



Univerzitet u Sarajevu
Elektrotehnički fakultet u Sarajevu
Odsjek za Računarstvo i informatiku



Detekcija pravca pogleda

Eye gaze detection

PREDMET: Vještačka inteligencija
– PROJEKT –

Studentice:
Dženeta Milić
Lamija Gutić
Fatima Kaljanac

Supervizor:
Elvedin Smajić

Sarajevo, juni 2025.

Sadržaj

1. Uvod.....	3
1.1 Opis problema.....	3
1.2 Osnovni pojmovi.....	3
1.3 Korist i primjena.....	4
1.4 Pregled postojećih dataset-ova.....	4
2. Pregled stanja u oblasti.....	5
2.1 Trenutno stanje u oblasti.....	5
2.2 Postojeći opseg problema.....	5
2.3 Prethodno korištene AI metode za detekciju pokreta oka.....	6
2.4 Dosadašnji postignuti rezultati.....	7
2.5 Identificirani pravci za poboljšanje.....	7
3. Izbor, analiza i pretprocesiranje dataset-a.....	8
3.1 Izbor dataset-a.....	8
3.2 Detaljna analiza dataset-a.....	8
3.3 Pretprocesiranje podataka.....	11
3.4 Identificirani rizici.....	12
4. Odabir, formiranje, treniranje i testiranje modela.....	12
4.1 Izbor metoda.....	12
4.2 Arhitektura modela.....	13
4.3 Tehnologije.....	17
4.4 Treniranje modela.....	17
4.5 Evaluacija modela.....	19
4.6 Implementacija za djecu.....	22
5. Zaključak.....	23
5.1 Analiza postignuća.....	23
5.2 Poređenje s postojećim rješenjima.....	23
5.3 Moguća poboljšanja.....	24
5.4 Društveni utjecaj.....	24
Literatura.....	26

1. Uvod

Detekcija pravca pogleda (eng. *eye gaze detection*) koristi prepoznavanje slika (eng. *image recognition*) sa dubokim učenjem za praćenje pokreta oka. Praćenje pokreta oka postaje veoma važna tehnika u mnogim disciplinama, poput sigurnosti, neuronauke, psihologije medicinske dijagnostike i slično. Praćenjem pokreta oka moguće je odrediti nivo pažnje posmatrane osobe, što može biti izuzetno korisno za razne primjene.

1.1 Opis problema

U ovom projektu, posebna pažnja bit će posvećena korištenju tehnike praćenja pravca pogleda u razumijevanju ponašanja djece sa autizmom ili drugim poteškoćama u razvoju. Ova djeca često se suočavaju sa problemima u verbalnoj komunikaciji, socijalnoj interakciji i pažnji. Takvoj djeci često bude teško izraziti svoje potrebe, osjećaje ili odgovore na uobičajene načine, što može prouzrokovati pogrešno razumijevanje njihovog ponašanja, kao i frustraciju. Zbog toga ovakvi alati igraju važnu ulogu u razumijevanju njihovih potreba i stanja bez oslanjanja na verbalnu komunikaciju.

1.2 Osnovni pojmovi

Osnovni pojmovi koji će se spominjati u ovom projektu su:

1. **Eye gaze tehnologija:** Tehnologija koja omogućava praćenje smjera i fokusa pogleda korisnika. Koristi se za analizu onoga što osoba gleda u realnom vremenu, što služi i kao ulazni signal za interakciju s uređajima ili razumijevanje pažnje i emocionalnog stanja [1].
2. **Praćenje pogleda** (eng. *eye tracking*): Proces snimanja i analiziranja pokreta očiju pomoću specijaliziranih kamera i softvera [2].
3. **Autizam** (eng. *Autism Spectrum Disorder* - ASD): Razvojni poremećaj koji utiče na komunikaciju, ponašanje i socijalnu interakciju. Djeca s autizmom često imaju probleme u komunikaciji i drugačije procesiraju ulazne senzorne informacije, uključujući vizualne informacije [3].
4. **Poremećaj pažnje s hiperaktivnošću** (eng. *Attention Deficit Hyperactivity Disorder* - ADHD): Neurološki razvojni poremećaj karakteriziran problemima s pažnjom, impulsivnošću i hiperaktivnošću. Kod djece s ADHD-om mogu se javiti problemi u održavanju vizualnog fokusa [4].
5. **Alternativna i augmentativna komunikacija** (eng. *Alternative and Augmentative Communication* - AAC): Skup metoda i tehnologija koje

pomažu osobama s poteškoćama u razvoju da izraze svoje misli, emocije i potrebe. Detekcija pravca pogleda može biti jedna od tehnika u okviru AAC sistema [5].

6. **Neverbalna komunikacija:** oblik komunikacije koji se ne oslanja na govor, već na geste, izraze lica, kontakt očima i druge vizualne signale [6].
7. **Senzorna preopterećenost** (eng. *sensory overload*): Stanje kada osoba postane preopterećena prevelikim brojem senzornih podražaja (zvuk, svjetlo, dodir, miris itd.) što može izazvati izuzetan stres ili povlačenje [7].

U nastavku teksta će se koristiti skraćenice za termine za koje su navedene.

1.3 Korist i primjena

Primjena praćenja pravca pogleda je višestruka, ali najveća korist ove tehnike vidi se u **medicinskoj dijagnostici**. Praćenje pogleda se koristi za detekciju specifičnih obrazaca kod osoba sa neurološkim oboljenjima, poput Parkinsonove bolesti, Alzheimerove bolesti, i drugih sličnih oboljenja. Osim toga, ova tehnologija može se koristiti i za praćenje bolesti poput glaukoma, makularne degeneracije i drugih sličnih oftalmoloških dijagnoza [8].

Osim medicine, ova tehnika se koristi i u automobilske industriji za praćenje pogleda vozača, detekciju umora i određivanje nivoa pažnje. Pored toga, praćenje pravca pogleda pokazalo se korisnim u oblastima kao što su marketing i istraživanje potrošačkih navika, u istraživanju angažmana učenika u obrazovnim procesima, kao i u virtualnoj i proširenoj stvarnosti (eng. *virtual reality* i *augmented reality*).

U ovom projektu, središte je zadržano na upotrebi detekcije pravca pogleda u medicinskoj dijagnostici za djecu sa poteškoćama u razvoju.

1.4 Pregled postojećih *dataset*-ova

Postoji nekoliko *online* dostupnih *dataset*-ova koji se mogu koristiti za svrhu ovog projekta. To su:

1. [ETDD70: Eye-Tracking Dyslexia Dataset](#) [9]
2. [Saliency4ASD: A dataset of eye movements for the children with autism spectrum disorder](#) [10]
3. [ChildPlay: A New Benchmark for Understanding Children's Gaze Behaviour](#) [11]
4. [Peekbank: An open, large-scale repository for developmental eye-tracking data of children's word recognition](#) [12]

2. Pregled stanja u oblasti

U ovom odjeljku bit će dat pregled trenutnog stanja u oblasti detekcije pravca pogleda, kao i opseg problema sa kojima će se baviti ovaj projekt.

2.1 Trenutno stanje u oblasti

Detekcija praćenja pogleda doživjela je veliki napredak razvojem integracije tehnika za duboko učenje. Nedavna istraživanja naglašavaju tranziciju sa metoda baziranih na modelima na metode bazirane na dubokom učenju, posebno metode zasnovane na izgledu (eng. *appearance-based methods*). Ove metode koriste konvolucijske neuralne mreže (eng. Convolutional neural networks - CNN) da mapiraju slike lica direktno na vektore pogleda. Ove metode su se posebno dobro pokazale u radu sa varijacijama u poziciji glave i osvjetljenosti [13].

Neke od zapaženih tehnika su:

- DHECA-SuperGaze: ova tehnika koristi mehanizam dvostruke ukrštene pažnje između glave i oka, u kombinaciji s tehnikama super-rezolucije za poboljšanje procjene pogleda u nekontroliranim okruženjima.
- RAGE-net: Rezidualna mreža za procjenu pogleda zasnovana na izgledu, koja eliminiše potrebu za kalibracijom što je čini pogodnom za primjenu u stvarnim uslovima.
- MGazeNet: podrazumijeva linearno adaptivnu normalizaciju serija (batch normalization) radi efikasnog kombiniranja karakteristika očiju i lica čime se postižu izuzetno dobri rezultati u zadacima procjene pogleda.

2.2 Postojeći opseg problema

Uprkos značajnom napretku u oblasti detekcije praćenja pogleda, postojeći pristupi se suočavaju s nekoliko ključnih problema i ograničenja:

- **Problemi vezani za podatke i okruženje**
 - **Varijabilnost uslova snimanja:** Postojeći sistemi često pokazuju smanjenu preciznost u različitim uslovima osvjetljenosti, pozicije kamere i ugla snimanja. Ovo posebno utiče na performanse u nekontroliranim okruženjima kao što su javni prostori ili domovi korisnika.
 - **Heterogenost populacije:** Trenutni modeli često pokazuju pristranost prema određenim demografskim grupama, što rezultira

nejednakim performansama za različite etničke grupe, uzraste ili osobe koje nose naočale.

- **Ograničena generalizacija:** Modeli obučeni na specifičnim *dataset*-ovima često ne generalizuju dobro na nove domene ili uslove koji nisu bili prisutni tokom obuke.
- **Tehnički izazovi:**
 - **Preciznost kalkulacije:** Postizanje visoke preciznosti u procjeni pravca pogleda ostaje izazov, posebno za aplikacije koje zahtevaju milimetarsku preciznost kao što su asistivne tehnologije.
 - **Obrada u realnom vremenu:** Balansiranje između složenosti modela i brzine obrade predstavlja značajan izazov za aplikacije koje zahtijevaju trenutni odgovor sistema.
 - **Robusnost na okluziju:** Postojeći sistemi često gube na preciznosti kada su dijelovi lica ili očiju zaklonjeni, što je čest slučaj u praktičnim scenarijima.

2.3 Prethodno korištene AI metode za detekciju pokreta oka

Razvoj metoda za detekciju pokreta oka još uvijek se odvija, pri čemu su tradicionalne pristupe potpuno zamijenili napredne tehnike dubokog učenja (eng. *deep learning*).

Tradicionalni pristup podrazumijevaju:

1. Metode bazirane na modelima; metode koje se oslanjaju na geometrijske modele oka i lica. Ove tehnike koristile su detekciju karakterističnih tačaka za lokalizaciju zjenice i rožnjače, kombinirajući 3D modele glave za procjenu pravca pogleda.
2. Support Vector Machines (SVM): Korišteni za klasifikaciju različitih pravaca na osnovu ekstraktovanih karakteristika kao što su Haar ili Local Binary Patterns (LBP).

Napredni pristup detekciji pokreta oka podrazumijevaju:

1. Konvolucijske neuralne mreže (eng. Convolutional neural networks): ovaj pristup koristi različite konvolucijske neuralne mreže za ekstrakciju karakteristika iz slika očiju u sivim tonovima i kombiniranje ovih karakteristika sa pozicijom glave.
2. Višemodalni pristup: kombiniranje informacija iz oba oka, pozicije glave i konteksta lica
3. Mehanizmi pažnje: implementacija mehanizma pažnje za fokusiranje na najznačajnije regije oko oka, što je rezultiralo boljim performansama u kompleksnijim scenarijima.

2.4 Dosadašnji postignuti rezultati

Procjena ovih metoda pokazala je kontinuiran napredak u preciznosti i robusnosti sistema. Osnovne metrike koje se koriste za mjerenje performansi su ugaona greška i preciznost klasifikacije. Ugaona greška predstavlja razliku između procijenjenog i stvarnog pravca pogleda. Istraživanja su pokazala da se srednja ugaona greška može smanjiti na 0.25° nakon korekcije procijenjenog pogleda.

Rezultati za pojedinačne *dataset*-ove:

1. MPIIGaze dataset: ključni dataset za procjenu, na kojem najbolje metode mogu postići [15]:
 - Ugaona greška: $4.3-6.2^\circ$
 - Cross-dataset generalizacija: $7.5-12.1^\circ$
2. EyeDiap dataset: dataset pripremljen za interakciju na daljinu [16]:
 - Najbolje CNN metode: $5.4-7.8^\circ$ ugaona greška
 - Značajno poboljšanje u odnosu na tradicionalne metode ($12-15^\circ$)
3. GazeCapture dataset: najveći dataset sa podacima prikupljenim preko mobilnih uređaja [17]:
 - State-of-the-art rezultati: $3.8-4.5^\circ$ ugaona greška
 - Posebno dobra generalizacija na različite uređaje

2.5 Identificirani pravci za poboljšanje

Analizirajući trenutno stanje i postignute rezultate, mogu se identificirati ključni pravci za buduća poboljšanja:

1. Poboljšanje preciznosti
 - **Tehnike adaptacije domena:** Razvoj metoda koje omogućavaju bolju generalizaciju između različitih dataset-ova i uslova korištenja. Cilj je smanjenje cross-dataset greške sa trenutnih $7-12^\circ$ na $3-5^\circ$.
 - **Few-shot learning pristupi:** Implementacija tehnika koje omogućavaju brzu adaptaciju na nove korisnike sa minimalnom količinom podataka za kalibraciju.
 - **Multi-task learning:** Istovremeno učenje više povezanih zadataka (detekcija pogleda, prepoznavanje emocija, detekcija pospanosti) radi poboljšanja generalne reprezentacije.
2. Tehnološka poboljšanja:
 - Jednostavnije arhitekture: Razvoj efikasnijih arhitektura neuralnih mreža koje su optimizirane za mobilne uređaje
 - Optimizacija za realno vrijeme; implementiranje tehnika za poboljšanje brzine.

3. Izbor, analiza i pretprocesiranje *dataset-a*

3.1 Izbor *dataset-a*

U okviru realizacije ovog projekta, pažljivo smo pristupili odabiru odgovarajućeg dataset-a koji će služiti kao osnova za našu analizu. Proces selekcije dataset-a predstavlja fundamentalan korak koji direktno utiče na kvalitet i validnost rezultata koje planiramo postići kroz projekat.

Prilikom evaluacije dostupnih opcija, postavili smo jasne kriterije za odabir. Dataset je trebao biti relevantan za temu našeg projekta, sadržavati potrebne varijable za analizu, imati dovoljnu veličinu uzorka za značajne statističke analize, te biti dobro dokumentovan s obzirom na metodologiju prikupljanja podataka.

Nakon detaljnog pregleda nekoliko potencijalnih dataset-ova, opredijelili smo se za "[Eye Tracking Autism](#)" dataset koji je dostupan na Kaggle platformi [14]. Ovaj dataset odgovara svim našim postavljenim kriterijima - sadrži relevantne podatke o praćenju pokreta očiju kod osoba s autizmom, ima solidnu dokumentaciju i pruža dovoljno širok spektar varijabli potrebnih za našu projektnu analizu.

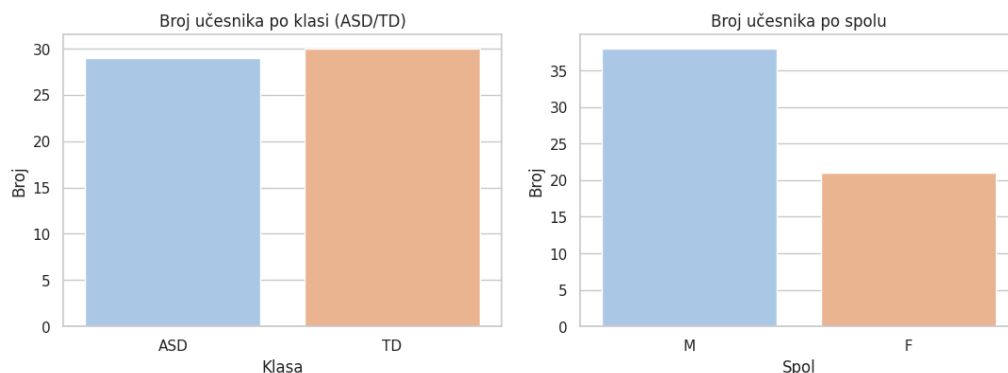
Dodatni faktori koji su utjecali na naš izbor uključuju verifikaciju dataset-a na Kaggle platformi, pozitivne korisničke recenzije, kao i činjenicu da je dataset javno dostupan što omogućava transparentnost našeg rada i mogućnost reprodukcije rezultata.

3.2 Detaljna analiza *dataset-a*

Dataset je distribuiran kroz 25 CSV datoteka, gdje svaka datoteka predstavlja izlaz jednog eye-tracking eksperimenta. Važno je napomenuti da pojedinačni eksperiment obično uključuje više učesnika, što je jasno označeno kroz 'Participant' kolonu koja sadrži jedinstveni identifikator svakog učesnika. Ova kolona takođe omogućava klasifikaciju učesnika u dvije glavne kategorije: tipično razvijajuće (Typically Developing) i osobe sa autizmom spektra poremećaja (ASD).

Analiza glavnog metadata fajla Metadata_Participants.csv pokazuje da dataset sadrži podatke o ukupno 59 učesnika. Struktura metapodataka uključuje pet ključnih kolona: ParticipantID (jedinstveni identifikator), Gender (pol), Age

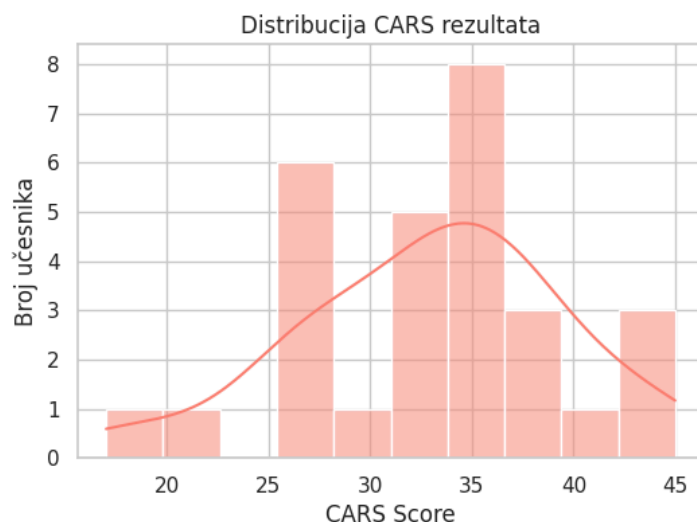
(uzrast), Class (klasifikacija učesnika) i CARS Score (rezultat na skali za procjenu autizma).



Slika : Distribucija učesnika po klasi i spolu

Analiza uzrasta učesnika pokazuje da se kreće u rasponu od 2,7 do 12,9 godina, sa prosječnim uzrastom od 7,9 godina. Ova distribucija je posebno relevantna za istraživanje jer pokriva ključne razvojne periode u djetinjstvu.

Što se tiče distribucije po spolu, dataset sadrži 38 učesnika muškog spola (64,4%) i 21 učesnika ženskog spola (35,6%). Interesantno je primijetiti razliku u distribuciji spola između grupa: u ASD grupi prevladavaju muškarci (25 od 29 učesnika), dok je u TD grupi distribucija po spolu relativno uravnotežena.



Slika : Distribucija CARS rezultata po broju učesnika

Dataset je gotovo savršeno balansiran između dvije glavne kategorije: 30 učesnika pripada tipično razvijajućoj grupi (TD), dok 29 učesnika pripada ASD

grupi. Ova uravnotežena distribucija je posebno vrijedna za statističke analize i omogućava objektivno poređenje između grupa.

CARS (Childhood Autism Rating Scale) rezultati su dostupni za 29 učesnika (isključivo iz ASD grupe), što predstavlja 49% ukupnog uzorka. Prosječni CARS rezultat iznosi 32,97, sa standardnom devijacijom od 6,55. Raspon rezultata kreće se od 17 do 45, što ukazuje na različite stepene težine autizma spektra poremećaja među učesnicima.

Struktura eksperimentalnih datoteka

Svaka od 25 eksperimentalnih CSV datoteka sadrži detaljne eye-tracking mjere organizovane kroz 37 kolona koje pokrivaju različite aspekte praćenja pokreta očiju. Struktura podataka omogućava precizno praćenje ponašanja pogleda tokom eksperimentalnih zadataka.

Tabela 1 prikazuje detaljne opise svih varijabli dostupnih u CSV datotekama:

Naziv kolone	Opis kolone
RecordingTime [ms]	Vrijeme snimanja u milisekundama
Time of Day [h:m:s:ms]	Vrijeme dana u formatu sati:minuti:sekunde:milisekunde
Trial	Identifikator pojedinačnog pokusa u eksperimentu
Stimulus	Identifikator stimulusa prikazanog učesniku
Export Start Trial Time [ms]	Početno vrijeme pokusa u milisekundama
Export End Trial Time [ms]	Završno vrijeme pokusa u milisekundama
Participant	Identifikator učesnika
Color	Boja stimulusa
Tracking Ratio [%]	Procenat uspješnog praćenja pogleda
Category Group	Kategorijska grupa stimulusa
Category Right	Kategorija za desno oko
Category Left	Kategorija za lijevo oko
Index Right	Indeks za desno oko
Index Left	Indeks za lijevo oko
Pupil Diameter Right [mm]	Promjer zjenice desnog oka u milimetrima
Pupil Diameter Left [mm]	Promjer zjenice lijevog oka u milimetrima
Point of Regard Right X [px]	X koordinata pogleda desnog oka u pikselima
Point of Regard Right Y [px]	Y koordinata pogleda desnog oka u pikselima
Point of Regard Left X [px]	X koordinata pogleda lijevog oka u pikselima

Point of Regard Left Y [px]	Y koordinata pogleda lijevog oka u pikselima
AOI Name Right	Naziv oblasti od interesa (AOI) za desno oko
AOI Name Left	Naziv oblasti od interesa (AOI) za lijevo oko
Gaze Vector Right X	X komponenta vektora pogleda desnog oka
Gaze Vector Right Y	Y komponenta vektora pogleda desnog oka
Gaze Vector Right Z	Z komponenta vektora pogleda desnog oka
Gaze Vector Left X	X komponenta vektora pogleda lijevog oka
Gaze Vector Left Y	Y komponenta vektora pogleda lijevog oka
Gaze Vector Left Z	Z komponenta vektora pogleda lijevog oka
Annotation Name	Naziv anotacije
Annotation Description	Opis anotacije
Annotation Tags	Oznake anotacije
Mouse Position X [px]	X koordinata pozicije miša u pikselima
Mouse Position Y [px]	Y koordinata pozicije miša u pikselima
Scroll Direction X	Smjer skrolovanja po X osi
Scroll Direction Y	Smjer skrolovanja po Y osi
Content	Sadržaj stimulusa

Tabela 1: Opis varijabli u CSV datotekama

3.3 Pretprocesiranje podataka

Prije bilo kakve ekstrakcije značajki i daljeg rada sa podacima, bilo je potrebno spojiti 25 datoteka CSV formata u jedan Pandas DataFrame radi jednostavnijeg i bržeg rada, te ujednačiti tipove podataka u svim datotekama kako bi bili kompatibilni. Zatim su učitani metapodaci o učesnicima iz datoteke `Metadata_participants.csv`, koja je sadržavala informacije poput dobi, ID-a učesnika, CARS score-a (koji opisuje gdje se osoba nalazi na autističnom spektru), kao i klase (ASD ili TD). Pošto je DataFrame uključivao i kolonu sa postotkom praćenja oka (`Tracking Ratio [%]`), uklonjeni su uzorci sa niskim koeficijentom pouzdanosti. Također je izvršena normalizacija x i y koordinata pogleda u odnosu na širinu i visinu ekrana, te su normalizirane vrijednosti spremljene u nove kolone `x_norm` i `y_norm`.

Dalje, kako bi se omogućila kvalitetna ekstrakcija značajki, izvršeno je dodatno čišćenje podataka — kolona `ParticipantID` je uklonjena radi izbjegavanja

redundancije, a osigurano je da informacije o učesnicima (spol, dob, klasa) budu jedinstvene po učesniku.

Za svaki par učesnik–trial, izračunate su statističke karakteristike poput srednjih vrijednosti širine zjenica za desno i lijevo oko, prosječnog postotka praćenja, dobi učesnika, kodiranog spola (0 za ženski, 1 za muški), te CARS score-a. Tako dobijeni vektori značajki standardizirani su korištenjem **StandardScaler** kako bi se omogućilo da sve značajke imaju jednak doprinos u kasnijem modeliranju.

Pored toga, izračunate su agregirane značajke i na nivou učesnik–stimulus, uključujući srednje vrijednosti i standardne devijacije za promjer zjenica, koordinate pogleda (x, y za oba oka), postotak praćenja i vrijeme snimanja. Nakon toga, klase su transformirane u numeričke ciljne varijable (**target**), gdje su ASD učesnici označeni sa 1, a TD učesnici sa 0, kako bi podaci bili spremni za primjenu različitih modela mašinskog učenja.

3.4 Identificirani rizici

Rizici koji se vežu za ovakve podatke su:

1. Niski koeficijenti pouzdanosti, iako se vrši otklanjanje onih koji su izrazito niski
2. Neki podaci nedostaju: uređaji za praćenje pokreta oka nekada prestanu pratiti oko ili bilježe neispravne koordinate
3. Greške u kalibraciji: ako uređaj za bilježenje nije bio dobro kalibriran, tačke pogleda mogu biti pogrešno zabilježene.
4. Nekonzistentna ID-evi učesnika: ako ovi identifikatori nisu standardizirani, javlja se greška pri spajanju sam etapodacima
5. Mali uzorci (za podgrupe ASD ili TD): mala količina podataka može ograničiti statističku moć i mogućnost generalizacije.

4. Odabir, formiranje, treniranje i testiranje modela

4.1 Izbor metoda

U ovom projektu korištene su četiri metode koje će biti prikazane u nastavku i kasnije upoređene po odgovarajućim metrikama. Četiri metode korištene u ovom projektu su:

1. Rekurentna neuralna mreža (RNN) sa pažnjom i *long short-term* memorijom (LSTM)
2. 1D vremenska konvolucijska neuralna mreža (1D TCNN)
3. Random forest algoritam
4. Višeslojni perceptron (eng. Multilayer perceptron - MLP)

Rekurentna neuralna mreža izabrana je jer zadržava **memoriju prošlih stanja**, modelira **dinamičke odnose kroz vrijeme**, i omogućuje učenje **kontinuiteta ponašanja korisnika**. Dvosmjerni LSTM odabran je jer **paralelno uči iz prošlosti i budućnosti** unutar svakog vremenskog prozora, bolje modelira **ambijent i tok pažnje** kod korisnika, i daje robusnije reprezentacije ponašanja.

Druga metoda odabrana je uzimajući u obzir vremensku dimenziju praćenja pokreta oka, budući da su podaci o pokretima očiju vremenske serije, a ovakva konvolucijska neuralna mreža služi upravo za takvu vrstu podataka. Također, kako pokreti oka zavise od prethodnih pokreta, ovakva mreža može vidjeti širi vremenski kontekst.

Treći navedeni algoritam odabran je budući da dobro radi sa veoma različitim tipovima podataka i jednostavno je moguće uvidjeti koji atributi (kao što su pozicija oka, frekvencija, trajanje) najviše doprinose klasifikaciji.

Četvrta metoda odabrana je budući da MLP može da modelira nelinearne odnose između ulaznih karakteristika i omogućava reprezentaciju podataka o pokretima oka kroz vektor karakteristika, te je jednostavniji za treniranje u odnosu na sekvencijalne modele. U ovom slučaju najčešće se koriste agregirani podaci o sekvencama.

4.2 Arhitektura modela

4.2.1 Arhitektura rekurentne neuralne mreže sa LSTM

Arhitektura ove rekurentne neuralne mreže izgleda ovako:

```
inputs = Input(shape=input_shape)
x = layers.SpatialDropout1D(0.1)(inputs)

x = layers.Bidirectional(layers.LSTM(64,
return_sequences=True))(x)
x = Attention()(x)
x = layers.Dense(64, activation='relu')(x)
```

```

x = layers.Dropout(0.2)(x)

x = layers.Dense(32, activation='relu')(x)
x = layers.Dropout(0.2)(x)
outputs = layers.Dense(1, activation='sigmoid')(x)
model = models.Model(inputs, outputs)

model.compile(
    optimizer=Adam(learning_rate=2e-3, beta_1=0.9,
beta_2=0.999),
    loss='binary_crossentropy',
    metrics=['accuracy', metrics.AUC()])

```

Model koristi:

1. **SpatialDropout1D** — Redukcija prenaučnosti.
2. **Bidirectional LSTM** (64 jedinica) — bilježi uzorke u oba vremenska smjera.
3. **FastAttention** — Složeniji mehanizam pažnje koji omogućuje modelu fokusiranje na relevantne vremenske korake.
4. Dva gusto povezana sloja (Dense) s ReLU i Dropout.
5. Izlazni sloj s **sigmoid** aktivacijom za binarnu klasifikaciju.

Model je optimiziran s **Adam** optimizatorom (LR = 2e-3) i koristi metrike accuracy i AUC.

4.2.2 Arhitektura 1D vremenske konvolucijske neuralne mreže

Arhitektura ovog modela izgleda ovako:

```

model = models.Sequential([
    Input(shape=input_shape),
    layers.Conv1D(64, kernel_size=5, padding='same',
activation='relu'),
    layers.BatchNormalization(),
    layers.Conv1D(128, kernel_size=5, padding='same',
activation='relu'),
    layers.BatchNormalization(),
    layers.GlobalAveragePooling1D(),
    layers.Dense(64, activation='relu'),
    layers.Dropout(0.3),

```

```
layers.Dense(1, activation='sigmoid')
])
```

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 333, 64)	2,304
batch_normalization (BatchNormalization)	(None, 333, 64)	256
conv1d_1 (Conv1D)	(None, 333, 128)	41,088
batch_normalization_1 (BatchNormalization)	(None, 333, 128)	512
global_average_pooling1d (GlobalAveragePooling1D)	(None, 128)	0
dense (Dense)	(None, 64)	8,256
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 1)	65

Ova mreža kao ulaz prima sekvence karakteristika praćenja oka (širina zjenica, pozicija oka, vrijeme), i kao izlaz daje binarnu klasifikaciju (sigmoid). U ovom slučaju, izlaz se vrši da li je predikcija da osoba ima autizam ili ne.

Detaljna arhitektura:

1. Ulazni sloj - predstavlja sekvencu mjerenja pokreta oka za neki vremenski period
2. Prvi konvolucijski sloj: Conv1D sa 64 filtera, veličinom kernela 5, 'same' paddingom i ReLu aktivacijskom funkcijom
3. Normalizacijski sloj: sloj koji normalizira prvi konvolucijski sloj da bi ubrzalo treniranje i povećalo stabilnost

4. Drugi konvolucijski sloj sa 128 filtera, veličinom kernela 5, 'same' paddingom i ReLU aktivacijskom funkcijom (bilježi vremenske karakteristike iz sekvence na visokom nivou)
5. Normalizacijski sloj
6. Global Average Pooling 1D sloj - smanjuje mapu svake značajke na jednu vrijednost tokom vremenske dimenzije
7. Dense sloj sa 64 jedinice i ReLU aktivacijskom funkcijom
8. Dropout sloj sa brzinom 0.3
9. Izlazni sloj sa jednom jedinicom i Sigmoid aktivacijskom funkcijom

4.2.3 Arhitektura Random Forest modela

Ovaj model se sastoji od ansambla (skupa) nasumično generisanih stabala odlučivanja. Svako drvo donosi svoju "odluku" (klasifikaciju), a konačna predikcija je rezultat **glasanjem većine** među stablima. Zahvaljujući nasumičnom uzorkovanju podataka i karakteristika prilikom učenja svake stabla, model smanjuje preprilagođavanje (overfitting) i postiže dobru opštu sposobnost (generalizaciju).

```
from sklearn.ensemble import RandomForestClassifier

import joblib

rf = RandomForestClassifier(random_state=42, n_jobs=-2)

rf.fit(X_train, y_train)
```

Arhitektura ovakvog klasifikatora izgleda ovako:

- Broj stabala (**n_estimators**): 100 stabala
- Maksimalna dubina stabala (**max_depth**): nije zadato → stabla mogu da rastu dok ne "pređu" listove (fully grown trees)
- Kriterij za dijeljenje čvorova (**criterion**): Gini impurity.
- Broj paralelnih radnih procesa (**n_jobs=-2**): koristi sve CPU jezgre osim jednog (optimalno za ubrzanje treniranja).
- Random state: 42
- Bootstrap sampling: omogućeno — svako stablo se trenira na uzorku podataka sa vraćanjem.

4.2.4 Arhitektura Multilayer perceptrona

```
mlp = MLPClassifier(hidden_layer_sizes=(300,150,100), max_iter=800,
early_stopping=True,
```



```

        validation_fraction=0.2, n_iter_no_change=15,
random_state=42, alpha=0.001)

mlp.fit(X_train_scaled, y_train)

```

Arhitektura MLP modela izgleda ovako:

1. Broj slojeva: 3 skrivena sloja
2. Veličina slojeva:
 - a. Prvi skriveni sloj: 300 neurona
 - b. Drugi skriveni sloj: 150 neurona
 - c. Treći skriveni sloj: 100 neurona
3. **Ulazni sloj:** broj neurona odgovara broju ulaznih osobina u X_train
4. Izlazni sloj: broj neurona odgovara broju klasa (automatski detektuje na osnovu y_train)
5. **Funkcija aktivacije:** ReLU (za skrivene slojeve), Softmax/Logistic za izlaz (zavisi da li je binarna ili višeklasna klasifikacija)
6. **Funkcija troška:** Log-loss (cross-entropy), što je podrazumevano za klasifikaciju
7. **Optimizator:** Adam (kombinacija RMSProp i momentum SGDa)
8. **Regularizacija:** L2 regularizacija (**alpha=0.001**) – sprečava overfitting
9. **Standardizacija podataka:** korišćen je **StandardScaler**, što je jako važno kod MLP-a zbog stabilnosti učenja

4.3 Tehnologije

Za implementaciju programskog koda korištene su Python biblioteke:

1. Pandas za obradu podataka
2. Scikit-learn za Random Forest i MLP klasifikatore, preprocesiranje i evaluaciju modela
3. Numpy za dalju numeričku analizu
4. Tensorflow Keras za implementaciju konvolucijske mreže.

4.4 Treniranje modela

4.4.1 Treniranje rekurentne neuralne mreže sa LSTM

```

callbacks_list = setup_fast_training()

history = model.fit(
    X_train, y_train,
    epochs=30,
    batch_size=128,

```

```
validation_data=(X_val, y_val),  
class_weight=class_weights,  
callbacks=callbacks_list,  
verbose=1,  
)
```

Model je treniran u 30 epoha, sa veličinom batch-a 128. Korišten je `class_weight.compute_class_weight()` iz `sklearn` da bi se ublažila neuravnoteženost među klasama.

4.4.2 *Treniranje 1D konvolucijske neuralne mreže sa vremenskom dimenzijom*

Treniranje 1D konvolucijske neuralne mreže korišten je Keras iz TensorFlow biblioteke. Korišteni su sljedeći parametri treniranja:

1. Optimizator: Adam (learning rate = 1e-4)
2. Funkcija gubitka: Binary cross-entropy
3. Metrika: Accuracy
4. Broj epoha: 20
5. Veličina batch-a: 32
6. Validacija: nad posebnim validacijskim skupom (X_val, y_val), 20% ukupnih podataka
7. Callback: TerminateOnNaN - ako se tokom treniranja pojavi NaN vrijednost u gubitku, trening se zaustavlja.

Model je treniran nad vremenskim prozorima (sliding windows), što omogućava CNN-u da uoči obrasce u vremenskom toku signala (npr. Dinamiku promjena u širini zjenica, tačke gledišta i slično). Prije treniranja vršena je standardizacija korištenjem `StandardScaler`. Ovakva arhitektura omogućila je automatsko učenje obrazaca u vremenskim sekvencama - npr. Tipične promjene u pokretim oka koje zavise od klase.

4.4.3 *Treniranje Random Forest modela*

Za treniranje Random Forest modela korištena je biblioteka `Scikit-learn`, odnosno njena klasa `RandomForestClassifier`. Model je obučen nad setom podataka za trening koji sadrži agregirane podatke za seriju pokreta oka. Parametri modela su navedeni u prethodnom odjeljku. Model je treniran nad skupom X_train, y_train bez potrebe za dodatnim skaliranjem. Validacija je vršena nad skupom X_val, y_val, a performanse modela prikazane su kroz tačnost, matricu konfuzije i klasifikacijski izvještaj (precision, recall, F1-score).

4.4.4 Treniranje MLP klasifikatora

Višeslojni perceptron (MLP) treniran je koristeći MLPClassifier klasu iz scikit-learn biblioteke.

Parametri treniranja su sljedeći:

1. Optimizator: Adam
2. Regularizacija: L2 regularizacija (alpha = 0.001)
3. Maksimalan broj epoha: 800
4. Early stopping: omogućen; trening se zaustavlja automatski ako nema poboljšanja na validacijskom skupu nakon 15 epoha, koristeći 20% trening podataka kao validacioni skup
5. Standardizacija ulaza: korišten je StandardScaler, što može biti korisno za MLP jer neuralne mreže bolje konvergiraju kada su ulazne vrijednosti centrirane i normalizirane.

4.5 Evaluacija modela

Nakon treniranja navedenih modela, izvršena je evaluacija sve tri metode i poređenje tačnosti i ostalih metrika za ovakvu vrstu binarnog klasifikatora.

4.5.1 Testiranje rekurentne neuralne mreže sa LSTM

Na validacijskom skupu ovaj model postigao je tačnost od 81.22%. Pri evaluaciji ovog modela na testnim podacima postignuti su sljedeći rezultati:

Test Loss: 0.4008

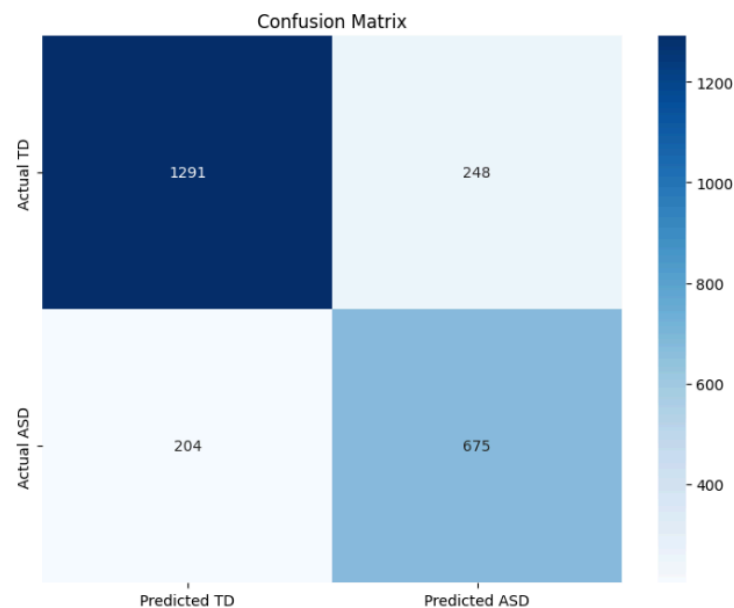
Test Accuracy: 0.8131

Test AUC: 0.8940

	precision	recall	f1-score	support
TD	0.86	0.84	0.85	1539
ASD	0.73	0.77	0.75	879
accuracy			0.81	2418
macro avg	0.80	0.80	0.80	2418
weighted avg	0.82	0.81	0.81	2418

Postignuta je tačnost od 81.31% na testnom skupu, čime se može zaključiti da model dobro generalizira, sa blagom prednosti u detekciji tipično razvijenih osoba, što je i za očekivati. Preciznost postignuta za tu klasu je 86%, dok je

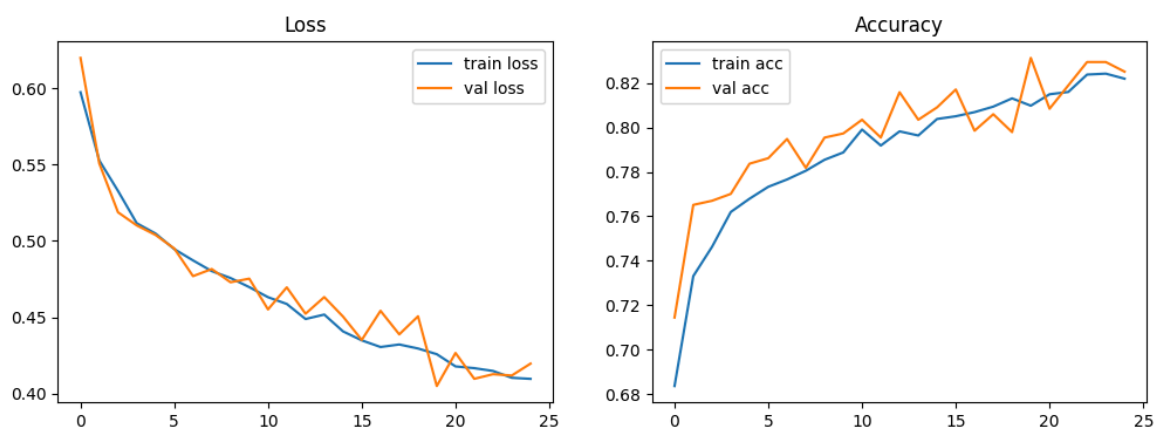
preciznost za ASD klasu 73%. AUC za ovaj model iznosi 89.4%. Matrica zbunjenosti izgleda ovako:



4.5.2 Testiranje 1D konvolucijske neuralne mreže

Na validacijskom skupu ovaj model postigao je tačnost od 82.51%. Ova mreža testirana je na izdvojenom testnom skupu i njeni rezultati su sljedeći:

Test Loss: 0.3968
Test Accuracy: 0.8263
Test Precision: 0.7391
Test Recall: 0.7486



Nažalost, unatoč prihvatljivoj tačnosti na testnom skupu ovaj model se nije pokazao stabilnim.

4.5.3 Testiranje Random Forest modela

Rezultati Random Forest modela su za validacijski skup izgledali ovako:

Validation Performance:

Accuracy: 0.8462

Classification Report:

	precision	recall	f1-score	support
0	0.85	0.85	0.85	128
1	0.84	0.84	0.84	119
accuracy			0.85	247
macro avg	0.85	0.85	0.85	247
weighted avg	0.85	0.85	0.85	247

Confusion Matrix:

```
[[109  19]
 [ 19 100]]
```

Rezultati ovog modela su za testni skup izgledali ovako:

Test Performance (Final Evaluation on Unseen Data):

Accuracy: 0.6606

Classification Report:

	precision	recall	f1-score	support
0	0.66	0.79	0.72	179
1	0.67	0.50	0.57	148
accuracy			0.66	327
macro avg	0.66	0.65	0.65	327
weighted avg	0.66	0.66	0.65	327

Confusion Matrix:

```
[[142  37]
 [ 74  74]]
```

Nažalost, iako model postiže zadovoljavajuću tačnost na validacijskom skupu, može se primijetiti značajan *overfitting* na testnom skupu, tj. da model ne generalizira najbolje na nepoznatim podacima.

Kako bismo riješili problem overfittinga, primijenili smo GridSearch optimizaciju hiperparametara, koristeći F1-score kao metriku za evaluaciju. Ova metrika omogućuje objektivniju procjenu performansi u odnosu na samu tačnost, posebno

u slučajevima neravnomjernog raspodjeljivanja klasa. Odabrani su sljedeći parametri za optimizaciju:

```
param_grid = {
    'n_estimators': [100, 150],
    'max_depth': [5, 7],
    'min_samples_split': [2, 5],
    'min_samples_leaf': [1, 2],
    'class_weight': ['balanced']
}
```

Nakon definisanja skupa vrijednosti hiperparametara za Random Forest klasifikator, korišten je **GridSearchCV** iz biblioteke **scikit-learn** s ciljem pronalaska optimalne kombinacije parametara. Ova metoda sistematski ispituje sve moguće kombinacije zadanih hiperparametara, pri čemu za svaku od njih trenira model koristeći k-struku unakrsnu validaciju (u ovom slučaju 5-struku).

Tokom svakog kruga validacije, model se trenira na četiri dijela trening skupa, dok se peti koristi za evaluaciju. Na taj način, **GridSearchCV** omogućava pouzdanu procjenu performansi za svaku kombinaciju hiperparametara.

```
grid_search = GridSearchCV(
    estimator=rf,
    param_grid=param_grid,
    cv=5,
    scoring='f1_weighted',
    verbose=2,
    n_jobs=-1
)
```

Nakon optimizacije hiperparametara pomoću **GridSearchCV**, dobijen je novi Random Forest model. Na validacionom skupu zabilježen je blagi pad u ukupnoj tačnosti, koja je smanjena s 0.8462 na 0.8057. Ipak, vidljivo je da novi model ostvaruje bolju ravnotežu između klasa, pri čemu je F1-score za klasu 1 porastao s 0.84 na 0.81, a recall sa 0.84 na 0.85.

Validation Performance:

Accuracy: 0.8057

Classification Report:

	precision	recall	f1-score	support
0	0.84	0.77	0.80	128
1	0.77	0.85	0.81	119
accuracy			0.81	247
macro avg	0.81	0.81	0.81	247
weighted avg	0.81	0.81	0.81	247

Confusion Matrix:

```
[[ 98  30]
```

```
[ 18 101]]
```

Nasuprot tome, rezultati na testnom skupu su se značajno poboljšali: tačnost je porasla sa 0.6606 na 0.7003, a F1-score za klasu 1 sa 0.57 na 0.63. Takođe, broj tačno klasifikovanih instanci klase 1 porastao je sa 74 na 82. Ovi rezultati ukazuju na bolju generalizaciju optimizovanog modela na neviđenim podacima, uprkos nešto slabijoj tačnosti na validacijskom skupu.

Test Performance (Final Evaluation on Unseen Data):

Accuracy: 0.7003

Classification Report:

	precision	recall	f1-score	support
0	0.69	0.82	0.75	179
1	0.72	0.55	0.63	148
accuracy			0.70	327
macro avg	0.70	0.69	0.69	327
weighted avg	0.70	0.70	0.69	327

Confusion Matrix:

```
[[147  32]
 [ 66  82]]
```

4.5.4 Treniranje MLP klasifikatora

MLP klasifikator je na validacijskom skupu postigao tačnost od 78.95%, dok su za testni skup rezultati bili sljedeći:

Test Accuracy: 0.7156

Test Classification Report:

	precision	recall	f1-score	support
0	0.70	0.83	0.76	179
1	0.74	0.57	0.65	148
accuracy			0.72	327
macro avg	0.72	0.70	0.70	327
weighted avg	0.72	0.72	0.71	327

Test Confusion Matrix:

```
[[149  30]
 [ 63  85]]
```

U ovom primjeru je postignuta manja tačnost na validacijskom skupu, ali može se primijetiti da je overfitting manji u odnosu na prethodni primjer, odnosno da je razlika između ova dva modela ~7%.

4.6 Implementacija za djecu

Budući da je RNN sa LSTM pokazao najbolje rezultate, u nastavku će biti prikazana implementacija upotrebe ovog modela u svrhe dijagnostike. Primjena ovog modela ostvarena je u vidu prosljeđivanja dokumenata .csv formata koji sadrže vremenske uzorke o pokretima očiju ispitanika. Takva datoteka prosljeđuje se modelu koji daje odgovor o tome da li taj uzorak pripada ASD (Autism Spectrum Disorder) ili TD (Typical Development) klasi. Za implementaciju je korišten Streamlit zbog jednostavne implementacije.

Eye-Tracking Autism Detection Classification

Disclaimer: This is a student project created for educational purposes only. It is *not* intended to be a medical diagnosis or substitute for professional medical advice. Please consult a qualified healthcare professional for any medical concerns.

Upload a CSV file with the same structure as training data:

Choose a CSV file



Drag and drop file here
Limit 200MB per file • CSV

Browse files

5. Zaključak

5.1 Analiza postignuća

Kada su u pitanju postignuća ovog istraživanja, testirane su različite metode za implementaciju modela za detekciju autizma na osnovu pravca pogleda, od čega se najboljom pokazala rekurentna neuralna mreža sa *long short-term* memorijom. Ovakva vrsta modela bolje modelira **ambijent i tok pažnje** kod korisnika, i daje robusnije reprezentacije ponašanja. Model je treniran sa značajkama kao što su širina zjenica, koordinate pogleda i vremenski period gledanja. Uzimajući u obzir da je ovaj dataset limitiran na samo 59 ispitanika i da postoji dosta neispravnih unosa sa NaN vrijednostima, može se zaključiti da je tačnost od ~82% prihvatljiva za takav skup. Ovakav model mogu koristiti

doktori za dalju analizu pacijenata kod kojih se sumnja da imaju autizam, kao i u ranoj detekciji ovog poremećaja.

5.2 Poređenje s postojećim rješenjima

U poređenju s postojećim pristupima za detekciju autizma na osnovu eye-tracking podataka, naš LSTM model pokazuje konkurentne rezultate. Tradicionalni pristupi korišćenjem Support Vector Machine (SVM) i Random Forest algoritama na sličnim skupovima podataka postižu tačnost između 75-80%, dok duboki neuronski modeli primjenjeni na eye-tracking podatke uglavnom dostižu tačnost oko 78-85%. Naš pristup s rekurentnim neuronskim mrežama postiže tačnost od ~82%, što ga pozicionira u gornji dio opsega postojećih rješenja. Ključna prednost RNN modela leži u mogućnosti modeliranja vremenskih sekvenci ponašanja pogleda kroz LSTM arhitekturu, za razliku od tradicionalnih pristupa koji se oslanjaju na statičke mjere poput ukupnog vremena fiksacije ili broja sakada. Međutim, ograničenja našeg istraživanja uključuju relativno mali dataset od 59 ispitanika i prisustvo NaN vrijednosti, što može utjecati na generalnost modela u odnosu na veća istraživanja u oblasti. Unatoč tome, praktična primjenjivost modela kao podrške dijagnostici u kliničkim postavkama predstavlja značajan korak prema objektivizaciji procesa detekcije autizma u ranim stadijumima.

5.3 Moguća poboljšanja

Proširenje skupa podataka: Trenutni dataset od 59 ispitanika predstavlja značajno ograničenje za generalnost modela. Buduća istraživanja trebalo bi usmjeriti na prikupljanje većeg i raznolikog skupa podataka koji uključuje različite uzrasne grupe, etničke pozadine i stepene težine autizma spektar poremećaja.

Poboljšanje kvaliteta podataka: Prisutnost NaN vrijednosti u trenutnom skupu podataka utiče na performanse modela. Implementacija naprednijih tehnika za interpolaciju nedostajućih vrijednosti ili korišćenje sofisticiranih metoda za čišćenje podataka moglo bi značajno poboljšati tačnost.

Arhitekturna poboljšanja: Eksperimentiranje s hibridnim arhitekturama koje kombinuju LSTM sa konvolucijskim neuronskim mrežama (CNN-LSTM) ili implementiranje attention mehanizma moglo bi dodatno poboljšati sposobnost modela da identifikuje ključne obrasce u eye-tracking sekvencama.

Proširenje skupa značajki: Trenutni model koristi širinu zjenica, koordinate pogleda i vremenske periode. Dodavanje dodatnih značajki kao što su brzina sakada, trajanje fiksacije, ili mikroekspresije lica moglo bi obogatiti model.

Real-time implementacija: Optimizacija modela za rad u realnom vremenu omogućila bi praktičnu primjenu u kliničkim postavkama tokom dijagnostičkih sesija.

Interpretabilnost modela: Implementacija tehnika poput LIME ili SHAP za objašnjavanje odluka modela omogućila bi liječnicima bolje razumijevanje osnova za automatske dijagnoze.

Multimodalni pristup: Kombinovanje eye-tracking podataka s drugim modalitetima kao što su audio analize govora, EEG signali ili behavioralne mjere moglo bi značajno poboljšati dijagnostičku tačnost.

5.4 Društveni utjecaj

Implementacija ovog modela u kliničku praksu može imati značajan društveni utjecaj kroz poboljšanje rane detekcije autizma, što je ključno za blagovremenu intervenciju i bolje razvojne ishode. Automatizovani screening može skratiti trenutna dijagnostička kašnjenja koja često traju mjesecima ili godinama, omogućavajući brži pristup potrebnim uslugama za djecu i njihove porodice. Model također omogućava objektivn i standardizovan pristup dijagnostici bez obzira na geografsku lokaciju ili dostupnost specijalizovanih stručnjaka, što je posebno važno u ruralnim područjima i zemljama s ograničenim resursima. Ekonomski benefiti uključuju značajne uštede u dugoročnim troškovima podrške, s obzirom da rana intervencija može uštedjeti 7-10 dolara u budućim troškovima za svaki dolar ulož en u ranu dijagnostiku. Korištenje objektivnih tehnoloških alata može doprinijeti destigmatizaciji autizma i promovirati razumijevanje neuroloških razlika, dok istovremeno omogućava bolje prilagođene edukacijske programe i podršku porodicama. Važno je naglasiti da ovaj model treba služiti kao pomoć stručnjacima u donošenju dijagnoza, a ne kao njihova zamjena, uz potrebu za odgovornim korišćenjem tehnologije s poštovanjem privatnosti i prava osoba s autizmom.

Literatura

- [1] “Eye Gaze,” *Callscotland.org.uk*, 2025.
<https://www.callscotland.org.uk/Information/eye-gaze/>
- [2] Tobii, “What is eye tracking? | How does eye tracking work,” *www.tobii.com*, Apr. 23, 2025.
<https://www.tobii.com/learn-and-support/get-started/what-is-eye-tracking>
- [3] National Institute of Mental Health, “Autism Spectrum Disorder,” *National Institute of Mental Health (NIMH)*, 2024.
<https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd>
- [4] National Institute of Mental Health, “Attention-Deficit/Hyperactivity disorder,” *www.nimh.nih.gov*, Dec. 2024.
<https://www.nimh.nih.gov/health/topics/attention-deficit-hyperactivity-disorder-adhd>
- [5] American Speech-Language-Hearing Association, “Augmentative and Alternative Communication (AAC),” *Asha.org*, 2025.
<https://www.asha.org/public/speech/disorders/aac> (accessed Jun. 01, 2025).
- [6] K. Akre, “nonverbal communication,” *www.britannica.com*, Oct. 21, 2022.
<https://www.britannica.com/topic/nonverbal-communication>
- [7] K. Watson, “What Is Sensory Overload?,” *Healthline*, Sep. 27, 2018.
<https://www.healthline.com/health/sensory-overload>
- [8] M. Tahri Sqalli, B. Aslonov, M. Gafurov, N. Mukhammadiev, and Y. Sqalli Houssaini, “Eye Tracking Technology in Medical practice: a Perspective on Its Diverse Applications,” *Frontiers in medical technology*, vol. 5, Nov. 2023, doi: <https://doi.org/10.3389/fmedt.2023.1253001>.
- [9] N. Dostalova *et al.*, “ETDD70: Eye-Tracking Dyslexia Dataset,” *Zenodo*, Aug. 2024, doi: <https://doi.org/10.5281/zenodo.13332134>.
- [10] H. Duan *et al.*, “[Saliency4ASD] A dataset of eye movements for the children with autism spectrum disorder,” *Zenodo*, Jun. 2019, doi: <https://doi.org/10.5281/zenodo.13960426>.
- [11] S. Tafasca, A. Gupta, and J.-M. Odobez, “ChildPlay: A New Benchmark for Understanding Children’s Gaze Behaviour,” *arXiv.org*, 2023.
<https://arxiv.org/abs/2307.01630> (accessed Jun. 02, 2025).

- [12] M. Zettersten *et al.*, “Peekbank: An open, large-scale repository for developmental eye-tracking data of children’s word recognition,” *Behavior Research Methods*, Aug. 2022, doi: <https://doi.org/10.3758/s13428-022-01906-4>.
- [13] P. Pathirana, S. Senarath, D. Meedeniya, and S. Jayarathna, “Eye gaze estimation: A survey on deep learning-based approaches,” *Expert Systems with Applications*, vol. 199, p. 116894, Aug. 2022, doi: <https://doi.org/10.1016/j.eswa.2022.116894>.
- [14] A. Hassoumi, V. Peysakhovich, and C. Hurter, “Uncertainty visualization of gaze estimation to support operator-controlled calibration,” Jan. 2017, doi: <https://doi.org/10.16910/jemr.10.5.6>.
- [15] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, “MPIIGaze: Real-World Dataset and Deep Appearance-Based Gaze Estimation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 1, pp. 162–175, Jan. 2019, doi: <https://doi.org/10.1109/tpami.2017.2778103>.
- [16] K. Alberto Funes Mora, F. Monay, and J.-M. Odobez, “EYEDIAP, A Database for the Development and Evaluation of Gaze Estimation Algorithms from RGB and RGB-D Cameras,” *ACM Symposium on Eye Tracking Research and Applications*, Mar. 2014, doi: <https://doi.org/10.1145/2578153.2578190>.
- [17] K. Krafka *et al.*, “Eye Tracking for Everyone,” 2016. Available: https://gazeCapture.csail.mit.edu/cvpr2016_gazeCapture.pdf