računalni agent uči donositi odluke putem ispitivanja i pogrešaka, a DRL unosi duboko učenje u ovo rješenje. Time omogućuje agentima donošenje odluka na temelju nestrukturiranih ulaznih podataka, bez potrebe za ručnim inženjeringom prostora stanja. Algoritmi dubokog pojačanog učenja mogu obraditi izuzetno velike ulazne skupove, poput svakog piksela na ekranu video igre, te donositi odluke koje optimiziraju određeni cilj, na primjer, maksimiziranje rezultata igre. [8]

Ovo područje ostvarilo je izvanredne rezultate u različitim područjima, uključujući robotiku, video igre, obradu prirodnog jezika, računalni vid, obrazovanje, promet, financije i zdravstvo.

Generative Adversarial Networks (GAN) predstavljaju inovativan pristup polu-nadziranom učenju putem sukoba između dvije mreže - generatora i diskriminatora. Generator stvara nove podatke, dok diskriminator ocjenjuje njihovu autentičnost. Ovaj sukob potiče oba modela da postanu sve učinkovitiji, rezultirajući boljom sposobnošću generiranja realističnih podataka iz nepotpunih skupova označenih podataka.

GAN predstavljaju značajan napredak u području strojnog učenja i ističu se kao važan okvir za istraživanje generativne umjetne inteligencije. Koncept, inicijalno razvijen od strane Iana Goodfellowa i njegovih kolega u lipnju 2014., temelji se na igri nulte sume gdje se dva neuronska mrežna agenta natječu jedan protiv drugoga.

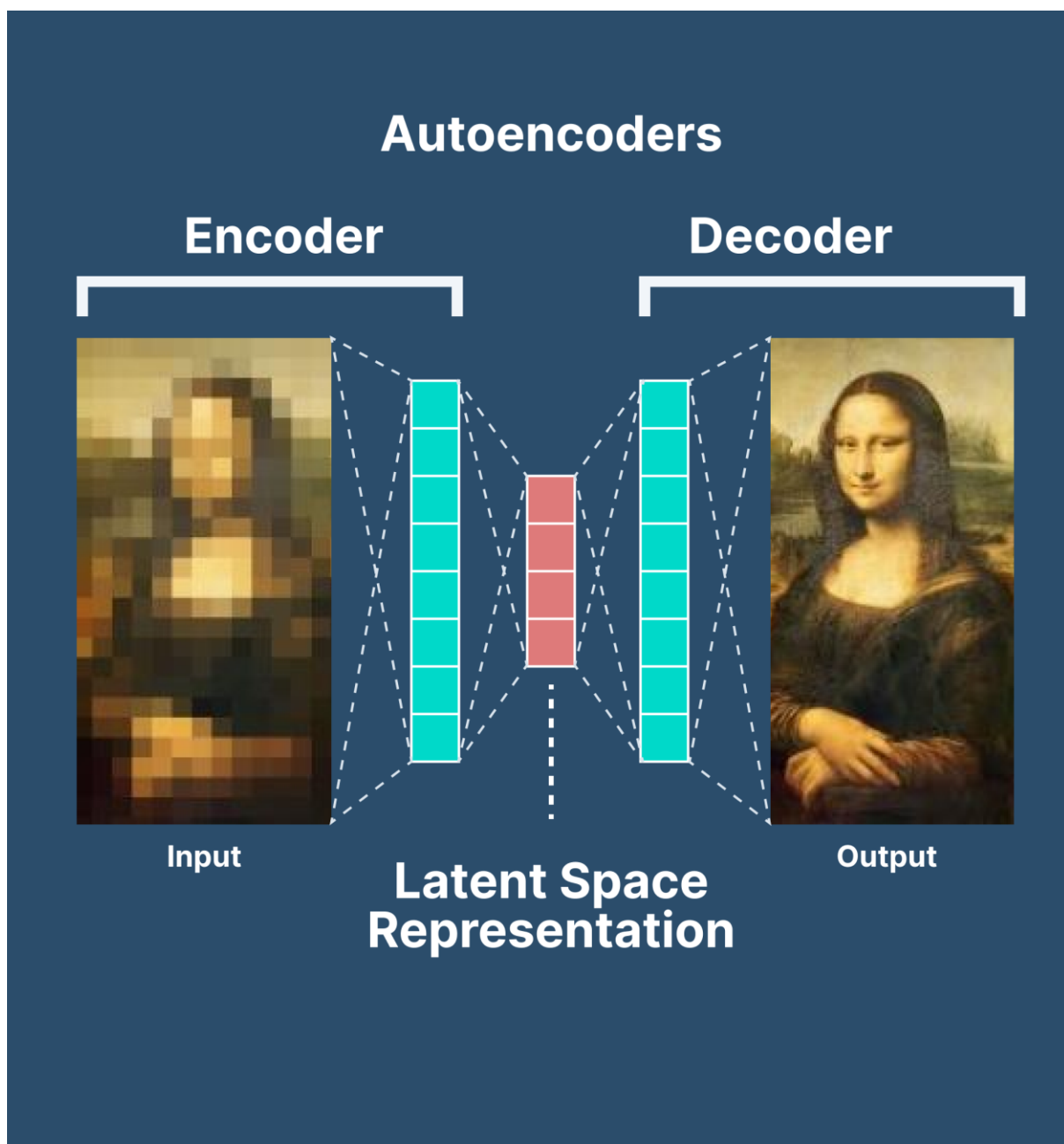
Glavna svrha GAN-ova je učiti iz zadanih skupova podataka te generirati nove podatke s istim statističkim karakteristikama kao i skup za učenje. Na primjer, GAN obučen na fotografijama može stvarati nove fotografije koje barem naoko autentično izgledaju ljudskim promatračima. Iako je prvotno predložen kao oblik generativnog modela za nenadzirano učenje, GAN-ovi su se pokazali korisnima i za polu-nadzirano učenje, potpuno nadzirano učenje i učenje pojačanja.

3.3. Duboko nenadzirano učenje

Duboko nenadzirano učenje ne koristi označene podatke. Model uči interne reprezentacije ili važne značajke kako bi otkrio nepoznate odnose ili strukture unutar ulaznih podataka. Ovdje se često koriste tehnike kao što su grupiranje, smanjenje dimenzionalnosti i generativne tehnike, uključujući *Auto-Encoders* (AE), *Restricted Boltzmann Machines* (RBM) i GAN. [6]

Autoenkoderi su izuzetno korisni u području nenadziranog strojnog učenja. Mogu se koristiti za komprimiranje podataka i smanjenje dimenzionalnosti. Autoenkoderi rekonstruiraju naš izvorni ulaz dajući samo komprimiranu verziju.

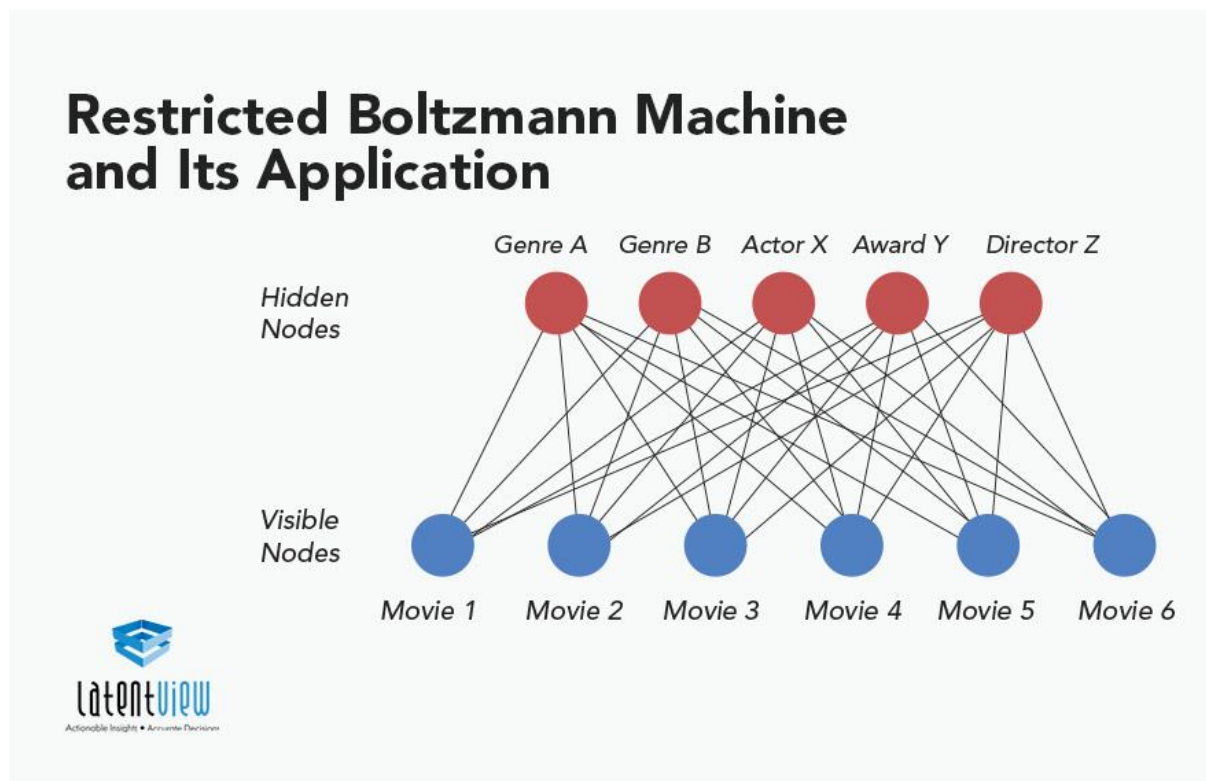
Ako netko treba izvorne podatke, može ih rekonstruirati iz komprimiranih podataka pomoću autoenkodera. Oni nude snažan alat za efikasno predstavljanje podataka i očuvanje bitnih informacija, čineći ih dragocjenim instrumentom u analizi i obradi podataka. [9]



Slika 1. Autoenkoderi [10]

Restricted Boltzmann Machine (RBM) je vrsta umjetne neuronske mreže koja se koristi za nenadzirano učenje. To je vrsta generativnog modela koji je sposoban naučiti vjerojatnosnu distribuciju nad skupom ulaznih podataka. [11]

RBM je predstavljen sredinom 2000-ih godina od strane Hintona i Salakhutdinova kao način rješavanja problema nenadziranog učenja. To je vrsta neuronske mreže koja se sastoji od dvije razine neurona - vidljivog sloja i skrivenog sloja. Vidljivi sloj predstavlja ulazne podatke, dok skriveni sloj predstavlja skup značajki koje mreža nauči. RBM se naziva "restricted" (ograničen) jer veze između neurona u istom sloju nisu dopuštene. Drugim riječima, svaki neuron u vidljivom sloju povezan je samo s neuronima u skrivenom sloju, i obrnuto. To omogućava RBM-u da nauči komprimirano predstavljanje ulaznih podataka smanjenjem dimenzionalnosti ulaza. RBM se trenira korištenjem procesa nazvanog "contrastive divergence," što je varijanta algoritma stohastičkog gradijentnog spusta. Tijekom treninga, mreža prilagođava težine veza između neurona kako bi maksimizirala vjerojatnost trening podataka. Nakon što je RBM treniran, može se koristiti za generiranje novih uzoraka. RBM se primjenjuje u različitim područjima, uključujući računalni vid, obradu prirodnog jezika i prepoznavanje govora. Također se koristi u kombinaciji s drugim arhitekturama neuronskih mreža, poput dubokih vjerodajnih mreža i dubokih neuronskih mreža, radi poboljšanja njihove učinkovitosti. [11]



Slika 2. Restricted Boltzmann Machine [12]

Duboko učenje igra ključnu ulogu u različitim disciplinama, posebno u računalnom vidu i prepoznavanju govora. Njegova sveprisutnost i svestranost čine ga iznimno važnim područjem unutar razvoja umjetne inteligencije.

3.4. Metode dubokog učenja

Duboko učenje obuhvaća različite metode, a neke od najpoznatijih metoda dubokog učenja su:

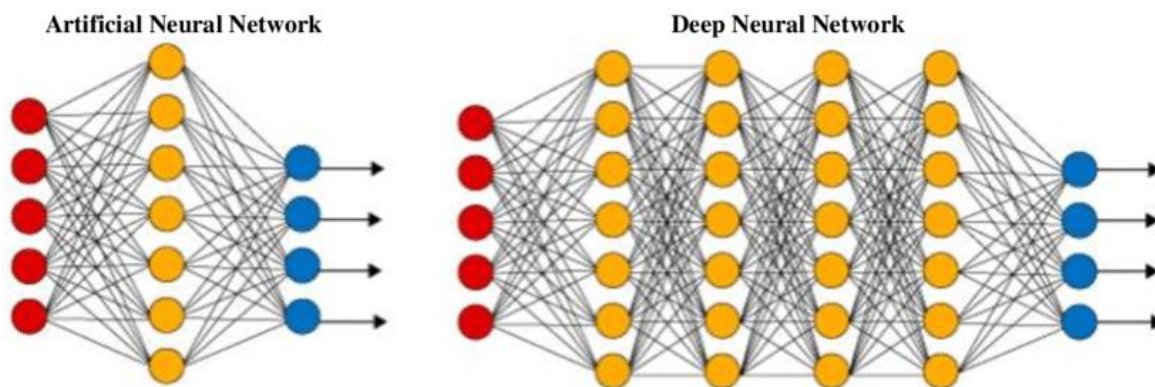
- *Deep Neural Networks* (DNN)
- *Convolutional Neural Networks* (CNN)
- *Recurrent Neural Networks* (RNN)
- *Long Short-Term Memory* (LSTM)
- *Generative Adversarial Networks* (GAN)
- *Probabilistic Neural Networks* (PNN)

3.4.1. Deep Neural Networks (DNN)

Duboka neuronska mreža (DNN) predstavlja sofisticiranu granu umjetne neuronske mreže (ANN) koja se odlikuje višeslojnom strukturom između ulaznog i izlaznog sloja. Ova vrsta mreže sastoji se od ključnih komponenata poput neurona, sinapsi, težina, pristranosti i funkcija, čime oponaša osnovne procese ljudskog mozga. Bitno je naglasiti da se ove komponente zajedno ponašaju na način koji omogućava treniranje mreže, što ju čini prilagodljivom. [1]

Jedan od ključnih primjera dubokih neuronskih mreža jest njihova sposobnost prepoznavanja kompleksnih uzoraka, poput razlikovanja pasmina pasa na fotografijama. U ovom procesu, mreža prolazi kroz sliku i izračunava vjerojatnost da li je pas određene pasmine. Krajnji korisnik ima mogućnost pregledati rezultate i odabrati koje vjerojatnosti mreža treba prikazati, što omogućuje prilagodbu prema specifičnim zahtjevima. [1]

Duboke arhitekture, poput DNN-a, proizvode kompozicijske modele gdje se objekt prikazuje slojevito, što omogućuje kompleksno modeliranje podataka koristeći manje jedinica u usporedbi s plitkim mrežama. To je posebno korisno u situacijama gdje je potrebno modelirati nelinearne odnose i obraditi složene skupove podataka. [1]



Slika 3. ANN vs DNN [13]

3.4.2. Convolutional Neural Networks (CNN)

Konvolucijska neuronska mreža (CNN) predstavlja vrstu feed-forward neuronske mreže koja samostalno usvaja inženjering značajki putem optimizacije filtara (jezgara). Ova tehnologija sprječava nestanak i eksploziranje gradijenata, problema koji su bili prisutni tijekom backpropagation-a u ranijim neuronskim mrežama, koristeći regulirane težine na manje veza. Na primjer, dok bi za svaki neuron u potpuno povezanom sloju bilo potrebno 10,000 težina za obradu slike veličine 100×100 piksela, primjenom kaskadnih konvolucijskih jezgara potrebno je samo 25 neurona za obradu pločica veličine 5×5 . Višeslojne značajke izvlače se iz širih kontekstualnih prozora u usporedbi s nižim slojevima. [14]

Primjene konvolucijskih neuronskih mreža su raznovrsne, uključujući prepoznavanje slika i videa, klasifikaciju slika, segmentaciju slika, analizu medicinskih slika, obradu prirodnog jezika, te analizu financijskih vremenskih nizova. [14]

U usporedbi s tradicionalnim algoritmima klasifikacije slika, CNN-ovi zahtijevaju relativno malo predprocesiranja. Ova mreža uči optimizirati filtre (jezgre) putem automatiziranog učenja, dok su u tradicionalnim algoritmima ovi filtri ručno oblikovani. Neovisnost o prethodnom znanju i ljudskom posredovanju u izdvajanju značajki predstavlja ključnu prednost ovog pristupa. [14] CNN se sastoji od tri sloja:

- Konvolucijski sloj (eng. *Convolution Layer*)
- Agregacijski sloj (eng. *Pooling Layer*)
- Potpuno povezani sloj (eng. *Fully Connected Layer*)

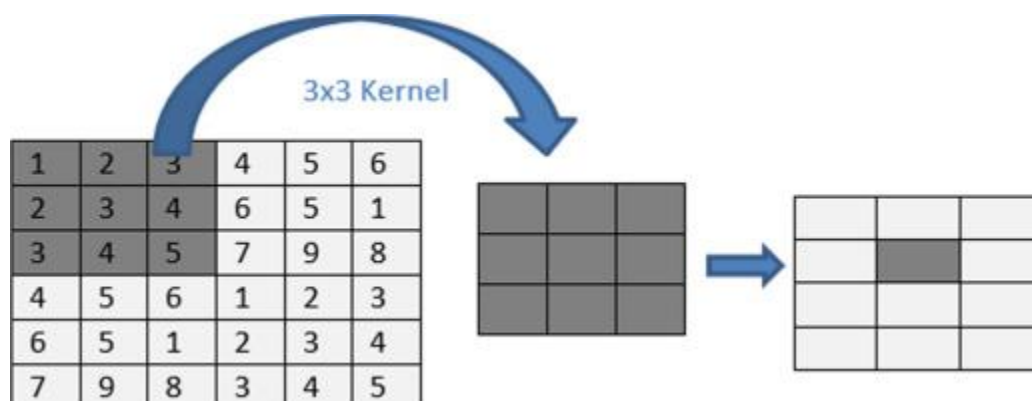
3.4.2.1. Convolution Layer (Konvolucijski sloj)

Konvolucijski sloj je ključna građevna jedinica CNN-a i nosi glavni dio računalnog opterećenja mreže. Ovaj sloj izvodi skalarni produkt između dva skupa matrica, gdje jedna

matrica predstavlja skup parametara (za učenje) poznatih kao jezgra (kernel), a druga matrica predstavlja ograničeni dio receptivnog polja.[15]

Tijekom prolaska prema naprijed, jezgra klizi duž visine i širine slike, stvarajući slikovno prikazivanje tog receptivnog područja. Ovo proizvodi dvodimenzionalno prikazivanje slike poznato kao aktivacijska mapa koja daje odgovor jezgra na svakom prostornom položaju slike. [15]

Konvolucija se oslanja na tri važne ideje: rijetku interakciju, dijeljenje parametara i prikazivanje. Rijetka interakcija postiže se činjenicom da je jezgra prostorno manja od slike, što znači da je potrebno pohraniti manje parametara, smanjujući memoriju potrebnu za model i poboljšavajući statističku učinkovitost. [15]



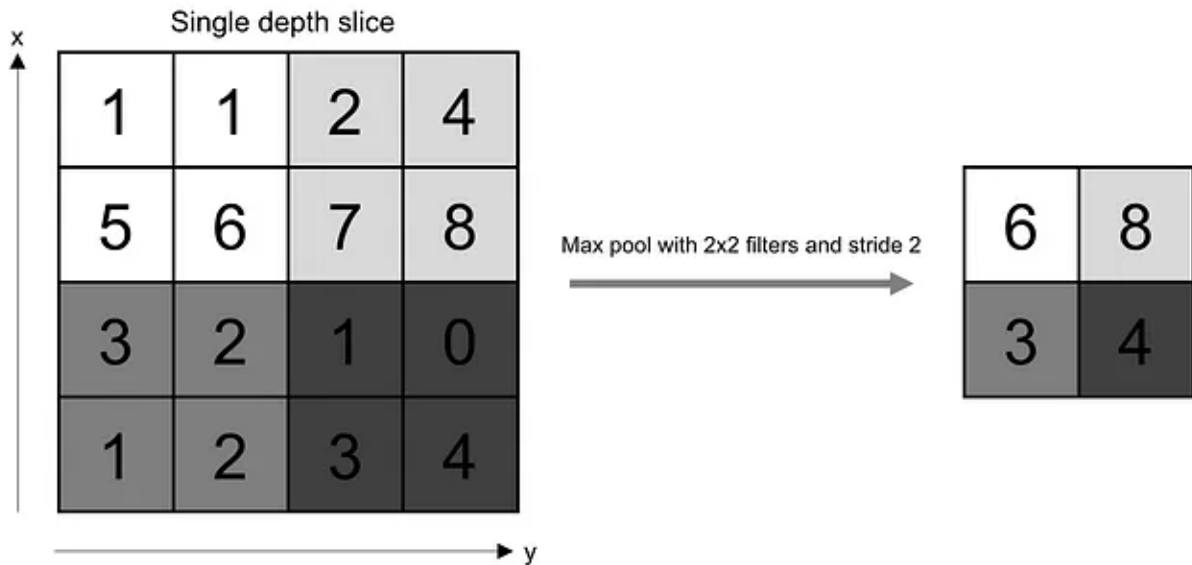
Slika 4. Convolution Layer [16]

3.4.2.2. Pooling Layer (Agregacijski sloj)

Agregacijski sloj jedna su od temeljnih građevnih jedinica konvolucijskih neuronskih mreža. Dok konvolucijski slojevi izvlače značajke iz slika, slojevi agregacije konsolidiraju značajke koje je CNN naučio. Njihova svrha je postupno smanjiti prostornu dimenziju reprezentacije kako bi se minimizirao broj parametara i računanja u mreži. [17]

Sloj agregacije zamjenjuje izlaz mreže na određenim lokacijama izvodeći sažetak statistike bliskih izlaza. To pomaže u smanjenju prostorne veličine reprezentacije, što smanjuje potrebnu količinu računanja. Operacija sažimanja obrađuje se na svakom sloju reprezentacije pojedinačno. [15]

Postoje različite funkcije sažimanja, poput prosjeka pravokutnog susjedstva, L2 norme pravokutnog susjedstva i ponderiranog prosjeka na temelju udaljenosti od središnjeg piksela. Međutim, najpopularniji postupak je maksimalno sažimanje, koje prijavljuje maksimalni izlaz iz susjedstva. [15]

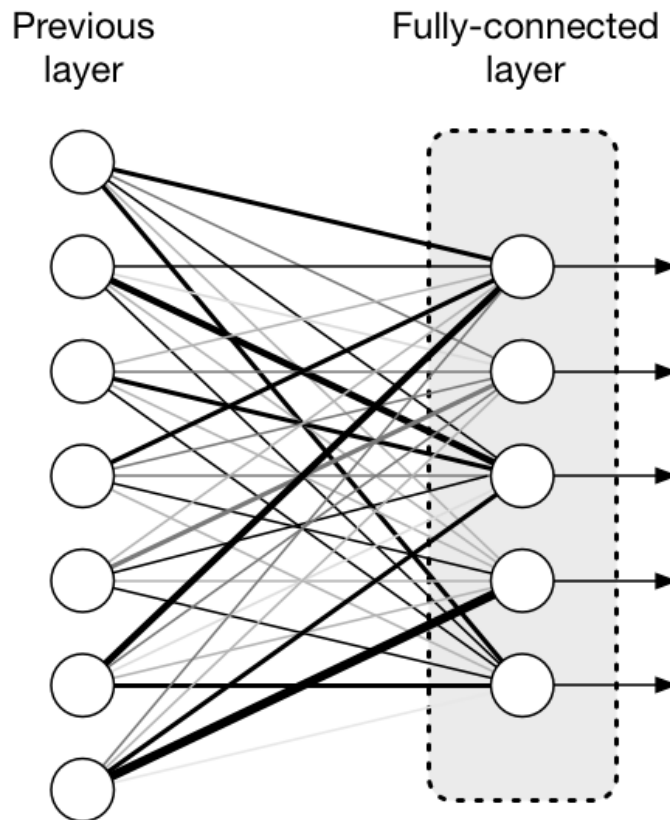


Slika 5. Pooling Layer [15]

3.4.2.3. Fully Connected Layer (Potpuno povezani sloj)

Potpuno povezani sloj (Fully Connected Layer) ključan je dio konvolucijskih neuronskih mreža (CNN) i igra značajnu ulogu u klasifikaciji i obradi podataka. Nakon što su ulazni podaci prošli kroz niz procesiranja kako bi izvukli relevantne značajke, potpuno povezani sloj preuzima zadnji izlazni skup značajki i obavlja konačnu klasifikaciju.

U potpuno povezanom sloju, svaki neuron ima veze sa svim aktivacijama iz prethodnog sloja. Ova potpuna povezanost omogućava modelu da nauči složene nelinearne kombinacije značajki. [18]



Slika 6. Fully Connected Layer [19]

3.4.3. Recurrent Neural Network (RNN)

Rekurentne neuronske mreže (RNN) sposobne su učiti značajke i dugoročne ovisnosti iz sekvencijalnih i vremenskih podataka. Rekurentne neuronske mreže imaju hrpu nelinearnih jedinica gdje barem jedna veza između jedinica tvori usmjereni ciklus. Dobro obučeni RNN može modelirati bilo koji dinamički sustav; međutim, obuka RNN-a često je pogođena problemima u učenju dugoročnih ovisnosti. [20]

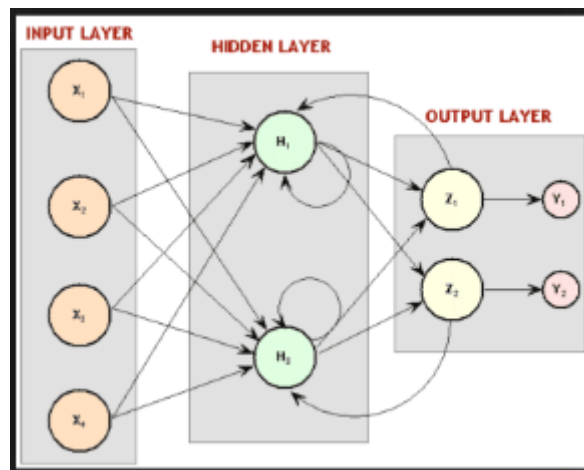
Osnovna značajka rekurzivnih neuronskih mreža (RNN) jest ta da mreža sadrži barem jednu povratnu vezu, omogućujući aktivacijama da kruže u petlji. To omogućuje mrežama izvođenje vremenske obrade i učenje sekvenci, primjerice, izvođenje prepoznavanja/reprodukcije sekvenci ili temporalne asocijacije/predikcije. [21]

Arhitekture rekurzivnih neuronskih mreža mogu imati mnogo različitih oblika. Jedan uobičajeni tip sastoji se od standardnog višeslojnog perceptrona (MLP) s dodanim petljama. Ove mreže mogu iskoristiti snažne sposobnosti nelinearnog mapiranja MLP-a i također imati neki oblik memorije. Druge imaju uniformnije strukture, potencijalno s svakim neuronu povezanim s ostalima, te mogu imati stohastičke aktivacijske funkcije. [21]

Za jednostavne arhitekture i determinističke aktivacijske funkcije učenje se može postići korištenjem sličnih postupaka gradijentnog spusta kao i kod algoritma propagacije unatrag za neuronske mreže s progresijom unaprijed. Kada su aktivacije stohastičke, pristupi simuliranog kaljenja mogu biti prikladniji. [21]

Postoje razne varijante RNN-a, a neke od njih su [22]:

- *Fully recurrent (FRNN)* - Povezuju izlaze svih neurona s ulazima svih neurona, predstavljajući najopćenitiju topologiju neuronske mreže
- *Elman networks and Jordan networks* - Elmanove mreže održavaju stanje za zadatke poput predviđanja sekvenci, dok Jordanove mreže imaju kontekstualne jedinice koje se hrane izlaznim slojem
- *Hopfield* - Pruža robusnu memoriju otpornu na promjene u vezama uz korištenje Hebbianovog učenja
- *Bidirectional associative memory* - Koristi matricu i njezinu transpoziciju za prolaz informacija, omogućujući dvosmjerno dohvaćanje asocijacija
- *Recursive* - Stvorene su primjenom istog skupa težina rekurzivno preko diferencijalne grafičke strukture, omogućujući procesiranje distribuiranih reprezentacija strukture
- *Neural history compressor* - Nesupervizirani stog od RNN-ova koji uči predviđati sljedeći ulaz i komprimira informacije na višim razinama
- *Long short-term memory* - Duboki sustav koji izbjegava nestanak gradijenta i sposoban je za učenje dugoročnih ovisnosti



Slika 7. Fully recurrent neural network [23]

4. Prepoznavanje lica kao biometrijska karakteristika

4.1. Biometrija

Izraz "Biometrija" koristi se od početka 20. stoljeća, a odnosi se na područje razvoja statističkih i matematičkih metoda primjenjivih na probleme analize podataka u biološkim znanostima. Nedavno se izraz "Biometrija" također koristi za označavanje novog područja tehnologije posvećenog identifikaciji pojedinaca na temelju njihovih bioloških obilježja, poput onih temeljenih na skeniranju mrežnice, uzorku šarenice, otiscima prstiju ili prepoznavanju lica. [24]

U današnjem svijetu različite aplikacije zahtijevaju pouzdane i sigurne metode autentifikacije kako bi potvrdile identitet osobe koja traži njihovu uslugu. Primjeri takvih aplikacija uključuju siguran pristup zgradama, računalnim sustavima, prijenosnim računalima, mobilnim telefonima, memorijskim uređajima poput USB stickova i mnoge druge. Osim toga, moguće je utvrditi identitet na temelju tko ste, a ne na temelju onoga što posjedujete (npr. identifikacijske kartice) ili čega se sjećate (npr. lozinke). [24]

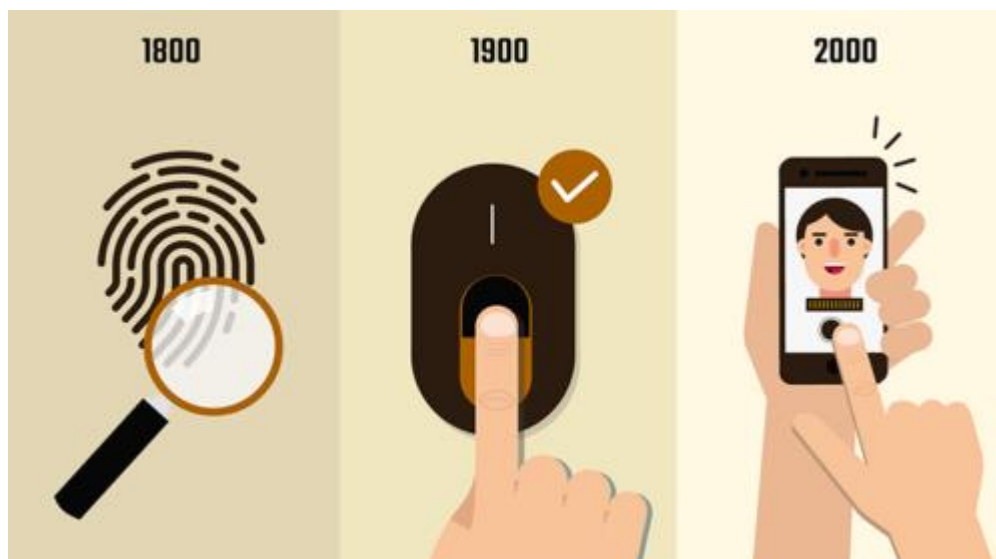
Biometrijski identifikatori su prepoznatljive, mjerljive karakteristike koje se koriste za označavanje i opisivanje pojedinaca. Biometrijski identifikatori često se kategoriziraju kao fiziološke karakteristike koje su povezane s oblikom tijela. Primjeri uključuju, ali nisu ograničeni na otiske prstiju, prepoznavanje lica, otiske dlana, geometriju ruke, mrežnicu, glas, hod i sl. [25]

Ponašajne karakteristike povezane su s uzorkom ponašanja osobe, uključujući, ali ne ograničavajući se na pokrete miša, ritam tipkanja, hod, potpis, ponašajno profiliranje i glas. [25]

Tradicijski načini kontrole pristupa uključuju sustave identifikacije temeljene na tokenima, poput vozačke dozvole ili putovnice, te sustave identifikacije temeljene na znanju, poput lozinke ili osobnog identifikacijskog broja. Budući da su biometrijski identifikatori jedinstveni za pojedince, pouzdaniji su u potvrđivanju identiteta u usporedbi s tokenima i metodama temeljenim na znanju; međutim, prikupljanje biometrijskih identifikatora postavlja pitanja o privatnosti u vezi s konačnom upotrebom tih informacija. [25]

4.2. Povijest biometrijskih sustava

Povijest biometrijskih sustava započinje još 1885. godine kada je Juan Vucetich počeo sakupljati otiske prstiju kriminalaca u Argentini. Josh Ellenbogen i Nitzan Lebovic tvrde da su biometrijski sustavi proizašli iz sustava identifikacije kriminalnih aktivnosti koje su razvili Alphonse Bertillon (1853–1914) te iz teorija otisaka prstiju i fizionomije Frances Galtona. Lebovic naglašava da je Galtonovo djelo potaknulo primjenu matematičkih modela na otiske prstiju, frenologiju i karakteristike lica, čime je postavljena osnova za "apsolutnu identifikaciju" i "ključ za uključivanje i isključivanje" populacija. S tim u vezi, biometrijski sustav opisuje kao apsolutno političko oružje današnjice i oblik "mekog nadzora". David Lyon ukazuje na to da su biometrijski sustavi posljednjih dvadeset godina prodrli na građansko tržište i zamaglili granice između državnog nadzora i privatnog korporativnog nadzora. Kelly A. Gates ističe da je 11. rujna bio ključni dan u kulturnom jeziku, gdje je "automatizirano prepoznavanje lica postalo tehnologija za nacionalnu sigurnost. [25]



Slika 8. Povijest biometrijskih sustava [26]

4.3. Evaluacija performansi

Izvedba biometrijskog autentikacijskog sustava može se procijeniti putem Stopa lažnih odbijanja (FRR) ili Stope lažnih prihvatanja (FAR), a mogu se izračunati na sljedeći način: [24]

$$FRR = \frac{\text{broj lažnih odbijanja}}{\text{ukupan broj pristupa klijenta}}$$

$$FAR = \frac{\text{broj lažnih prihvatanja}}{\text{ukupan broj pristupa klijenta}}$$

Idealan biometrijski autentikacijski sustav imao bi $FRR = 0$ i $FAR = 0$, što je vrlo teško postići u stvarnosti. Zanimljivo je napomenuti da se bilo koja od vrijednosti FRR ili FAR može svesti na proizvoljno malu vrijednost, ali uz nedostatak povećanja druge vrijednosti. Vrijednost TER može se računati na sljedeći način: [24]

$$TER = \frac{\text{broj lažnih prihvatanja} + \text{broj lažnih odbijanja}}{\text{ukupan broj pristupa}}$$

Važno je naglasiti da ove mjere mogu biti značajno pristrane prema jednoj od vrsta pogrešaka (FAR ili FRR), ovisno o broju pristupa korištenih za dobivanje tih pogrešaka. To znači da će TER uvijek biti bliži vrsti pogreške koja je dobivena s najvećim brojem pristupa. Ukupna izvedba biometrijskog autentikacijskog sustava ne bi se trebala mjeriti prema TER-u, već prema Receiver Operation Characteristic (ROC) krivulji, koja prikazuje FAR kao funkciju FRR. [24]

4.4. Biometrijski identifikatori

Biometrijski identifikatori koriste se za autentifikaciju pojedinaca putem jedinstvenih fizičkih ili ponašajnih karakteristika. Oni omogućuju visoku razinu sigurnosti i jednostavnosti u raznim sektorima života. Umjesto tradicionalnih metoda temeljenih na znanju (npr. Upisivanje lozinke), biometrija koristi jedinstvene osobine pojedinaca kako bi ih prepoznala i verificirala. Jedni od najpoznatijih identifikatora su: otisak prsta, prepoznavanje lica, prepoznavanje šarenice oka, prepoznavanje glasa i prepoznavanje hoda.

4.4.1. Otisak prsta

Otisak prsta najstarija je i jedna od najuspješnijih biometrijskih tehnika koja se koristi u raznim područjima. Ova metoda temelji se na jedinstvenim i nepromjenjivim obrascima koji se nalaze na vrhovima prstiju, čineći je fizičkim biometrijskim sustavom. Unikatnost otiska prsta određena je uzorcima brda i dolina, kao i točkama na površini prsta, poznatim kao minutije. Minutije su lokalne karakteristike brda koje se javljaju na razdjelnicama brda ili na krajevima brda. [24]

Otisci prstiju pružaju pouzdanu metodu identifikacije jer osnovne karakteristike otisaka ostaju nepromijenjene tijekom vremena. Ove karakteristike uključuju uzorke brda, dolina i minutija na površini prsta. Otisci prstiju široko su prepoznati po svojoj jedinstvenosti i postali su neizostavan dio forenzičkih laboratorija i odjela za identifikaciju diljem svijeta, te se već gotovo stoljeće prihvaćaju u pravosudnim postupcima. [24]

Dok su otisci prstiju redovito korišteni u forenzičkim laboratorijima, njihova primjena u civilnom području postala je relevantnija od 1980-ih zbog povećane točnosti i smanjenja cijena uređaja za otiske prstiju. Primjeri upotrebe uređaja za otiske prstiju u civilnom području uključuju omogućavanje prijave temeljenih na otiscima prstiju. [24]

Tehnologija je jednostavna za implementaciju, ali može biti skupa. Performanse mogu biti smanjene ako je sam otisak prsta znojav ili suh ili ima neki defekt na koži. Problemi uključuju i temperaturu okoline, prljavštinu, znoj, vlažnost ili suhoću. Prst se može ne postaviti uvijek pod istim kutom, a može se primijeniti i veći pritisak nego uobičajeno, čime dolazi do izobličenja. Međutim, otisci prstiju mogu se uzeti za manje od dvije minute, što pridonosi jednostavnosti. U otiscima prstiju se uzima otprilike 35 točaka. [27]



Slika 9. Otisak prsta [28]

4.4.2. Prepoznavanje šarenice oka

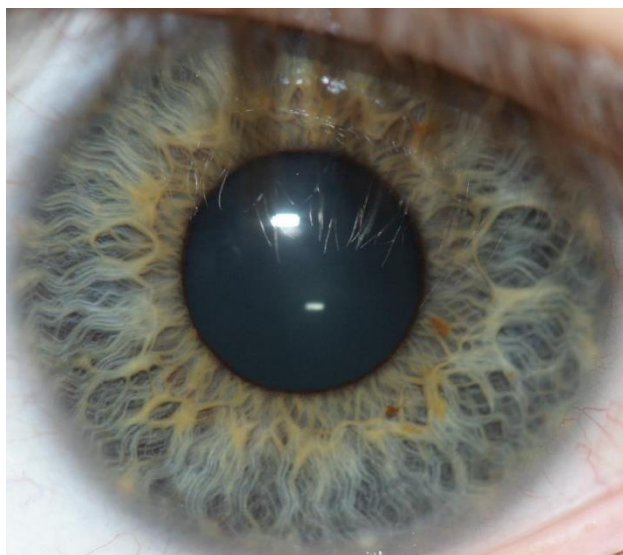
Biometrijska tehnologija skeniranja šarenice koristi jedinstvene karakteristike i značajke ljudskih šarenica, koje ostaju nepromijenjene tijekom čitavog života pojedinca, kako bi potvrdila identitet osobe. Šarenica je područje oka gdje obojeni krug, obično smeđi, zeleni, sivi ili plavi, okružuje tamnu zjenicu oka. Šarenica je dobro zaštićena zbog ljudske anatomije, pa su ozljede rijetke. [24]

Od 1997. godine koriste se tehnike identifikacije putem šarenice. Nijedna dva pojedinca nemaju istu šarenicu. Postoje embriološki faktori koji utječu na razvoj ljudske šarenice, tako da čak i identični blizanci mogu biti razlikovani. Jedinstveni uzorci u šarenici određeni su DNK subjekta, koji se razlikuju čak i između lijevog i desnog oka iste osobe, te ostaju konstantni tijekom čitavog života osobe. Šarenica se stoga smatra inherentnom jedinstvenom karakteristikom. [27]

Tipično, postupak skeniranja šarenica počinje fotografijom koja se snima pomoću posebne kamere blizu ispitanika. Korisnik treba biti udaljen najviše oko 1 metar od čitača. Kamera koristi infracrveni senzor za osvjetljavanje oka i snimanje fotografije visoke razlučivosti. [24]

Unutarnji rub šarenice određuje se pomoću algoritma za skeniranje šarenica koji mapira različite obrasce i karakteristike šarenica. Sustavi koji koriste biometrijsku tehnologiju šarenica čak funkcioniraju s naočalama, čineći ovu tehnologiju jednom od rijetkih biometrijskih tehnologija koja dobro djeluje u identifikacijskom načinu rada. [24]

Oblici šarenica su iznimno složeni, noseći zadivljujuću količinu informacija i imaju preko 200 jedinstvenih točaka. Te točke kategorizirane su prema tkivu, što daje dojam dijeljenja šarenice u radijalnom smjeru, prstenovima, brazdama, pjegama i krunom. [24]



Slika 10. Prepoznavanje šarenice oka [29]

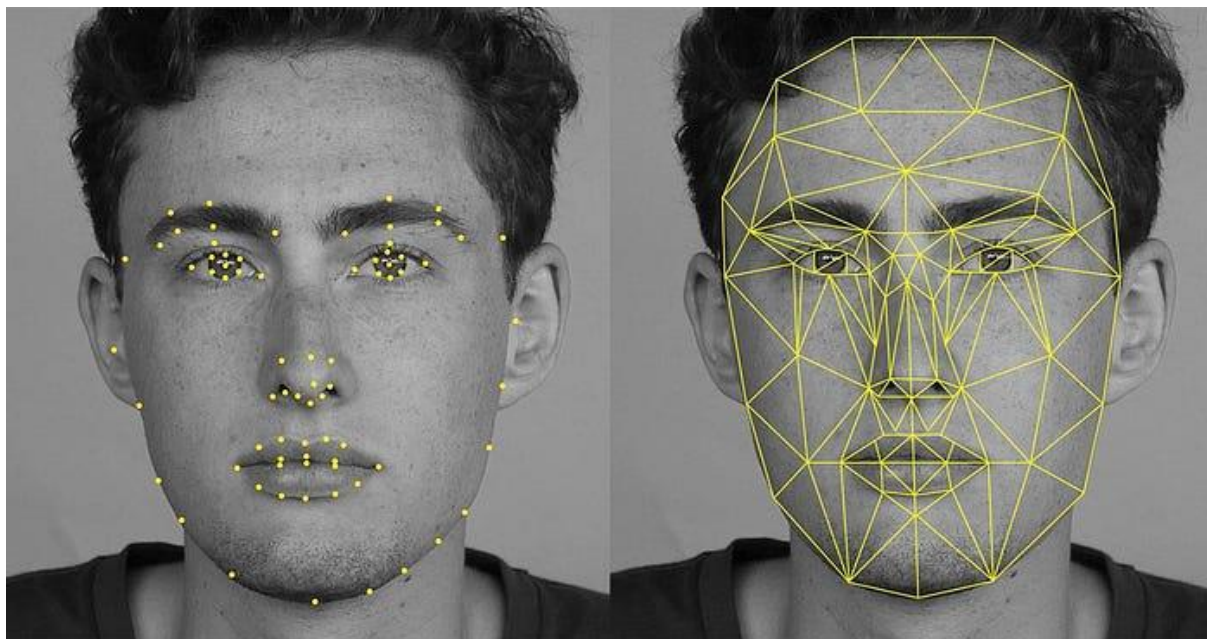
4.4.3.Prepoznavanje glasa

Svaki glas jedinstven je po tonu, ritmu i naglascima za različite glasovne jedinice. Obično možemo prepoznati osobu s druge strane telefona čim ta osoba počne govoriti "Bok". Identifikacija govora omogućuje prepoznavanje određenog govornika. Prepoznavanje govora analizira što osoba stvarno govori. To je jedina biometrija koja koristi akustične podatke i može stoga raditi s javnim telefonom. Mogu se uhvatiti uzorci glasa čak i pomoću pametnih telefona ili osobnih digitalnih pomoćnika (PDA uređaja) sve dok postoji mikrofoni. [27]

4.4.4.Prepoznavanje lica

Prepoznavanje lica koristi one karakteristike lica koje se ne mijenjaju lako. Analiziraju se različite značajke lica, poput očiju, jagodičnih kostiju i usana. Obično se bilježi oko 20 točaka, a slike se obično snimaju video kamerama, što omogućuje upotrebu ove biometrije u svrhu nadzora. Kada se pogledaju performanse prepoznavanja lica, korištene su različite velike baze podataka, ali preciznost nije tako dobra kao kod prepoznavanja šarenice ili otiska prsta. Također treba imati na umu da prepoznavanje lica utječe na svjetlo, položaj, veličinu pozadine, buku, prekrivanje lica (frizuram i šminkom) i izraze lica. Postoje zabrinutosti da bi prepoznavanje lica trebalo biti provedeno pod kontroliranim uvjetima kako bi bilo uspješno. Ako osoba nije okrenuta ravno prema kameri, nosi naočale ili kapu ili dođe do promjene osvjetljenja na subjektu, tada je teško identificirati osobu. Utjecaj tih čimbenika treba smanjiti kako bi se postigle najbolje performanse. Prepoznavanje lica je neinvazivno i stoga lako prihvatljivo za korisnike. Ova tehnologija manje je nametljiva, tj. ne očekuje da osoba "pritisne

lice uz čisto staklo", primjerice. Umjesto toga, tehnologija se može koristiti s udaljenosti za prepoznavanje lica. [27]



Slika 11. Prepoznavanje lica [30]

5. Implementacija algoritma za prepoznavanje lica na slici

U praktičnom dijelu rada naglasak je bio na kreiranju modela za prepoznavanje lica na slici koristeći duboko učenje. Kako bismo riješili problem potrebno je odabrati bazu koja će sadržavati fotografije lica, predprocesuirati slike, implementirati odabranu metodu dubokog učenja te napraviti evaluaciju performansi.

5.1. Odabir i priprema baze

Pronašli smo nekoliko javno dostupnih baza slika koje sadrže lica kao što su: „Labeled Faces in the Wild (LFW)“, „CelebA“ te „IMDB-WIKI“. Sve navedene baze (eng. *Dataset*) sadrže slike s licima, ali se razlikuju u svojim strukturama. Naime, neke od njih sadrže .csv datoteku u kojoj je popis svih slika te x i y koordinate na kojima se nalaze lica, neke od njih imaju .mat datoteke i sl. S obzirom da mi moramo prepoznati osobu na slici, odlučili smo se na dataset „Labeled Faces in the Wild (LFW)“.

5.1.1.Labeled Faces in the Wild

Navedeni dataset popularna je baza podataka koja se često koristi u području računalnog vida i prepoznavanja lica. Sadrži slike lica iz realnih svakodnevnih situacija, obuhvaćajući raznolike pozadine, osvjetljenja i izraze lica. Baza sadrži više od 13000 slika koje predstavljaju preko 5000 različitih osoba. Odabrali smo navedenu bazu slika zbog idealne strukture za naš projekt. Baza sadrži poseban „folder“ za svaku osobu koji je nazvan imenom te osobe, a to nam će nam pojednostaviti kreiranje našeg modela, jer ne moramo podatke pronalaziti u .csv datotekama i sl.

Dataset se nalazi na sljedećem linku. <https://www.kaggle.com/datasets/jessicali9530/lfw-dataset>

5.2. Učitavanje i predprocesuiranje slika

Za uspješnu izgradnju modela prepoznavanja lica, ključno je provesti kvalitetno učitavanje i predprocesuiranje slika. Ovaj korak ima ključnu ulogu u oblikovanju učinkovitog skupa podataka koji ćemo koristiti za treniranje našeg modela.

Prvi je korak bio učitavanje baze koja sadrži slike postavljajući bazu u korijenski direktorij projekta, a nakon toga definiranje putanje do same baze. Koristeći python biblioteke TensorFlow i Keras implementirali smo učitavanje slika iz zadane baze podataka. Direktoriji su strukturirani prema klasama, a slike unutar njih prikupili smo pomoću osnovnih python funkcija za manipulaciju datoteka.

S obzirom da neke osobe sadrže samo jednu sliku, potrebno je filtrirati slike kako bi naš model bio što točniji, pa smo se iz tog razloga odlučili odabrati osobe koje sadrže više od 50 slika lica.

5.2.1.Predprocesuiranje slika

Jedan od ključnih koraka u pripremi baze za model prepoznavanja lica bio je augmentacija. Ona predstavlja skup tehnika koje se primjenjuju na slike kako bi se povećala njihova raznolikost, a time i sposobnost modela da se nosi s različitim uvjetima. Kako bismo predprocesuirali slike, koristili smo ImageDataGenerator iz Keras biblioteke. ImageDataGenerator koristi različite tehnike kao što su:

- raspodjela vrijednosti piksela (eng. *Rescale*) – služi nam za skaliranje vrijednosti piksela na raspon [0,1], čime se olakšava konvergencija modela tijekom treniranja. Na taj način sprječavamo dominaciju visokih vrijednosti

- izobličenje (eng. *Shear range*) – ova tehnika pomaže našem modelu kod prepoznavanja lica u različitim perspektivama
- povećanje (eng. *Zoom range*) – Koristimo povećanje kako bismo simulirali promjene udaljenosti osobe od kamere, a to nam koristi u situacijama gdje slike lica mogu varirati po veličini i udaljenosti od kamere
- horizontalno zrcaljenje (eng. *Horizontal flip*) – Primijenili smo horizontalno zrcaljenje slika kako bismo stvorili varijante slika sa zamijenjenim lijevom i desnom stranama, a samim time ćemo unaprijediti model koji ćemo istrenirati jer neće biti osjetljiv na horizontalne orijentacije lica
- podjela na skup za obuku i validaciju (eng. *Validation split*) – pomoću ovih parametara odvojili smo 20% podataka za validaciju, dok je ostatak podataka za treniranje. Samim time spriječili smo „prenaučnost“ modela te omogućili precizniju evaluaciju jer se radi o podacima s kojima se trenirani model još nije susreo

Processed Image 3
Ariel_Sharon\Ariel_Sharon_0018.jpg



Slika 12. Primjer procesuirane slike; autorski rad

5.3. Implementacija konvolucijske neuronske mreže (CNN)

Treniranje modela za prepoznavanje lica predstavlja ključan korak u razvoju sustava dubokog učenja. Korištenjem konvolucijskih neuronskih mreža (CNN), moguće je ekstrahirati složene značajke iz slika te naučiti reprezentacije koje su ključne za prepoznavanje lica u različitim uvjetima.

Konvolucijske neuronske mreže su vrsta neuronskih mreža prilagođena za obradu slika. One koriste konvolucijske slojeve za automatsko izdvajanje značajki iz vizualnih podataka. U kontekstu prepoznavanja lica, ove mreže su iznimno učinkovite jer mogu naučiti hijerarhiju značajki, prepoznati obrasce i prilagoditi se različitim orijentacijama i veličinama lica.

Kako bismo izgradili naš model, koristili smo razne slojeve:

- Konvolucijski slojevi – Oni detektiraju lokalne obrasce na slikama kao što su rubovi i teksture, a uz njih koristili smo i funkciju ReLU s ulaznim veličinama slika (100,100,3). ReLU funkcija ostavlja pozitivne vrijednosti nepromijenjenima, dok sve negativne postavlja na nulu, a to pomaže modelu pri učenju nelinearnih značajki. Veličina slika je 100x100 piksela, a brojka 3 označava tri kanala boja, odnosno crvenu, plavu i zelenu (RGB). Uz pomoć konvolucijskog sloja model može lakše prepoznati rubove, teksture i oblike
- Agregacijski slojevi – Nakon svakog konvolucijskog sloja koristili smo agregacijski sloj kako bismo smanjili prostornu dimenzionalnost i sačuvali bitne značajke. Tako će primjerice ako konvolucijski sloj ima veličinu 4x4 piksela, agregacijski sloj ga može smanjiti na 2x2 piksela
- Sloj ispadanja (eng. *Dropout*) – ovaj nam sloj pomaže u sprječavanju pretreniranosti modela, odnosno isključuje određeni postotak neurona tijekom treniranja čime se poboljšava robusnost modela. U funkciju za „Dropout“ potrebno je proslijediti postotak koji označava vjerojatnost da će neuron biti isključen (u našem smo primjeru koristili 0.3)
- Potpuno povezani slojevi – predstavljaju završni dio neuronskih mreža, a u njima se povezuju neuroni jedan za drugi. Koristili smo „Dense(256, activation='relu')“ kako bi prvi potpuno povezani sloj imao 256 neurona te koristio ReLU funkciju

5.4. Evaluacija performansi

Evaluacija performansi pomaže nam razumjeti koliko je model dobro istreniran, a izvršava se na podacima koje nije vidio tijekom treninga. Za evaluaciju performansi koristili smo „False Acceptance Rate (FAR)“ i „False Rejection Rate (FRR)“.

False Acceptance Rate predstavlja postotak lažno pozitivnih identifikacija u sustavu za prepoznavanje. Lažno prepoznavanje događa se kada sustav krivo prihvaća nekoga tko se ne podudara s biometrijskim podacima. Niža FAR vrijednost označava bolju preciznost sustava jer manje lažnih prihvaćanja ukazuje na manje greške identifikacije.

False Rejection Rate predstavlja postotak lažno negativnih identifikacija u sustavu za prepoznavanje. FRR se događa kada sustav ne uspije identificirati osobu, odnosno odbija priznati pravu osobu. Niži False Rejection Rate ukazuje na bolju osjetljivost sustava jer manje lažnih odbijanja znači da sustav rijetko propušta prave korisnike.

Kako bismo izvršili evaluaciju performansi koristili smo funkciju koja je ugrađena u Keras. Također, kreirali smo i matricu „zabune“ (eng. *Confusion matrix*) koja prikazuje broj stvarnih i predviđenih instanci za svaku klasu. Svaki element (i,j) u matrici predstavlja broj instanci koje pripadaju klasi i, a model je predvidio klasu j. *Confusion matrix* prikazali smo pomoću „Heatmap“ iz Seaborn biblioteke. Elementi matrice prikazuju se bojama, a brojevi unutar ćelija predstavljaju stvarne brojeve instanci za svaki par stvarne i predviđene klase. Takav vizualni prikaz olakšava nam analizu gdje model može imati poteškoća.

6. Zaključak

U ovom seminarskom radu istraživali smo područje prepoznavanja lica korištenjem dubokog učenja. Duboko učenje predstavlja inovativni pristup u razvoju sustava prepoznavanja lica, omogućujući napredne analize biometrijskih karakteristika koje su nezamjenjive u suvremenim sigurnosnim i identifikacijskim sustavima.

Biometrijske karakteristike, poput jedinstvenih obilježja lica, postale su esencijalne u borbi protiv prijetnji i poboljšanju sigurnosti. Duboko učenje, posebice koncepti kao što su konvolucijske neuronske mreže (CNN) i rekurentne neuronske mreže (RNN), omogućuje prepoznavanje lica pružajući sposobnost sustavima da uče iz velikih skupova podataka.

Naš je istreniran model uspješno prepoznao osobu sa slike uz vrlo visoku točnost korištenjem CNN na predprocesuiranim slikama. Zbog točnosti prepoznavanja, odredili smo da model može prepoznati samo osobe koje sadrže više od 50 slika lica u datasetu.

7. Popis literature

- [1] „Deep learning“. Wikipedia. Pristupljeno: 06. prosinac 2023. [Na internetu]. Dostupno na: https://en.wikipedia.org/wiki/Deep_learning
- [2] „Frank Rosenblatt“. Wikipedia. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: https://en.wikipedia.org/wiki/Frank_Rosenblatt
- [3] „Perceptron“. Wikipedia. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://en.wikipedia.org/wiki/Perceptron>
- [4] Jürgen Schmidhuber, „Who Invented Backpropagation?“. AI Blog, 2014. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://people.idsia.ch/~juergen/who-invented-backpropagation.html>
- [5] Jürgen Schmidhuber, „2011: DanNet triggers deep CNN revolution“. AI Blog, veljača 2021. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://people.idsia.ch/~juergen/DanNet-triggers-deep-CNN-revolution-2011.html>
- [6] M.Z. Alom *i ostali*, „A state of the Art Survey on Deep Learning Theory and Architectures“. MDPI, 05. ožujak 2019. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://www.mdpi.com/2079-9292/8/3/292>
- [7] Fumić Patrik, „Duboko učenje: pregled područja“. Nacionalni repozitorij završnih i diplomskih radova ZIR, 24. rujan 2021. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://zir.nsk.hr/islandora/object/foi:6745>
- [8] „Deep reinforcement learning“. Wikipedia. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: https://en.wikipedia.org/wiki/Deep_reinforcement_learning
- [9] „What are Autoencoders? Introduction to Autoencoders in Deep Learning“. Simplilearn. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://www.simplilearn.com/tutorials/deep-learning-tutorial/what-are-autoencoders-in-deep-learning>
- [10] Hrishav Bandyopadhyay, „Autoenkoderi“. v7labs. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://www.v7labs.com/blog/autoencoders-guide>
- [11] „Restricted Boltzmann Machine“. Geeksforgeeks. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://www.geeksforgeeks.org/restricted-boltzmann-machine/>
- [12] Prerna Sodani, „Restricted Boltzmann Machine“. Latentview, 08. rujan 2023. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://www.latentview.com/blog/restricted-boltzmann-machine-and-its-application/>
- [13] K. Bharath, „Introduction to Deep Neural Networks“. Datacamp. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://www.datacamp.com/tutorial/introduction-to-deep-neural-networks>
- [14] „Convolutional neural network“. wikipedia. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: https://en.wikipedia.org/wiki/Convolutional_neural_network
- [15] Mayank Mishra, „Convolutional Neural Networks, Explained“. TowardsDataScience, 26. kolovoz 2020. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939>
- [16] „Convolution Layer“. Sciencedirect. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://www.sciencedirect.com/topics/engineering/convolutional-layer>
- [17] „Introduction To Pooling Layers In CNN“. Editorial Team, 16. kolovoz 2022. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://towardsai.net/p/l/introduction-to-pooling-layers-in-cnn>
- [18] „Fully Connected Layers in Convolutional Neural Networks“. IndianTechWarrior. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://indiantechwarrior.com/fully-connected-layers-in-convolutional-neural-networks/>

- [19] „Fully-Connected Layer with dynamic input shape“. Medium, 31. srpanj 2019. Pristupljeno: 12. lipanj 2023. [Na internetu]. Dostupno na: <https://medium.com/@tecokids.monastir/fully-connected-layer-with-dynamic-input-shape-70c869ae71af>
- [20] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, i S. Valaee, „Recent Advances in Recurrent Neural Networks“. Arxiv, 29. prosinac 2017. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: <https://arxiv.org/abs/1801.01078>
- [21] J. A. Bullinaria, „Recurrent Neural Networks“. 2015. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: <https://www.cs.bham.ac.uk/~jxb/INC/112.pdf>
- [22] „Recurrent neural network“. Wikipedia. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: https://en.wikipedia.org/wiki/Recurrent_neural_network
- [23] M. R. Mahmood, „Fully recurrent“. ResearchGate. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: https://www.researchgate.net/figure/Figure-6-The-fully-Recurrent-Neural-Network_fig4_320853485
- [24] M. D. Femila i A. A. Irudhayaraj, „Biometric System“. IEEE Xplore, 07. srpanj 2011. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: <https://ieeexplore.ieee.org/abstract/document/5941580>
- [25] „Biometrics“. Wikipedia. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: <https://en.wikipedia.org/wiki/Biometrics>
- [26] „Biometrics: History, Origin & How It Affects Our Lives“. Protelhrms, 13. listopada 2020. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: <https://protelhrms.com/biometrics-history-origin-how-it-affects-our-lives-2/>
- [27] H. van de Haar, D. van Greunen, i D. Pottas, „The Characteristics of a Biometric“. IEEE Xplore. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6641037>
- [28] „Who What Why: How durable is a fingerprint?“. BBC, 29. rujan 2012. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: <https://www.bbc.com/news/magazine-19740979>
- [29] „Iris recognition“. Wikipedia. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: https://en.wikipedia.org/wiki/Iris_recognition
- [30] Cole Murray, „Building a Facial Recognition Pipeline with Deep Learning in Tensorflow“. Medium, 07. kolovoz 2017. Pristupljeno: 12. srpanj 2023. [Na internetu]. Dostupno na: <https://medium.com/hackernoon/building-a-facial-recognition-pipeline-with-deep-learning-in-tensorflow-66e7645015b8>

8. Popis slika

Slika 1. Autoenkoderi [10]	6
Slika 2. Restricted Boltmann Machine [12]	7
Slika 3. ANN vs DNN [13]	9
Slika 4. Convolution Layer [16].....	10
Slika 5. Pooling Layer [15]	11
Slika 6. Fully Connected Layer [19]	12
Slika 7. Fully recurrent neural network [23]	14
Slika 8. Povijest biometrijskih sustava [26]	16
Slika 9. Otisak prsta [28].....	18
Slika 10. Prepoznavanje šarenice oka [29].....	19
Slika 11. Prepoznavanje lica [30].....	20
Slika 12. Primjer procesuirane slike; autorski rad.....	22