

# Raspoznavanje visokorizičnih mentalnih stanja pilota na temelju obrade fizioloških signala metodama dubokog učenja

---

Nekić, Sven

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:168:262339>

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

Download date / Datum preuzimanja: **2025-03-30**



Repository / Repozitorij:

[FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repository](#)



SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 402

**RASPOZNAVANJE VISOKORIZIČNIH MENTALNIH STANJA  
PILOTA NA TEMELJU OBRADJE FIZIOLOŠKIH SIGNALA  
METODAMA DUBOKOG UČENJA**

Sven Nekić

Zagreb, lipanj 2024.

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 402

**RASPOZNAVANJE VISOKORIZIČNIH MENTALNIH STANJA  
PILOTA NA TEMELJU OBRADJE FIZIOLOŠKIH SIGNALA  
METODAMA DUBOKOG UČENJA**

Sven Nekić

Zagreb, lipanj 2024.

## DIPLOMSKI ZADATAK br. 402

Pristupnik: **Sven Nekić (0036524893)**  
Studij: Računarstvo  
Profil: Računarska znanost  
Mentor: izv. prof. dr. sc. Siniša Popović

Zadatak: **Raspoznavanje visokorizičnih mentalnih stanja pilota na temelju obrade fizioloških signala metodama dubokog učenja**

### Opis zadatka:

Proučiti izabranu literaturu o primjenama dubokog učenja u detekciji visokorizičnih mentalnih stanja pilota, kao što su visoka razina stresa, kognitivno preopterećenje, iznenadni prepad (engl. startle) uzrokovan neočekivanom ugrozom itd. Definirati metodologiju obrade fizioloških signala primjenom metoda dubokog učenja u cilju raspoznavanja visokorizičnih stanja pilota u raspoloživom podatkovnom skupu snimljenom na simulatoru leta aviona. Ispitati više načina reprezentacije fizioloških podataka u oblik ulaznog vektora za modele dubokog učenja, koji podrazumijevaju različite metode predobrade snimljenih podataka. Provesti komparativne analize nekoliko izabranih metoda dubokog učenja na problemima binarnog i višeklasnog raspoznavanja promatranih visokorizičnih stanja pilota. U komparativnim analizama primijeniti uobičajene evaluacijske metrike modela dubokog učenja, kao što su konfuzijske matrice, točnost, preciznost, odziv, ROC krivulje itd.

Rok za predaju rada: 28. lipnja 2024.

*Ovim putem želio bih zahvaliti mentoru izv. prof. dr. sc. Siniši Popoviću na potpori u pisanju diplomskog rada. Također zahvaljujem svojoj obitelji na potpori tijekom studija.*

# SADRŽAJ

<b>1. Uvod</b>	<b>1</b>
<b>2. Osvrt na srodnu literaturu</b>	<b>2</b>
<b>3. Opis raspoloživog podatkovnog skupa</b>	<b>4</b>
<b>4. Metodologija razvoja dubokih modela za raspoznavanje visokorizičnih mentalnih stanja pilota</b>	<b>8</b>
4.1. ResNet . . . . .	11
4.1.1. Rezidualno učenje . . . . .	11
4.1.2. ResNet autoenkoder i klasifikator . . . . .	12
4.1.3. ResNet klasifikator . . . . .	13
4.2. U-NET 1D CNN . . . . .	14
<b>5. Rezultati i diskusija</b>	<b>16</b>
5.1. Primijenjene evaluacijske mjere . . . . .	16
5.2. ResNet autoenkoder i klasifikator . . . . .	18
5.2.1. Autoenkoder . . . . .	18
5.2.2. Klasifikator . . . . .	19
5.3. ResNet klasifikator . . . . .	20
5.4. U-NET 1D CNN . . . . .	21
<b>6. Zaključak</b>	<b>23</b>
<b>Literatura</b>	<b>26</b>

# 1. Uvod

Sigurnost u zrakoplovstvu ključna je briga, a raspoznavanje mentalnog stanja pilota tijekom letačkih operacija igra ključnu ulogu u osiguravanju sigurnosti [1]. Piloti se često nalaze u situacijama visokog rizika, što uzrokuje stres, kognitivno preopterećenje i iznenadne reakcije preplašenosti zbog neočekivanih prijetnji. Brzo otkrivanje ovih stanja može značajno povećati sigurnost leta i spriječiti potencijalne nesreće. Brojna istraživanja usredotočila su se na praćenje mentalnog i fizičkog stanja pilota koristeći razna fiziološka mjerenja poput otkucaja srca, disanja i elektrodermalne aktivnosti [2]. Središnji zadatak uključuje prepoznavanje i klasificiranje mentalnih stanja visokog rizika na temelju fizioloških signala dobivenih tijekom simulacija letenja. Ovi signali uhvaćeni su putem raznih senzora i daju dragocjen uvid u psihičko stanje pilota. Međutim, tumačenje ovih signala složen je zadatak.

Osiguravanje dobrobiti pilota izravno utječe na sigurnost putnika. Mentalno stanje pilota utječe na donošenje odluka, svijest o situaciji i ukupnoj izvedbi u letu. Rano otkrivanje mentalnih stanja visokog rizika omogućuje pravovremenu intervenciju, što može spriječiti incidente i nesreće. Osim sigurnosti, raspoznavanje mentalnog opterećenja pilota i razina stresa može optimizirati letačke operacije. Prepoznavanjem trenutaka kognitivnog preopterećenja ili prekomjernog stresa, zrakoplovne tvrtke mogu prilagoditi programe obuke, poboljšati ukupnu učinkovitost pilota. Pristup usmjeren na metode dubokog učenja za obradu fizioloških signala mogao bi pomoći u rješavanju ovih problema.

Koristit će se fiziološki podaci iz skupa podataka Flight Crew Physiological Data for Crew State Monitoring (NASA) koji su dobiveni tijekom simulacija leta. Ti podaci sadrže očitavanja elektrokardiograma (EKG), očitavanja elektroencefalograma (EEG), respiraciju te galvanski kožni odgovor. Cilj je prvo smanjiti šum u podacima koristeći različite tehnike obrade signala, poput frekvencijske reprezentacije i vremensko-frekvencijske reprezentacije. Nakon toga podaci će se koristiti kao ulazni vektor za različite duboke modele strojnog učenja. Izlazi iz modela, tj. klasifikacija bit će u nekoliko varijanti: višestruka klasifikacija (sva moguća kognitivna stanja), binarna klasifikacija "No Event" ili rizična stanja, ternarna klasifikacija No Event ili SS (Startle/Surprise) ili ostala rizična stanja.

Koristit će se uobičajene mjere za evaluaciju modela kao što su matrice zabune, točnost, preciznost, odziv itd.

## 2. Osvrt na srodnu literaturu

Autori Terwilliger, Paul i ostali predviđali su kognitivna stanja koristeći tehnike strojnog učenja, tako da ih primijene na vremenske okvire podataka [3]. Koriste skup podataka Flight Crew Physiological Data for Crew State Monitoring (NASA).

Kako bi se simulirali scenariji iz stvarnog života, predviđanja se rade samo za buduća stanja. Iz četiri izvorne klase (SS, CA, DA, NE) klasificira se "Događaj" ili "Nema događaja".

Shvatili su da klasifikacija svakog vremenskog okvira bez ikakve obrade ulaznih podataka dovodi do prenaučenosti, pa su pokušali predvidjeti kognitivna stanja koristeći samo jedan vremenski okvir s izgladivanjem signala i plitkom neuronskom mrežom, ali nisu postigli uspjeh. Zbog preklapajućih vremenskih okvira, autori su pokušali koristiti konvolucijske neuronske mreže (CNN) za klasifikaciju vremenskih prozora. S velikim vremenskim prozorima ( $N=128$ ) događa se prenaučenost jer neuronska mreža vidi svaku vremensku oznaku više puta po epohi. Autori zatim koriste Residual Network (ResNet) autoenkoder za treniranje na latentnoj reprezentaciji vremenskih prozora, čime su izbjegli problem prenaučenosti.

Autori uspoređuju performanse triju modela: ResNet klasifikator, gusti klasifikator i ResNet autoenkoder-klasifikator. ResNet klasifikator pokazao je prenaučenost, a gusti klasifikator rješavao je ovaj problem treniranjem na pojedinačnim vremenskim okvirima. Nadalje, ResNet klasifikator nadmašio je gusti klasifikator u smislu gubitka (engl. loss) validacije, naglašavajući važnost uključivanja prošlih i budućih vremenskih okvira za moć predviđanja. ResNet autoenkoder-klasifikator postigao je najbolje rezultate, učeći osam puta više epoha i postigavši manji gubitak od druga dva modela (0.34) naspram ResNet klasifikatora (0.47) i gustog klasifikatora (0.57).

Autori Han, Soo-Yeon i ostali istraživali su mogućnost detekcije mentalnih stanja pilota pomoću multimodalnih biosignala (tj. EEG, EKG, respiracija i elektrodermalna aktivnost) [4] i multimodalnog dubokog učenja. Pristup je višeklasna klasifikacija između mentalnih stanja normalnosti, distrakcije, radnog opterećenja te umora. Osam zdravih ispitanika sudjelovalo je u simulatoru leta više od 100 sati. Eksperimenti su koristili sustav simulatora leta opremljen raznim kontrolama i bežičnom tipkovnicom za snimanje odgovora pilota.

Za EEG signale koristili su CNN arhitekturu s tri skrivena sloja (tj. dva konvolucijska sloja i potpuno povezani sloj). Za EKG, respiraciju i elektrodermalnu aktivnost koristili su



LSTM modele. U radu za signale koji nisu EEG koriste nazivlje PPM (engl. peripheral physiological measures)

Rezultati klasifikacije temeljeni samo na EEG-u nadmašili su rezultate temeljene samo na PPM-ovima (EKG, respiracija i elektrodermalna aktivnost). Također, korištenje EEG-a i PPM-ova zajedno općenito je dalo bolje rezultate nego korištenje bilo kojeg od ovih podataka pojedinačno. Model temeljen na EEG i PPM signalima na višeklasnoj klasifikaciji postigao je najveću točnost ( $85.2\% \pm 4.3\%$ ) u usporedbi s drugim modelima.

Autori AR Harrivel i drugi željeli su istražiti točnost predviđanja tzv. AHPLS (engl. attention-related human performance limiting states) korištenjem multimodalnih psihofizioloških mjera kao ulaznih značajki za modele klasifikatora [5]. Dvadeset i četiri pilota sudjelovala su u istraživanju. Nosili su senzore dok su obavljali zadatke u simulatoru leta koji su mjerili EEG, EKG, respiraciju, GSR i fNIRS (funkcijsku blisku infracrvenu spektroskopiju). Zadaci su izazivali AHPLS i sadržavali su varijacije radnog opterećenja.

Sirovi fiziološki podaci pretprocesirani su kako bi se uklonili šumovi.

Ti su podaci zatim korišteni kao ulazne značajke za treniranje nekoliko klasifikatora strojnog učenja; autori su implementirali XGBoost, slučajne šume te dva različita modela SVM-a.

Rezultati su pokazali da korištenje svih signala s odvojenim treniranjem po modalitetu daje najbolju točnost (95% s EEG-om, respiracijom i GSR-om). Multimodalno treniranje s EKG-om i GSR-om također je dalo visoku točnost (93% s EKG-om i GSR-om).

Multimodalno treniranje općenito daje nešto bolje performansi klasifikatora u usporedbi s treniranjem zasebnih klasifikatora specifičnih za modalitet ( $64\% \pm 14\%$  naspram  $57\% \pm 9\%$ ).

Kombiniranje modaliteta u treniranju omogućuje korištenje sinergijskih ili međusobnih informacija, dok odvojeno treniranje nudi robusnije rješenje u slučaju gubitka signala u jednom modalitetu.

### 3. Opis raspoloživog podatkovnog skupa

Skup podataka Flight Crew Physiological Data for Crew State Monitoring (NASA) koji će se koristiti u ovom radu sadrži stvarne fiziološke podatke osamnaest pilota.

Podaci su prikupljeni od parova pilota i kopilota u simulatoru letenja (LOFT - engl. Line Oriented Flight Training) i izvan njega, a služit će za treniranje dubokih modela strojnog učenja za raspoznavanje visokorizičnih mentalnih stanja pilota. Referentni podaci (datoteke koje završavaju na CA, DA i SS) prikupljeni su izvan simulatora letenja. S druge strane, LOFT simulacija sastoji se od polijetanja, leta i slijetanja u simulatoru leta.

Događaji koje su subjekti doživjeli trebali su potaknuti jedno od slijedeća tri kognitivna stanja: kanalizirana pozornost (CA - engl. Channelized Attention); preusmjerena pozornost (DA - engl. Diverted Attention) i zaprepaštenje/iznenađenje (SS - engl. Startle/Surprise).

Kanalizirana pozornost je, grubo rečeno, usredotočenost na jedan zadatak. Subjekti igraju zanimljivu video igru temeljenu na “zagonetkama” (engl. puzzles), a pojedini eksperiment traje otprilike 360 sekundi, tj. 5 minuta. Preusmjerena pažnja stanje je u kojem je pažnja subjekta skrenuta ili ometana. To je inducirano tako da se povremeno pojavi matematički problem koji se mora riješiti usred drugog, zahtjevnijeg zadatka. Zaprepaštenje/iznenađenje izaziva se tako što subjekti gledaju filmske isječke sa neočekivanim “jumpscare” scenama. Tijekom LOFT simulacije zadaci također induciraju kognitivna stanja. Postoje dva odvojena događaja dizajnirana da induciraju SS, dva da induciraju CA i dva ili više da induciraju DA stanje, a cijeli eksperiment traje otprilike 4054 sekunde, tj. 67 minuta.

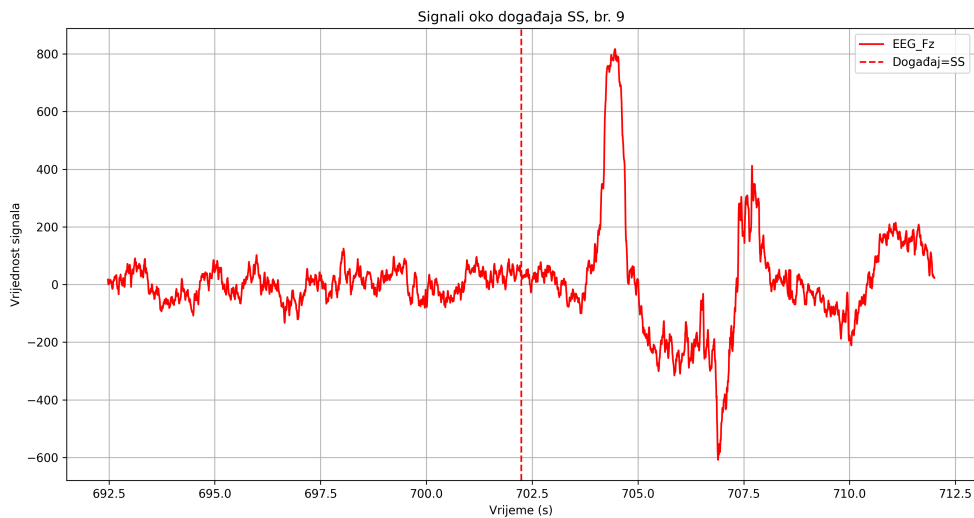
Referentne vrijednosti izvan simulatora leta prikupljeni su prijepodne, a LOFT poslijepodne.

Svaki eksperiment sprema se u posebnu .csv datoteku u kojoj je stupac određeno mjerenje, vrijeme ili oznaka događaja, a red je puni uzorak s mjerenjima sa svakog senzora povezanog s vremenskom oznakom. Datoteka sadrži 25 stupaca podataka; 20 kanala EEG-a (elektroencefalogram), respiracije (mjere širenja i skupljanja prsnog koša), galvanskog kožnog odgovora (mjera elektrodermalne aktivnosti) i EKG-a (elektrokardiogram). Datoteka također sadrži i globalni indeks podatka te globalno vrijeme eksperimenta. Ova dva podatka ne koriste se u modelima te se izbacuju.

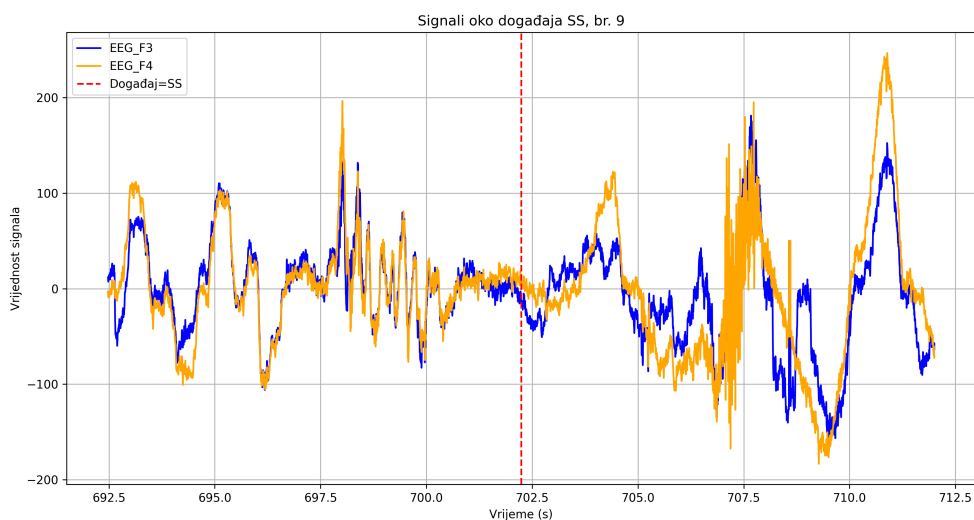
Problem ovog skupa podataka je nebalansiranost klasa. Naime, klasa No Event pred-

stavlja čak 83% podataka. Ostalih 17% predstavljaju ostale klase; kanalizirana pozornost obuhvaća oko 14%, preusmjerena pažnja čini oko 2%, a zaprepaštenje/iznenađenje čini 1% podataka.

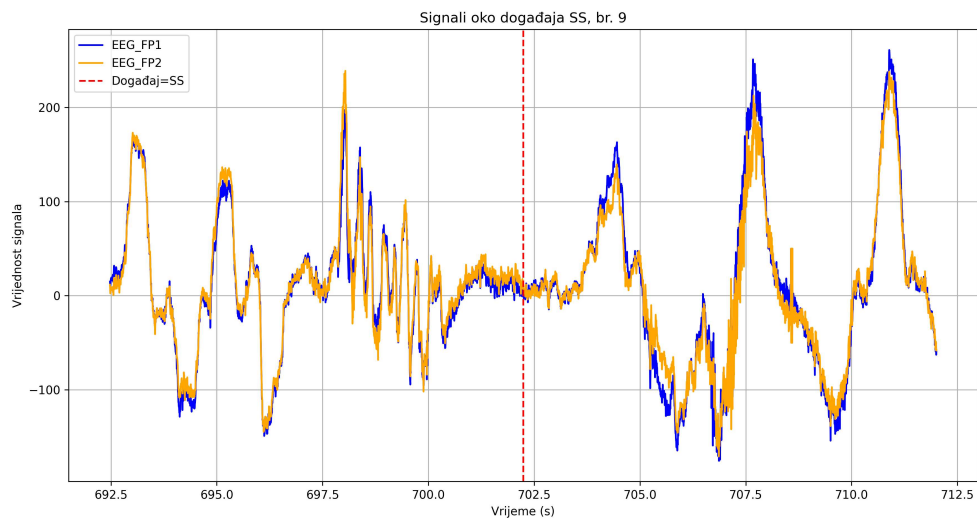
EEG podaci prikupljeni su B-Alert X24 EEG sustavom na glavi subjekta. Senzori imaju razlučivost po bitu od  $.012215 \mu\text{V}$  i raspon od  $-100\text{mV}$  do  $+100\text{mV}$ . Senzori za respiraciju i GSR imaju rezoluciju po bitu od  $.2384186 \mu\text{V}$  i raspon od  $-2,0 \text{ V}$  do  $+2,0 \text{ V}$ . Podaci za EKG, R i GSR prikupljeni su sustavom NeXus-10, a svi senzori prikupljali su podatke na frekvenciji od 256 Hz.



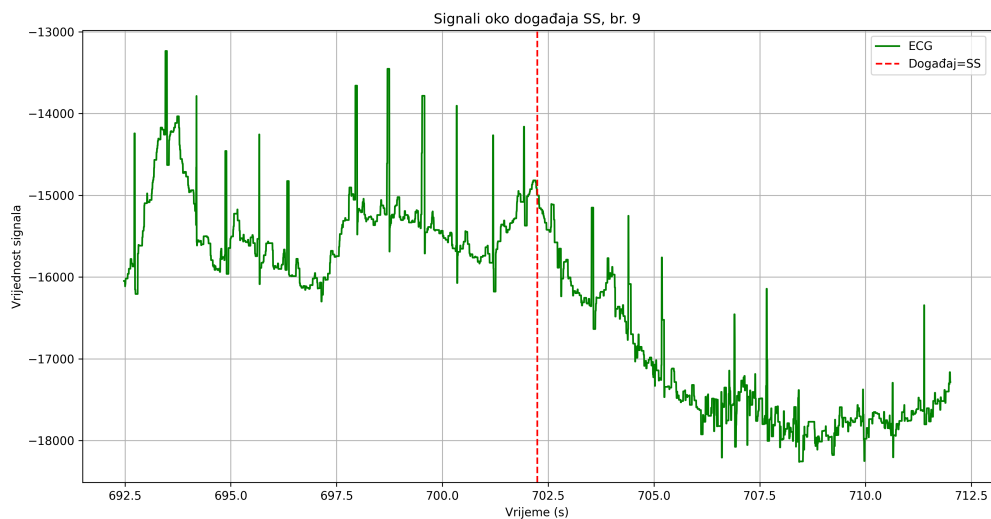
Slika 3.1: Mjerenja EEG Fz kanala



Slika 3.2: Mjerenja EEG F3 i F4 kanala



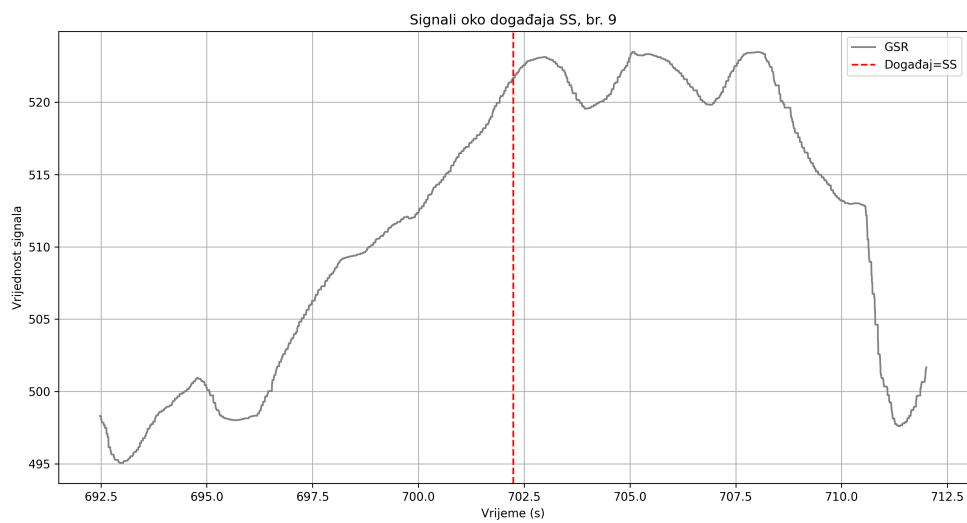
**Slika 3.3:** Mjerenja EEG FP1 i FP2 kanala



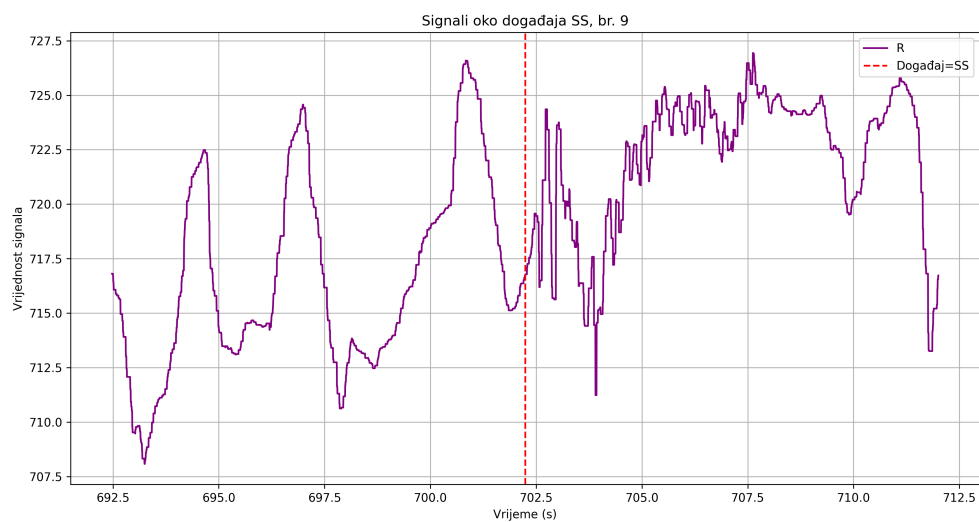
**Slika 3.4:** Mjerenja EKG kanala (elektrokardiogram)

Na grafovima 3.1 - 3.6 prikazane su izmjerene vrijednosti EEG kanala te kanala EKG, galvanskog kožnog odgovora te respiracije.

EEG F3 i F4 kanali mjere aktivnosti lijevog, odnosno desnog frontalnog režnja, Fz kanal mjeri aktivnost središnjeg frontalnog korteksa, a FP1 i FP2 kanali mjere aktivnost prefrontalnog korteksa.



**Slika 3.5:** Mjerenja GSR kanala (galvanski kožni odgovor)



**Slika 3.6:** Mjerenja R kanala (respiracija)

## 4. Metodologija razvoja dubokih modela za raspoznavanje visokorizičnih mentalnih stanja pilota

U ovom radu ispitana su tri duboka modela strojnog učenja vrlo različite arhitekture.

Implementirani su u programskom jeziku Python sa pomoćnim bibliotekama SciKit-Learn, Pandas, Numpy i PyTorch.

Kao skup podataka koristi se prethodno naveden Flight Crew Physiological Data for Crew State Monitoring. Koriste se podaci snimljeni izvan simulatora leta te unutar simulatora leta.

Uzimajući u obzir nebalansiranost podataka, koristit će se podskup podataka dobiven poduzorkovanjem (engl. undersampling). Teoretski, može se koristiti i metoda prekomjernog uzorkovanja (engl. oversampling), ali onda bi morali generirati mnogo podataka koji bi mogli narušiti stabilnost modela.

Svi podaci su normalizirani z-transformacijom.

Svaki signal (EEG, EKG, R, GSR...) normaliziran je z-transformacijom svojeg stupca podataka, na način:

$$Z = \frac{x - \mu(x)}{\sigma(x)} \quad (4.1)$$

gdje je  $\mu(x)$  prosjek, a  $\sigma(x)$  standardna devijacija.

Za treniranje i testiranje modela vrijedi slijedeće:

Za oba ResNet modela normalizirani podaci ulaze u model kao trodimenzionalni vektor veličine [veličina minigrupe, veličina vremenskog prozora, broj signala]. Veličina minigrupe podesiva je ovisno o memoriji računala, a veličina vremenskog prozora je 256, što označava jednu sekundu. Broj signala je 23, što je 20 kanala EEG-a, respiracija, galvanski kožni odgovor i EKG. Pomak između vremenskih prozora je 1, što znači da se vremenski prozori preklapaju na 255 uzoraka svakog od 23 signala. Izlaz ResNet modela daje za binarnu klasifikaciju dvodimenzionalni vektor veličine 2x1 (vjerojatnost Event, tj. No Event), a za višeklasnu klasifikaciju dvodimenzionalni vektor veličine 4x1 (vjerojatnost SS, CA, DA ili

No Event).

Za model U-NET ulaz je isti kao kod ResNet modela, ali izlaz je drukčiji. U-NET model zbog drugačije arhitekture ne računa vjerojatnost jedne klase po vremenskom prozoru, nego računa onoliko vjerojatnosti koliko ima unosa u vremenskom prozoru. Za binarnu klasifikaciju, dvodimenzionalni vektor veličine [veličina vremenskog prozora, 2], te za višeklasnu klasifikaciju dvodimenzionalni vektor veličine [veličina vremenskog prozora, 4].

U-NET i ResNet modeli validirani su metodom sličnom LOSO (engl. Leave one subject out); no umjesto da se iz skupa podataka izdvoji samo jedan pilot za testiranje, izdvaja se njih 5, a ostalih 13 koristi se za treniranje modela. Ovi skupovi za treniranje i testiranje ciklički se izmjenjuju. U daljnjem tekstu ova će se metoda po analogiji označavati kraticom LFSO (engl. Leave five subjects out)." Koristila se metoda LFSO umjesto LOSO zbog velike neravnoteže podataka i manjka podataka klasa CA, DA i SS naspram klase No Event.

Model U-NET dodatno je validiran s K-Fold unakrsnom validacijom zbog znatno drugačije arhitekture od ResNet modela. Koristi se skup podataka tako što su se očitavanja svakog pilota koristila slijedno u obliku pilot1 - [eksperiment CA, eksperiment DA, eksperiment SS, eksperiment LOFT], pilot2 - [eksperiment CA, eksperiment DA, eksperiment SS, eksperiment LOFT] itd. Naravno, svaki od ovih skupova podataka prethodno je poduzorkovan. Fold je 5, što znači da će se skup podataka podijeliti na 5 dijelova bez ikakvih preklapanja, te će se ovih 5 skupova ciklički izmjenjivati za treniranje i testiranje.

80% jednog folda koristi se za skup treniranja, a 20% koristi se za skup testiranja.

Dodatno, u svim modelima koristi se normalizacija podataka u minigrupi (engl. batch normalization) po slijedećoj formuli:

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} \quad (4.2)$$

gdje je:

$x_i$  : Ulazni vektor

$\mu$  : Medijan vektora

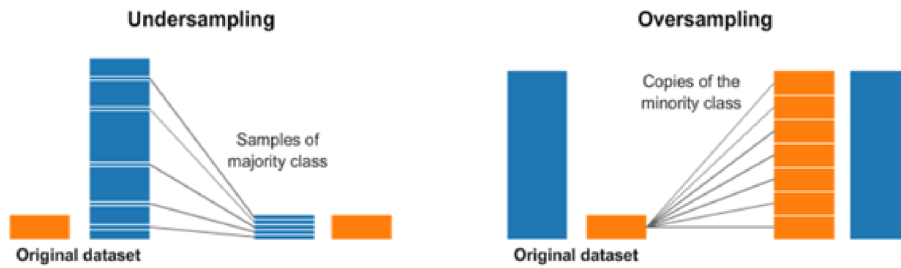
$\sigma^2$  : Varijanca vektora

$\epsilon$  : Mala konstanta dodana kako bi se izbjeglo dijeljenje s nulom

$\hat{x}_i$  : Normalizirani vektor

Slika implicira da su sve klase svakog pilota poduzorkovane da odgovaraju veličini njegove najmanje klase. Iz skupa podataka svakog pojedinog ispitanika slučajno su izbačena mjerenja svih klasa, kako bi reprezentacija klasa bila podjednaka.

Ukoliko radimo binarnu klasifikaciju Event/No Event, klase CA, DA i SS pretvorit ćemo u klasu Event, odnosno 1, a klasu No Event ostavit ćemo kao 0. Poduzorkovanjem će se broj



**Slika 4.1:** Metoda uzorkovanja podataka - preuzeto iz [3]

podataka značajno smanjiti, pogotovo za višeklasnu klasifikaciju, no ovaj korak je nužan kako bismo mogli dobiti modele sa boljim performansama.

Nakon obrade, podaci se šalju u modele kao ulaz te se modeli treniraju. Nakon treniranja modela, provodi se komparativna analiza. Koriste se mnoge uobičajene metrike evaluacije modela, kao što su konfuzijske matrice, točnost, preciznost, odziv itd.

Zbog vremenske kompleksnosti treniranja modela, točnost i gubitak treniranja računa se na podskupu podataka za treniranje.



## 4.1. ResNet

Duboke neuronske mreže revolucionirale su klasifikaciju slika jer su postigle značajan napredak u točnosti naspram modela pliće arhitekture. Nedavna istraživanja naglašavaju ključnu ulogu dubine mreže u poboljšanju performansi, s vodećim modelima koji koriste dubine u rasponu od šesnaest do trideset slojeva [6]. Međutim, jednostavno slaganje više slojeva ne dovodi uvijek do boljih performansi zbog problema poput nestajućih ili eksplodirajućih gradijenata.

### 4.1.1. Rezidualno učenje

Rezidualno učenje koncept je koji pokušava ublažiti poteškoće u treniranju vrlo dubokih neuronskih mreža. Glavna je ideja preformulirati cilj učenja kako bi se usredotočio na rezidualne ciljne funkcije, a ne na samu ciljnu funkciju. Ovaj pristup pojednostavljuje proces optimizacije i pomaže u ublažavanju problema koji su uobičajeni u dubokim mrežama.

U tradicionalnom dubokom učenju, očekuje se da će svaki sloj u neuronskoj mreži naučiti izravno preslikavanje od ulaza do izlaza. Kako se dubina mreže povećava, učenje ovog izravnog mapiranja postaje izazovnije. Rezidualno učenje preformulira ovaj problem tako što svaki sloj uči rezidualnu funkciju. Funkcija reziduala predstavlja razliku između željenog preslikavanja i funkcije identiteta ulaza [6]:

$$H(x) = F(x) + x \quad (4.3)$$

gdje  $H(x)$  predstavlja željeno preslikavanje koju mreža želi naučiti,  $F(x)$  je rezidualna funkcija, a  $x$  je ulaz.

Temeljni građevni blok rezidualnog učenja je rezidualni blok. U ovom radu koristit će se slijedeća arhitektura:

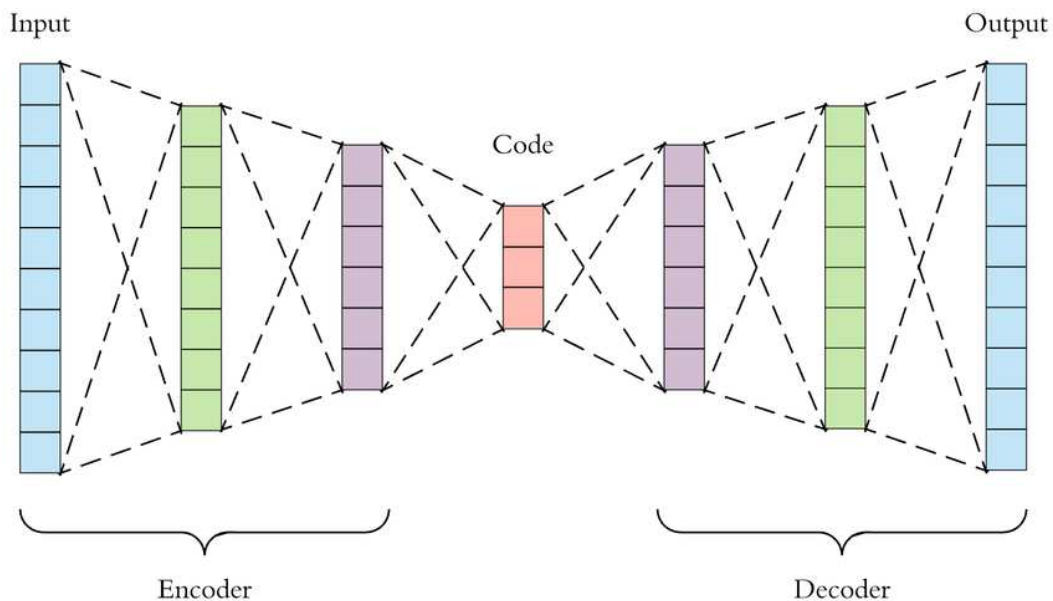
- Konvolucijski slojevi: 1D konvolucijski sloj sa 64 izlazna kanala, veličinom jezgre 9 i korakom od 1 te 1D konvolucijski sloj s istim brojem ulaznih i izlaznih kanala, veličinom jezgre 9 i korakom od 1
- Slojevi normalizacije minigrupe
- Slojevi stiskanja i pobude (engl. Squeeze-Excitation, SE)
- Aktivacijska funkcija Swish

Vektor ulaza  $x$  prolazi kroz prvi konvolucijski sloj koji primjenjuje 1D konvoluciju za izdvajanje lokaliziranih značajki. Izlaz zatim prolazi kroz prvi SE sloj koji ponovno kalibrira odgovore značajki po kanalu. Izlaz se zatim normalizira te se primjenjuje funkcija aktivacije Swish. Izlaz prolazi kroz drugi konvolucijski sloj, a zatim prolazi kroz drugi SE sloj za dodatnu ponovnu kalibraciju značajki po pitanju kanala. Ponovno kalibrirani izlaz prolazi još

jednu normalizaciju serije i Swish aktivaciju.  $x$  se zbraja s izlazom drugog konvolucijskog bloka, čime postizemo rezidualnu vezu.

#### 4.1.2. ResNet autoenkoder i klasifikator

Autoenkoderi sažimaju ulazne podatke u kompaktnu latentnu prostornu reprezentaciju pomoću kodera i zatim ih rekonstruiraju natrag u izvorni oblik pomoću dekodera. Tijekom treniranja nastoje minimizirati "neslaganje" između ulaza i izlaza, usavršavajući svoju sposobnost za prepoznavanje bitnih značajki podataka.



**Slika 4.2:** Ilustracija slične ResNet arhitekture - preuzeto iz [7]

Model Resnet autoenkoder i klasifikator hibridna je arhitektura duboke neuronske mreže koja kombinira rezidualni autoenkoder s klasifikacijskom neuronskom mrežom. U ovom modelu, autoenkoder je baziran na ResNet arhitekturi. Inicijalni sloj jednodimenzionalne konvolucije prima 23-dimenzionalni vektor podataka sa veličinom jezgre = 9. Nakon toga mreža ima 9 sekvencijalnih tzv. Res-blokova. Svaki blok sadrži jednodimenzionalnu konvoluciju, normalizaciju minigrupe te sloj "Squeeze-Excitation" dvaput zaredom. Sloj Squeeze-Excitation (SE) osmišljen je za poboljšanje reprezentativne moći neuronske mreže eksplisnim modeliranjem međuovisnosti između kanala njezine konvolucijske mreže. SE sloj izvodi dinamičku ponovnu kalibraciju značajki po kanalu, što pomaže mreži da selektivno naglasi informativne značajke i potisne one manje korisne. Na kraju, aktivacijska funkcija je Swish, odnosno:

$$\text{Swish}(x) = x \cdot \sigma(x) \quad (4.4)$$

gdje je  $\sigma(x)$  sigmoidalna funkcija,  $x$  je vektor ulaza. Kada je  $x$  vektor, sigmoidalna funkcija primjenjuje se na svaki element, a produkt Swish funkcije je umnožak element po

element (engl. element-wise).

Za razliku od tradicionalnih aktivacijskih funkcija kao što je ReLU (Rectified Linear Unit), Swish je glatka i nemonotona. Ova glatkoća može pomoći u optimizaciji temeljenoj na gradijentima budući da pruža nijansiranije gradijente, što može dovesti do učinkovitijeg i stabilnijeg treniranja neuronske mreže.

Nakon Res-blokova slijedi još jedna jednodimenzionalna konvolucija te Swish aktivacijska funkcija. Izlaz ovog autoenkodera prosljeđuje se u klasifikatorsku potpuno povezanu mrežu. Ulaz u klasifikator je vektor duljine  $8 \times 64$  te se on prosljeđuje kroz tri potpuno povezana sloja duljine  $512 \times 1$ , a izlaz mreže je vektor duljine  $2 \times 1$  koji predstavlja vjerojatnost klase Event ili No Event (0 ili 1). Model može raditi i višeklasnu klasifikaciju, stoga izlaz može biti i vjerojatnost neke od klasa CA, DA, SS ili No Event.

### 4.1.3. ResNet klasifikator

Primarna razlika između ovog i prethodnog modela je što se u prethodnom modelu koristi autoenkoder za dobivanje kompaktne reprezentacije ulaznih podataka te se umjesto korištenja dekodera za rekonstrukciju, izlaz kodera, tj. latentna reprezentacija unosi u klasifikator. ResNet klasifikator, s druge strane, izravno uči klasificirati ulazne podatke.

ResNet klasifikator sačinjavaju:

- Jednodimenzionalni konvolucijski sloj sa 64 filtra.
- 5 Res-blokova kao u prethodnom modelu.
- Jednodimenzionalni konvolucijski sloj sa 64 filtra.
- Potpuno povezani sloj aktiviran Swish funkcijom.

U početku se ulazni podaci obrađuju kroz konvolucijski sloj i prosljeđuju kroz Swish aktivacijsku funkciju. Nakon prolaska kroz Res-blobove, podatke dalje obrađuje drugi konvolucijski sloj, koji agregira naučene značajke smanjenjem broja kanala sa 64 na 1.

Izlaz se izravna, pretvarajući vektor iz dimenzija (veličina minigrupe, broj signala, vremenski prozor) u (veličina minigrupe, broj signala \* vremenski prozor). Spljošteni izlaz prolazi kroz potpuno povezani sloj koji ga transformira u vektor veličine  $2 \times 1$  za binarnu klasifikaciju.

## 4.2. U-NET 1D CNN

1D konvolucijske neuronske mreže (CNN) specijalizirani su oblik konvolucijskih neuronskih mreža. Dizajnirane su za obradu i analizu jednodimenzionalnih sekvencijalnih podataka, za razliku od tradicionalnih 2D CNN, koji rade na dvodimenzionalnim mrežama podataka (kao što su slike). Zbog toga su posebno prikladni za zadatke koji uključuju vremenske serije podataka, poput obrade signala.

U-NET model je jednodimenzionalna konvolucijska neuronska mreža sa tzv. U-NET arhitekturom, koja je inicijalno smišljena za segmentaciju slika u biomedicini. Srž U-NET modela njegova je arhitektura u obliku slova U, koja se sastoji od sažimanja (tj. kodera) i širenja (tj. dekodera).

Implementacija se sastoji od slijedećih blokova:

Konvolucijski blokovi:

- Jednodimenzionalni konvolucijski sloj.
- Normalizacija minigrupe.
- ReLU aktivacijska funkcija.

Ovaj blok primjenjuje operacije konvolucije na ulazne podatke, nakon čega slijedi normalizacija i ReLU aktivacijska funkcija za uvođenje nelinearnosti.

SE blok: Poboljšava osjetljivost mreže na relevantne značajke.

- Dva konvolucijska sloja.
- Sloj uprosječivanja (engl. average pooling).
- Dva potpuno povezana sloja.
- ReLU aktivacijska funkcija.
- Sigmoidna aktivacijska funkcija.

Sloj uprosječivanja provodi se slijedećom formulom:

$$\text{out}(N_i, C_j, l) = \frac{1}{k} \sum_{m=0}^{k-1} \text{input}(N_i, C_j, \text{stride} \times l + m) \quad (4.5)$$

gdje je:

input : Ulazni vektor

$N_i$  :  $i$ -ti element u seriji

$C_j$  :  $j$ -ti kanal

$l$  : Indeks trenutnog izlaznog elementa

$k$  : Veličina jezgre

stride : Veličina pomaka prozora preko vektora

$m$  : Indeks unutar trenutnog prozora vektora

out : Izlazni vektor nakon uprosječivanja

Sloj provodi uprosječivanje, a zatim obrađuje vektor s konvolucijskim slojevima i aktivacijskim funkcijama. Rezultantni vektor se zbraja s ulaznim vektorom, čime se stvara rezidualna ili preskočna veza između slojeva.

Rezidualni blok: Uvodi preskočne veze koje pomažu s problemom nestajanja gradijenta. Sastoji se od:

- Dva konvolucijska sloja.
- SE bloka.

Na kraju ovog bloka ponovno se rezultatni vektor zbraja s ulaznim.

Izlazni konvolucijski sloj - smanjuje broj kanala na željeni broj izlaznih kanala, čime radimo binarnu ili višeklasnu klasifikaciju.

## 5. Rezultati i diskusija

### 5.1. Primijenjene evaluacijske mjere

U ocjenjivanju performansi modela koristit će se sljedeće mjere: točnost, Cohenova Kappa mjera, odziv, preciznost i F1 mjera.

Točnost je omjer točno predviđenih primjera prema ukupnom broju primjera. To je jedna od najintuitivnijih metrika evaluacije, ali može biti zavaravajuća ako su klase podataka ne-balansirane.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.1)$$

- $TP$  (True Positives): Broj pozitivnih uzoraka koje je model ispravno identificirao.
- $TN$  (True Negatives): Broj negativnih uzoraka koje je model točno identificirao.
- $FP$  (False Positives): Broj negativnih uzoraka koje je model pogrešno identificirao kao pozitivne.
- $FN$  (False Negatives): Broj pozitivnih uzoraka koje je model pogrešno identificirao kao negativne.

Cohenova kappa mjera, simbolizirana malim grčkim slovom  $\kappa$  statistička je mjera slaganja među ocjenjivačima za kategoričke stavke. Općenito se smatra da je to robusnija mjera od jednostavnog izračuna postotka slaganja jer uzima u obzir slaganje koje se dogodilo slučajno.

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (5.2)$$

- $p_o$ : Uočeno slaganje, tj. točnost modela izračunata po formuli 5.1
- $p_e$ : Očekivani slučajni dogovor, izračunat kao:

$$p_e = \frac{(TP + FP)(TP + FN) + (TN + FN)(TN + FP)}{(TP + TN + FP + FN)^2} \quad (5.3)$$

Može se kretati od -1 do +1, gdje:

- -1 predstavlja slaganje gore od slučajnosti
- 0 predstavlja da nema slaganja osim onoga što bi se slučajno očekivalo
- 1 predstavlja potpuno savršeno slaganje.

Odziv je omjer uzoraka podataka koje model strojnog učenja ispravno identificira kao pripadnost klasi od interesa, odnosno pripadnost "pozitivnoj klasi" - od ukupnog broja uzoraka za tu klasu.

$$R = \frac{TP}{TP + FN} \quad (5.4)$$

Preciznost je metrika koja mjeri koliko često model strojnog učenja točno predviđa pozitivnu klasu. Preciznost možete izračunati dijeljenjem broja točnih pozitivnih predviđanja (pravih pozitivnih) s ukupnim brojem instanci koje je model predvidio kao pozitivne (pravih i lažnih pozitivnih).

$$P = \frac{TP}{TP + FP} \quad (5.5)$$

Predstavlja harmonijsku sredinu preciznosti i odziva. Najviša moguća vrijednost F1 mjere je 1,0. što ukazuje na savršenu preciznost i odziv, a najniža moguća vrijednost je 0. ako su preciznost i odziv jednaki nuli.

$$F1 = 2 \cdot \frac{P \cdot R}{P + R} \quad (5.6)$$

## 5.2. ResNet autoenkoder i klasifikator

ResNet autoenkoder i klasifikator ne pokazuje obećavajuće rezultate.

### 5.2.1. Autoenkoder

Unatoč svojoj sofisticiranoj arhitekturi, autoenkoder pokazuje minimalne promjene u gubicima treniranja i validacije tijekom vremena. Ova stagnacija ukazuje na to da je model vrlo rano prestao učiti.

Autoenkoder i klasifikator su zapravo dva modela koji rade sekvencijalno. Izlaz dekodera se ignorira, a latentna reprezentacija koju dobijemo iz autoenkodera prosljeđuje se u klasifikator. Kod autoenkodera i klasifikatora gubitak se optimira metodom SMORMS3.

Algoritam optimizacije SMORMS3 ažurira parametre pomoću sljedećih koraka:

1. Izračuna se gradijent funkcije gubitka s obzirom na parametre:

$$g_t = \nabla_{\theta} \mathcal{L}(\theta_t) \quad (5.7)$$

2. Ažurira se pomični prosjek gradijenata i gradijenata na kvadrat:

$$m_t = (1 - \beta)m_{t-1} + \beta g_t \quad (5.8)$$

$$v_t = (1 - \beta)v_{t-1} + \beta g_t^2 \quad (5.9)$$

3. Izračuna se recipročna vrijednost pomičnog prosjeka kvadratnih gradijenata:

$$r_t = \frac{1}{v_t} \quad (5.10)$$

4. Ažuriranje parametara:

$$\theta_{t+1} = \theta_t - \eta \frac{m_t}{\sqrt{v_t + \epsilon}} \quad (5.11)$$

gdje je:

$g_t$  : Gradijent funkcije gubitka u vremenskom koraku  $t$

$m_t$  : Pokretni prosjek gradijenata

$v_t$  : Pokretni prosjek kvadratnih gradijenata

$r_t$  : Recipročna vrijednost pomičnog prosjeka kvadratnih gradijenata

$\theta_t$  : Parametri u vremenskom koraku  $t$

$\theta_{t+1}$  : Ažurirani parametri u vremenskom koraku  $t + 1$

$\beta$  : Parametar izgladivanja za pokretne prosjeke

$\eta$  : Stopa učenja

$\epsilon$  : Mala konstanta za izbjegavanje dijeljenja s nulom



Gubitak treniranja počinje na približno 0.9756 i samo se malo smanjuje na oko 0.9586 do kraja procesa treniranja. Slično tome, gubitak validacije počinje na 0.1733 i završava na oko 0.1649. Minimalne promjene u ovim vrijednostima gubitaka pokazuju da je sposobnost modela da uči iz podataka o treniranju i generalizira na podatke o validaciji dosegla plato. Ova stagnacija sugerira da trenutni hiperparametri i arhitektura modela nisu u potpunosti učinkoviti u izdvajanju i učenju temeljnih obrazaca u podacima. Sukladno tome, ovaj model ne pokazuje bolje ponašanje od slučajnog pogađanja prilikom testiranja.

### 5.2.2. Klasifikator

Gubitak se kao i u prethodnom modelu optimira metodom SMORMS3. Klasifikator postiže točnost treniranja od približno 52.74% i točnost validacije od oko 56.01%. Ove vrijednosti upućuju na to da, iako model radi malo bolje na skupu za validaciju, njegova je ukupna točnost relativno niska. Gubitak treniranja je 0.6908, a gubitak validacije je 0.6859, što ukazuje da model nije pretreniran jer su gubici prilično slični za oba skupa podataka.

F1 mjera treniranja je 0.6239, dok je za skup validacije 0.6585. Ovi su rezultati sugeriraju da model nema optimalnu sposobnost ispravnog klasificiranja pozitivnih i negativnih instanci. Vrijednosti preciznosti (0.5910) i odziva (0.5606) sugeriraju da je klasifikator bolji u ispravnom identificiranju pozitivnih instanci (manje FP vrijednosti), ali zato dobivamo više FN vrijednosti.

Cohenova kappa vrijednost od 0.1210 ukazuje na blago slaganje izvan slučajnosti između predviđanja modela i stvarnih oznaka. Ova niska vrijednost naglašava da je potrebno dalje istraživati te mijenjati arhitekturu i/ili hiperparametre modela.

Matrica zabune daje još jednu perspektivu u performanse modela:

$$\begin{bmatrix} 13.59\% & 36.49\% \\ 7.50\% & 42.41\% \end{bmatrix} \quad (5.12)$$

Veliki broj lažno negativnih rezultata sugerira da model ima tendenciju netočno klasificirati pozitivne slučajeve kao negativne. Suprotno tome, postoji puno manje lažno pozitivnih rezultata, ali su još uvijek značajni.

Za zadatak klasifikacije s više klasa, metrike izvedbe modela znatno su niže. Točnost treniranja je 27.96%, a točnost validacije je 30.57%, što je znatno niže od onih u scenariju binarne klasifikacije. Gubitak treniranja je 1.3953, a gubitak validacije je 1.3865, oba veća od gubitaka uočenih u binarnoj klasifikaciji. Rezultati F1 su 0.2796 za treniranje i 0.3057 za validaciju, što ukazuje da se model više bori s klasifikacijom u više klasa. Preciznost je 0.3221, odziv je 0.3051, a Cohenova kappa mjera je 0.0897.

$$\begin{bmatrix} 7.02\% & 22.16\% & 0.04\% & 6.15\% \\ 2.92\% & 27.72\% & 0.00\% & 4.71\% \\ 4.56\% & 20.88\% & 0.04\% & 9.86\% \\ 4.96\% & 20.98\% & 0.03\% & 8.05\% \end{bmatrix} \quad (5.13)$$

Velik broj nedijagonalnih elemenata u matrici pokazuje značajne pogreške u klasifikaciji.

### 5.3. ResNet klasifikator

ResNet klasifikator pokazuje bolje rezultate od prethodno navedenog ResNet autoenkodera i klasifikatora za binarnu klasifikaciju. Za višeklasnu klasifikaciju model ne pokazuje bolje ponašanje od slučajnog pogađanja prilikom testiranja.

Ovaj model, treniran tijekom samo jedne epohe, pokazao je obećavajuće rezultate. Rezultat F1 na treniranju dosegao je 0.6119, dok je rezultat F1 na validaciji postigao vrijednost od 0.6793. Preciznost i odziv bili su relativno visoki, s rezultatima od 0.7045 i 0.7008. Ove metrike ukazuju da model može relativno uspješno razlikovati klase.

Cohenova kappa mjera daje rezultat 0.4017, što sugerira mnogo bolje od slučajnog slaganja između predviđenih i stvarnih oznaka klase.

$$\begin{bmatrix} 37.50\% & 11.34\% \\ 17.85\% & 33.31\% \end{bmatrix} \quad (5.14)$$

Sa 37.50% pravih negativnih i 33.31% pravih pozitivnih rezultata u matrici konfuzije, klasifikator je ispravno identificirao većinu podataka. Međutim, prisutnost 11.34% lažno pozitivnih i 17.85% lažno negativnih ukazuje da postoji prostor za poboljšanje.

Točnost od 64.96% na skupu za treniranje i 68.91% na skupu za validaciju pokazuju slično. Vrijednosti gubitaka treniranja i validacije s treniranjem su se konstantno smanjivale zahvaljujući planeru stopi učenja "ReduceLROnPlateau" koja iterativno smanjuje stopu učenja ukoliko se gubitak na validaciji ne smanjuje. Konačne vrijednosti gubitaka iznose 0.6462 za skup treniranja te 0.6291 za skup validacije.

## 5.4. U-NET 1D CNN

Funkcija gubitka je unakrsna entropija (BCEWithLogitsLoss) koja kombinira binarnu unakrsnu entropiju sa sigmoidom.

Planer stope učenja (ReduceLRonPlateau) koristi se za prilagodbu stope učenja na temelju gubitka testiranja, čime se sprječava pretreniranje i osigurava se bolja konvergencija mreže.

Korištenje LFSO unakrsne validacije na U-NET modelu pokazuje da model ne može naučiti razlikovati klase s točnošću većom od slučajne niti za binarnu niti za višeklasnu klasifikaciju.

U okviru K-Fold unakrsne validacije, model U-NET ostvaruje rezultate raspoznavanja klasa koji su bolji od nasumičnog pogađanja.

Konačne vrijednosti gubitaka za skupove podataka za treniranje i validaciju bile su 0.3202 odnosno 0.4029. Ove vrijednosti odražavaju dobro generalizirani model koji učinkovito ublažava pretreniranje.

Rezultati F1 za skupove podataka za treniranje i validaciju bili su 0.6761 odnosno 0.6310. Rezultat F1 ukazuje na to da model dosljedno dobro radi i na viđenim i na neviđenim podacima. Dodatno, ocjena točnosti od 0.6344 naglašava umjerenu sposobnost klasifikacije. Rezultati preciznosti i odziva, koji su 0.6295 odnosno 0.6353, dodatno to potvrđuju. Cohenov kappa rezultat od 0.6224 ilustrira značajno slaganje i time pojačava pouzdanost našeg modela.

No, model radi bolje na binarnoj klasifikaciji. Gubitak treniranja bio je 0.3847 i gubitak validacije 0.3231. F1 rezultat treniranja bio je 0.8052, dok je F1 rezultat validacije bio relativno viši, 0.8408. Ove metrike pokazuju da model ima značajno bolju izvedbu u razlikovanju dviju klasa, sa snažnom generalizacijom za validacijski skup.

$$\begin{bmatrix} 27.24\% & 4.89\% & 2.11\% & 0.60\% \\ 1.48\% & 25.99\% & 1.93\% & 4.63\% \\ 2.07\% & 4.08\% & 20.53\% & 10.70\% \\ 4.00\% & 4.50\% & 10.25\% & 15.00\% \end{bmatrix} \quad (5.15)$$

Matrica konfuzije otkriva da model održava dobar stupanj točnosti u većini klasa, s prostorom za poboljšanje. Iako su prave pozitivne stope zadovoljavajuće, model povremeno pogrešno klasificira određene slučajeve, kao što se vidi u nedijagonalnim elementima matrice.

Korišteni planer stope učenja ReduceLRonPlateau bio je učinkovit, ali eksperimentiranje s drugim metodama raspoređivanja kao što su Cosine Annealing ili OneCycleLR moglo bi dodatno poboljšati konvergenciju i performanse modela. Osim toga, eksperimentiranje s različitim veličinama i dubinama slojeva moglo bi otkriti arhitekturne konfiguracije koje

poboljšavaju točnost i generalizaciju.

Korištenje K-Fold unakrsne validacije omogućilo je sveobuhvatnu procjenu performansi modela u različitim podskupovima podataka. Svako preklapanje konzistentno pokazuje poboljšanje svih metrika evaluacije čime se potvrđuje sposobnost modela da relativno dobro generalizira na neviđenim podacima.

Točnost za zadatak binarne klasifikacije bila je 0.8415, što pokazuje da model ispravno klasificira veliku većinu podataka. Preciznost i odziv iznosili su 0.8485 odnosno 0.8417, što ukazuje na uravnoteženu izvedbu u smislu identificiranja pravih pozitivnih rezultata i smanjenja lažno pozitivnih i negativnih rezultata. Cohenova kappa mjera iznosila je 0.6832, što ukazuje na značajno bolje slaganje od puke slučajnosti.

$$\begin{bmatrix} 46.62\% & 4.47\% \\ 11.77\% & 37.14\% \end{bmatrix} \quad (5.16)$$

Matrica zabune pokazuje performanse modela s 46.62% pravih pozitivnih i 37.14% pravih negativnih rezultata, što ukazuje da je većina instanci ispravno klasificirana. Međutim, još uvijek postoji 4.47% lažno pozitivnih i 11.77% lažno negativnih.

## 6. Zaključak

U ovom radu ispitana su tri duboka modela strojnog učenja različitih arhitektura na problemu raspoznavanja visokorizičnih mentalnih stanja pilota: ResNet autoenkoder i klasifikator, ResNet klasifikator te U-NET 1D. Modeli su implementirani u programskom jeziku Python te su testirani na skupu podataka "Flight Crew Physiological Data for Crew State Monitoring" od NASA-e.

ResNet autoenkoder i klasifikator pokazao je najlošije rezultate. Model nije uspio značajno smanjiti gubitak tijekom treniranja, što ukazuje na rani prestanak učenja. Metrike na validaciji i treniranju bile su dosta niske jer je model radio dosta grešaka s klasificiranjem. Rezultati za višeklasnu klasifikaciju bili su još niži, što ukazuje na potrebu za daljnjim istraživanjem ove arhitekture.

ResNet klasifikator pokazao je bolje rezultate od prethodnog modela, ali je i dalje imao prostor za poboljšanje. Točnost treniranja i validacije bila je relativno niska, s relativno visokim brojem lažno pozitivnih i lažno negativnih rezultata. Rezultati Cohenove kappa mjere sugeriraju da model ima bolje slaganje od slučajnog, ali, kao i kod prethodnog modela, potrebno je dalje istraživati.

U-NET 1D pokazao je najbolje rezultate. Ispitana je i binarna i višeklasna klasifikacija, a model je pokazao značajno bolje rezultate u binarnoj klasifikaciji. Točnost treniranja i validacije bila je viša nego kod prethodna dva modela te je model pokazao dobru sposobnost generalizacije na neviđenim podacima. Korištenje K-Fold unakrsne validacije dalo je drukčiji uvid u procjenu performansi modela, dok je LFSO validacija dala lošije rezultate. Planer stope učenja ReduceLRonPlateau pokazao se učinkovitim, no valja eksperimentirati i s drugim planerima koji bi mogli dodatno poboljšati konvergenciju modela.

Analiza izvedbe modela pokazala je da U-NET 1D ima najveći potencijal za rješavanje problema klasifikacije na ovom skupu podataka. Međutim, potrebno je daljnje istraživanje i optimizacija arhitektura i hiperparametara kako bi se poboljšala moć klasifikacije svih modela.

Također, prikupljanje novog skupa podataka gdje postoji više instanci klasa CA, DA i SS moglo bi znatno poboljšati rezultate.

Istraživanje novih arhitektura dubokih modela moglo bi bolje izvući značajke iz poda-

taka. S obzirom da hiperparametri značajno utječu na performanse modela, sustavnim podešavanjem i optimizacijom moguće je postići bolje rezultate. Istraživanje drugih metoda validacije moglo bi dati bolji uvid u prednosti, no i nedostatke modela.

## **Raspoznavanje visokorizičnih mentalnih stanja pilota na temelju obrade fizioloških signala metodama dubokog učenja**

### **Sažetak**

U ovom radu primijenjena su tri modela dubokog učenja za klasificiranje visokorizičnih mentalnih stanja pilota. Ispitane arhitekture su ResNet autoenkoder i klasifikator, ResNet klasifikator i U-NET 1D. Za procjenu su korištene metrike evaluacije točnosti, preciznosti, odziva, F1 i Cohen Kappa. ResNet Autoenkoder i klasifikator te U-NET 1D pokazuju vrlo ograničenu sposobnost klasificiranja korištenjem LFSO (engl. Leave Five Subject Out) unakrsne validacije, dok ResNet klasifikator daje bolje rezultate. ResNet klasifikator LFSO validacijom postigao je najbolju točnost od 68.91% na skupu za validaciju, ali i dalje postoji velik broj lažno pozitivnih klasificiranja. U-NET 1D model pokazao je loše rezultate korištenjem LFSO validacije, ali znatno bolje rezultate korištenjem K-Fold validacije, posebno u binarnoj klasifikaciji, gdje postiže točnost od 84.15%.

**Ključne riječi:** Duboko učenje, visokorizična mentalna stanja, obrada fizioloških signala, komparativna analiza

### **Recognition of pilots' high-risk mental states based on physiological signal processing by deep learning methods**

#### **Abstract**

In this paper, three deep learning models were applied to classify high-risk mental states of pilots. The tested architectures are ResNet autoencoder and classifier, ResNet classifier and U-NET 1D. The evaluation metrics of accuracy, precision, responsiveness, F1 and Cohen Kappa were used for evaluation. ResNet Autoencoder and classifier and U-NET 1D show very limited accuracy using LFSO (Leave Five Subject Out) cross-validation, while ResNet classifier gives better results. The ResNet classifier with LFSO validation achieved the best accuracy of 68.91% on the validation set, but there are still a large number of false positive classifications. The U-NET 1D model showed poor results using LFSO validation, but significantly better results using K-Fold validation, especially in binary classification, where it achieves an accuracy of 84.15%.

**Keywords:** Deep learning, high-risk mental states, physiological signal processing, comparative analysis

# LITERATURA

- [1] Lenka Hanakova, Vladimir Socha, Lubos Socha, Stanislav Szabo, Jaroslav Kozuba, Andrej Lalis, Peter Vittek, Jakub Kraus, Robert Rozenberg, Peter Kalavsky, et al. Determining importance of physiological parameters and methods of their evaluation for classification of pilots psychophysiological condition. In *2017 International Conference on Military Technologies (ICMT)*, pages 500–506. IEEE, 2017.
- [2] Oliver Faust, Yuki Hagiwara, Tan Jen Hong, Oh Shu Lih, and U Rajendra Acharya. Deep learning for healthcare applications based on physiological signals: A review. *Computer methods and programs in biomedicine*, 161:1–13, 2018.
- [3] Paul Terwilliger, Jack Sarle, Shannon Walker, and Angela Harrivel. A resnet autoencoder approach for time series classification of cognitive state. *Proceedings of the MODSIM World, Norfolk, VA, USA*, pages 1–11, 2020.
- [4] Soo-Yeon Han, No-Sang Kwak, Taegeun Oh, and Seong-Whan Lee. Classification of pilots’ mental states using a multimodal deep learning network. *Biocybernetics and Biomedical Engineering*, 40(1):324–336, 2020.
- [5] Angela R Harrivel, Chad L Stephens, Robert J Milletich, Christina M Heinich, Mary Carolyn Last, Nicholas J Napoli, Nijo Abraham, Lance J Prinzel, Mark A Motter, and Alan T Pope. Prediction of cognitive states during flight simulation using multimodal psychophysiological sensing. In *AIAA Information Systems-AIAA Infotech@ Aerospace*, page 1135. 2017.
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [7] Luca Varotto, Marco Fabris, Giulia Michieletto, and Angelo Cenedese. Visual sensor network stimulation model identification via gaussian mixture model and deep embedded features. *Engineering Applications of Artificial Intelligence*, 114:105096, 2022.