

Проектна задача по предметот „Вовед во науката за податоци“

на тема

„ Препознавање на гестови базирани на електромиографски сигнали со примена на машинско учење“

Изработиле :

Ана Ѓурчинова 222004

Виолета Капчевска 223047

Христина Здравеска 222001

Ментор :

Проф. Др Соња Гиевска

Проф. Др Слободан Калајџиски

Финки, септември 2025

Содржина

Апстракт	3
Вовед	4
Методологија	5
Визуелизации	9
Дизајн и архитектура на модел	12
Тренирање на моделот и евалуација	13
Споредба на модели	14
Заклучок	16

Апстракт

Во овој научен труд нашиот фокус воглавно го ставаме на презентација на напредни системи за препознавање на различни гестови на рака базирани на електромиографски (ЕМГ) сигнали, користејќи техники на длабоко учење. За оваа цел искористен е јавно достапниот податочен сет GrabMyo од PhysioNet кој содржи во себе електромиографски сигнали снимени при изведба на разновидни моторни движења.

Предложениот пристап е комбинација на софистицирано претпроцесирање со иновативна LightWeightCNN архитектура. Претпроцесирањето вклучува селекција на канали, 50 Hz notch филтер, bandpass филтрирање, како и нормализација на податоците. За секој временски прозорец од 200 мс се извлекуваат соодветно седум карактеристики (MAV, RMS, VAR, WL, ZC, SSC, и WAMP). Со примената на лесната конволутивна невронска мрежа извршени се комбинации на информациите од суровиот сигнал и извлечените карактеристики. Во врска со спречувањето на дисбалансот на податочниот сет искористени се и доаѓаат до израз Focal Loss и Weighted Random Sampler.

Резултатите кои се добиени во оваа задача се еден вид демонстрација на ефективност на спој на различни традиционални карактеристики со длабоко учење за препознавање на гестови кои се ЕМГ-базирани, односно постигнат е ефикасен пристап кон препознавање на рачни гестови во реално време, носејќи погодност за практични примени.

Вовед

Способноста да се разберат и објаснат човечките намери преку физиолошки сигнали како во овој случај несомнено претставува клучен чекор во развојот на интелигентни системи. Во сферата на интеракција меѓу човек и компјутер, препознавањето на гестови на рака базирано на ЕМГ сигнали се истакнува како мошне ветувачко поле. ЕМГ сигналите кои ја мерат електричната активност од скелетните мускули, нудат можност за директен увид во моторните цели на поединецот, овозможувајќи прецизно и моментално препознавање на движењата.

Традиционалните методи за препознавање на гестови честопати се потпираат на визуелни податоци или инерцијални сензори, па ова е проследено со појава на препреки како чувствителност при услови на осветлување или потреба од постојан визуелен контакт. Од друга страна ЕМГ-базираниот пристап ги надминува овие препреки, обезбедувајќи робусност. Но, комплексноста на овој вид сигнали и нивната варијабилност бараат софистицирани техники за нивна обработка како и ефикасни модели за машинско учење, кои дел од нив ги применивме во нашата задача.

Нашата цел е да се развие и демонстрира систем за идентификација на осум различни гестови на рака со користење на податочниот сет GrabMyo. Нашиот предлог се состои од хибриден пристап кој е спој на традиционалната обработка на сигналите со моќта на длабокото учење. Пристапот започнува со детално неопходно претпроцесирање на суровите ЕМГ сигнали, подоцна проследено со сегментација на истите во временски прозорци. На самиот крај следува наша примена на иновативна архитектура на лесна конволутивна невронска мрежа. Нашата цел е да направиме обид и да покажеме дека со примена на техники за справување со дисбаланс на податоци и подобрување на добиените предвидувања може да се постигне висока точност и сигурност во препознавање на гестовите.

Методологија

За потребите на оваа задача беше користен податочен сет GrabMyo – сет на податоци креиран со цел за анализа и класификација на електромиографски сигнали за препознавање на движења на раце и гестови. Податочниот сет во себе содржи сигнали добиени преку MyoArmband , која е еден вид сензорска направа што се поставува во пределот на подлактицата на секој учесник. Секој учесник изведува осум различни гестови во текот на три сесии , со по три обиди за секој гест во секоја сесија.

Податоците во себе опфаќаат разни видови на гестови , односно различни позиции и движења на раце. Секој гест е снимен во повеќе повторувања , што воедно овозможува да се има репрезентативен примерок за тренирање на модели за машинско учење. Фреквенцијата на примероците изнесува 2048Hz што значи нуди можност за детекција на брзи промени во мускулната активност.

Во рамките на датасетот , секој примерок може да се разгледува како временска серија со повеќе канали, каде секој канал соодветсвува на еден од сензорите на MyoArmband. Со оглед на тоа што ЕМГ сигналите по својата природа содржат шум и артефакти кои се последица на движење , контакт на електроди , тие претставуваат еден голем предизвик за анализа и класификација.

Овој датасет е солидна основа за тестирање на разновидни пристапи и модели за класификација. Во рамките на ова истражување искористени се два различни модела со цел за анализа на истиот сет од податоци, првиот како иницијален експеримент , а вториот како финално решение, секој по себе давајќи различни финални резултати. Искористен е повеќе од еден модел со цел да видиме дали можеме да стигнеме до подобри перформанси при предвидувањето, односно дали можат да се добијат подобри резултати како точност(accuracy), прецизност(precision), повик (recall) и f1-мерка. Со следењето на двата модела во однос на начинот на нивно справување со податоците, врз кои подоцна ќе биде извршена подетална споредба , се роди можноста да се увиди кој модел дава поконзистентни и попрецизни резултати. Од друга страна анализата на повеќе од еден модел овозможи да се забележат сите слаби точки и обратно – потенцијални точки кои водат кон подобрување , како и различни параметри и нивно влијание врз нив. Не помалку важно и бенефит од користењето на разновидни модели е тоа што носи проверка во сигурноста и стабилноста на предвидувањата како и

проценка за нивна применливост во реални сценарија за препознавање на разни гестови. Податоците од сетот на сам почеток – ЕМГ сигнали се прикажани како физиолошки записи кои се читаат со помош на WFDB софтверот за Python, специјализиран за читање и анализа на овој вид сигнали. По самото вчитување сигналите поминуваат низ низа на филтрирања и трансформации, при што се обезбедува отстранување на евидентното шум и подготовка на карактеристики за класификација.

Претпроцесирањето е клучен чекор во секое истражување од областа на машинското учење, постоејќи опасност податоците да содржат шум, во наш случај сигналите или пак пропуштени вредности што сите овие ризици водат кон нарушување на добри перформанси на даден модел. Па така најпрвен во однос на каналите кои во целина се 32 на број извршена е нивна селективна употреба, базирана врз основа на квалитетот на сигналот и негова анатомска локација во првиот обид на градење на модел, додека во финалниот искористени се сите канали без никаква селекција. Имплементирани се три режими на селекција вклучувајќи:

- Подлактица режим – селектира 16 канали кои се позиционирани на подлактицата.
- Зглоб режим – селектира 12 канали позиционирани во предел на зглоб.
- Режим на сите останати канали – искористени се 28 канали, елиминирајќи ги каналите со лош квалитет на сигнал (16,23,24,31)

Следен чекор кој е извршен е филтрирањето на шум составен од два дела, прв дел во кој се отстранува интерференција на електрична мрежа на 50 Hz во кој е искористен IIR-notch филтер кој е конфигуриран со централна фреквенција, Q-фактор кој има вредност 30 и широчина на појасот. Употребата на `filtfilt()` функцијата обезбедува zero-phase филтрирање, што значи дека не се воведува временско одложување во сигналот, што е критично за одржување на временската синхронизација меѓу каналите. Во вториот дел применето е bandpass-филтрирање (20-450Hz), долната граница која има вредност 20 Hz служи за отстранување на движечки артефакти, додека горната граница во вредност од 450 Hz ги елиминира високофреквенциските пречки.

```
[69]: def remove_50hz_noise(signal, fs):
      b, a = iirnotch(50, 30, fs)
      return filtfilt(b, a, signal, axis=0)

[70]: #to filter only the signals in frequency[20,450]
      def bandpass_filter(data, low_cut=20, high_cut=450, fs=2000, order=4):
          nyq = fs / 2
          if high_cut >= nyq:
              high_cut = nyq-1
          if low_cut >= high_cut:
              low_cut = high_cut / 2
          low = low_cut / nyq
          high = high_cut / nyq
          b, a = butter(order, [low, high], btype='band')
          return filtfilt(b, a, data, axis=0)
```

Слика бр.1 – процес на филтрирање на податоците

По филтрирање на сигналите, неопходна е нивна нормализација. Секој сигнал се нормализира со цел да се израмнат амплитудите на различни канали. Ова практично значи дека за секој канал ја одземаме нивната средна вредност(mean) и го делиме со стандардна девијација, согласно дефиницијата на StandardScaler. Значајноста на нормализацијата е тоа што носи сигурност дека сите карактеристики ќе имаат сличен распон и нема една карактеристика да доминира друга.

Потоа ЕМГ-сигналите се сегментираат со техника на лизгачки прозорци(sliding windows) со следните параметри : должина на прозорец = 0.02 секунди, чекор(hop) = 0.05 секунди и максимум прозорци по обид = 60. За секој временски прозорец традиционални временски карактеристики како што се : средна апсолутна вредност (MAV), корен квадратен на средна вредност (RMS), должина на бран (Waveform Length) и броење на пресечувања низ нула (Zero Crossings) по канал. Овие карактеристики даваат опис за обликот и варијабилноста на сигналот и се користат заедно со суровиот сигнал во процесот на класификација. Па така во наш случај со 28 канали и 4 карактеристики по канал, секој прозорец се претставува како вектор од 112 карактеристики.

Генерирањето на датасетот е означено со функцијата make_dataset() која е централна компонента во подготовка на податоците за машинско учење. Целта на функцијата е да изврши систематска трансформација на суровите ЕМГ-сигнали во структури кои соодветствуваат за тренирање на модел за класификација. Притоа, процесот опфаќа повеќе чекори: вчитување на сигналите, филтрирање според зададени критериуми, претроцесирање и сегментација, екстракција на карактеристики и групирање на податоците во стандардизирани матрици. Оваа функција прифаќа параметри како што се :

- root_folder – патека до податочниот сет
- sessions – кои сесии да бидат вклучени

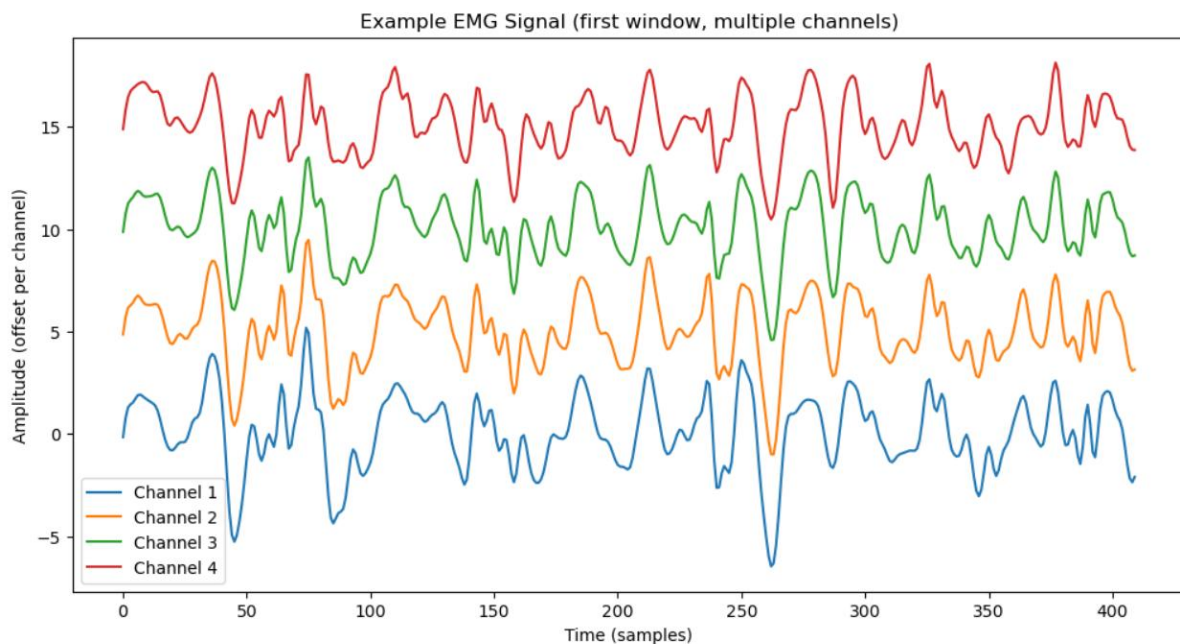
- `gestures_to_keep` – овозможува селекција на специфични гестови
- `max_participants` `max_trials_per_gesture` – максимален број на учесници и максимален број на обиди по гест
- `window_seconds` и `hop_seconds` – дефиниција на времетраењето на прозорците и нивно преклопување
- `max_windows_per_trial` – прави ограничување на бројот на прозорци извлечени од еден обид.

Процес на генерирање на датасет:

- иницијализација на структури: се креираат празни листи во кои ќе се зачувуваат суровите прозорци, карактеристиките и ознаките.
- Итерација низ сите фајлови: со помош на специјализирани функции се добиваат сите записи од коренската папка.
- филтрирање на примероци според критериуми: Секој запис се парсира за да се издвојат информации за сесијата, учесникот, гестот и обидот. Доколку не е исполнет некој од критериумите (на пример, гестот не е во листата на дозволени, бројот на учесници е надминува или бројот на обиди за даден гест е веќе достигнат), записот се прескокнува.
- читање и претпроцесирање на сигналите: Секој вчитан сигнал се подложува на филтрирање и нормализација во согласност со дефинираниот `channel_mode`. Овој чекор обезбедува чистење на шумот и стандардизација на сигналите.
- сегментација во временски прозорци: Со помош на методот на прозорци (`windowing`) сигналот се дели на кратки преклопувачки сегменти. На овој начин се овозможува динамичко претставување на сигналот во реално време.
- извлекување на карактеристики и подготовка на податоци - За секој временски прозорец се издвојуваат два вида податоци:
 - суров сигнал (матрица со временски серии по канали)
 - вектор со статистички и сигнални карактеристики добиени преку функцијата `extract_features`.

Дополнително, секој прозорец добива и класа-ознака што е синоним на категоријата на гестот.

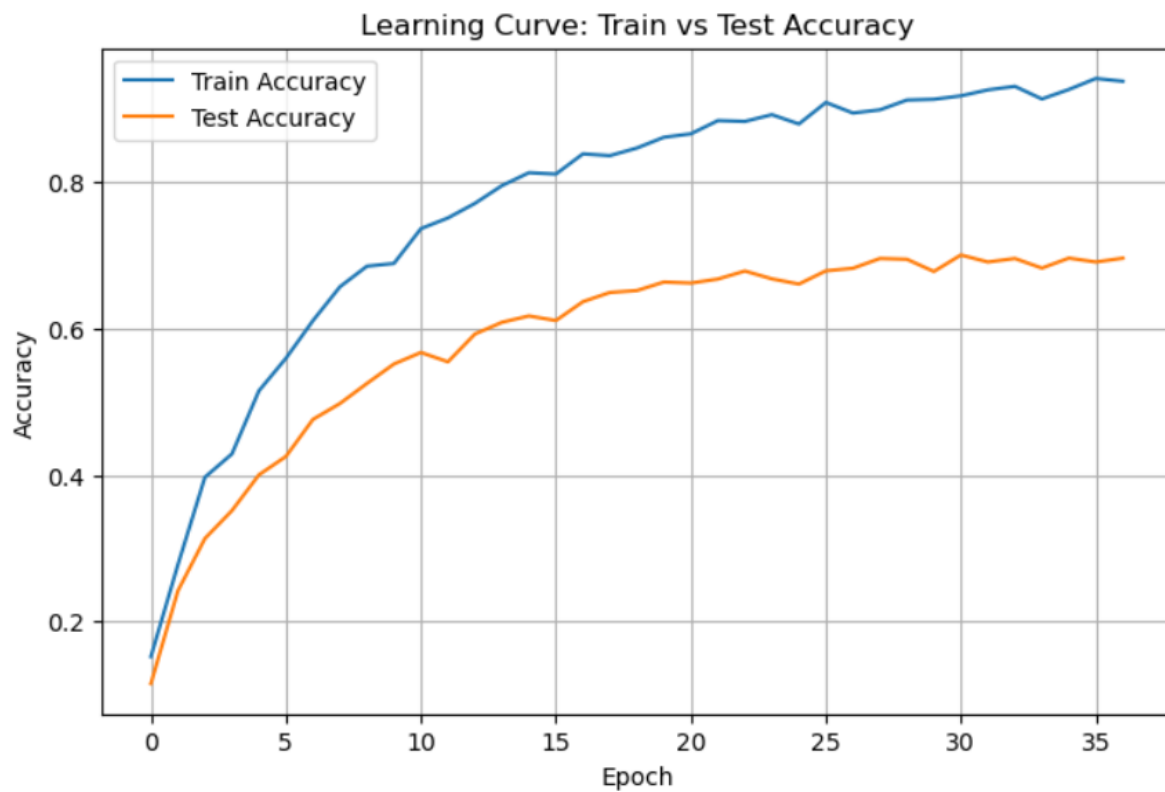
- агрегација - по обработката на сите записи, сите прозорци, нивните карактеристики и соодветните ознаки се конкатенираат во финални матрици.



Слика бр. 2 - графикон на сурови ЕМГ сигнали

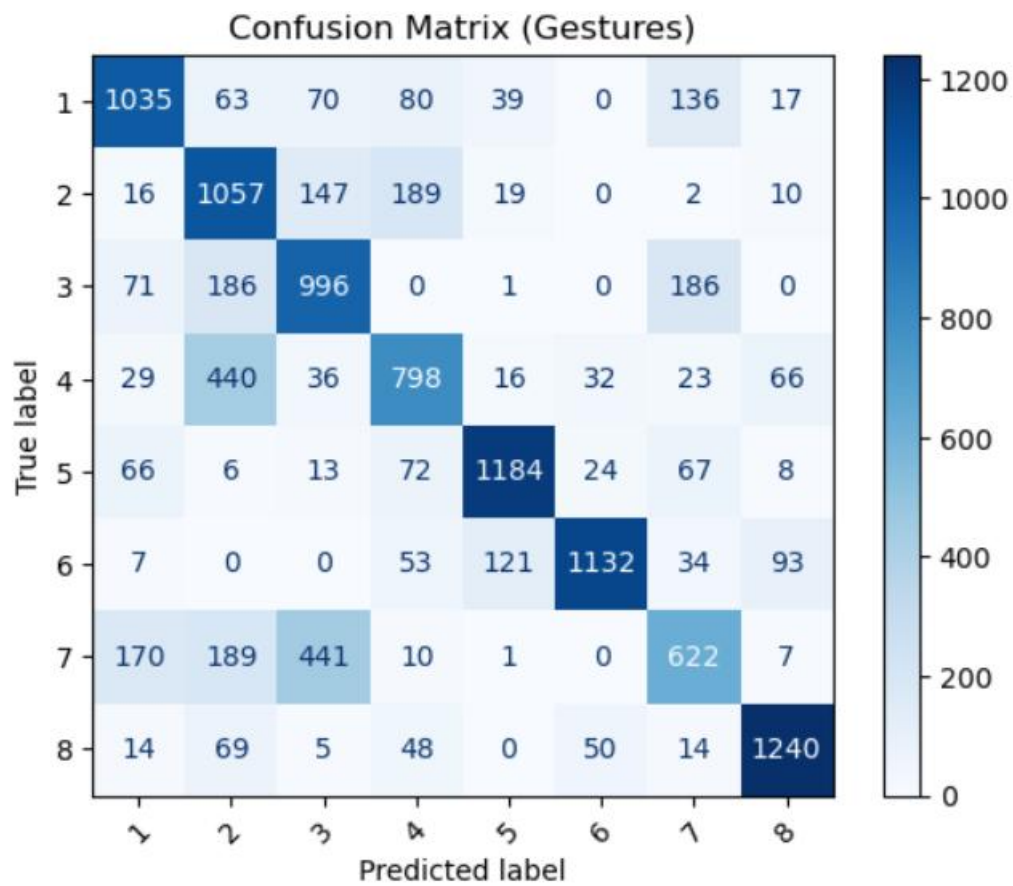
На графиконот имаме приказ од сурови ЕМГ сигнали од повеќе канали низ временски интервали. Можеме да забележиме периоди на поголема мускулна активност и делови без активност, како и разлики во амплитудите меѓу каналите. Оваа визуелизација е значајна од причина да се покаже квалитетот на податоците и потребата од филтрирање, нормализација како и

понатамошна обработка пред примена на моделите.



Слика бр. 3 – крива на точност при тренирање и тестирање низ епохи

Графиконот ја покажува точноста при тренирање и тестирање низ различни епохи. Иако е предвидено да се искористат поголем број на епохи(во наш случај имаме 50 епохи вкупно) , сепак примената на early stopping овозможува автоматски да се избере најдобар модел пред да настане overfitting , без разлика на максималниот број на епохи.



Слика бр.4 – Confusion Matrix

Confusion Matrix ни покажува точните и погрешни предвидувања за секоја класа. Вкупната точност е околу 70%, па така некои класи како 5,6 и 8 се препознаваат со голема точност.

Дизајн и архитектура на модел

Архитектурата на моделот што го развивме се темели на класичен пристап за обработка и класификација на биомедицински сигнали, при што основната цел е да се постигне прецизна категоризација на EMG сигналите добиени од предвидениот датасет. Како финален модел изградивме длабока невронска мрежа – LightweightCNN со 1D конволутивни слоеви.

```
class LightweightCNN(nn.Module):
    def __init__(self, in_ch, num_classes, use_features=True, feat_dim=0):
        super().__init__()
        self.use_features = use_features

        self.conv = nn.Sequential(
            DepthwiseSeparableConv1D(in_ch, 64, 3, dilation=1),
            DepthwiseSeparableConv1D(64, 128, 3, dilation=2),
            DepthwiseSeparableConv1D(128, 256, 3, dilation=4),
        )
        self.gap = nn.AdaptiveAvgPool1d(1)

        in_dim = 256
        if use_features:
            in_dim += feat_dim

        self.fc = nn.Sequential(
            nn.Linear(in_dim, 256),
            nn.ReLU(),
            nn.Dropout(0.4),
            nn.Linear(256, 128),
            nn.ReLU(),
            nn.Dropout(0.3),
            nn.Linear(128, num_classes)
        )
```

Слика бр. 5 – LightweightCNN модел

За да се постигне ефикасна обработка искористен е depthwise separable конволуција која прави поделба на стандардната конволуција на два дела :

- Depthwise конволуција - врши конволуција на секој влезен канал независно.
- Pointwise конволуција – комбинира резултати од различни канали.

Придонесот на овој вид на техника е тоа што значително го намалува бројот на параметри на моделот во однос на конвенционалните конволутивни слоеви со тоа што ја задржува репрезентативната моќ. Ова значи на пример стандарден $3 \times 3 \times 64$ конволутивен слој може да има 1792 параметри, додека еквивалентниот одделен слој има само 286. Во имплементацијата, мрежата содржи три последователни слоја DepthwiseSeparableConv1D со 64, 128 и 256 филтри (различно зададени дијаметри на филтрите/dilation), по што следи глобален Adaptive AvgPool1D слој. Овој pool слој врши апсорпција на временските димензии , а потоа трансформација на секој канал во една агрегирана вредност. Излезот што доаѓа од овој слој (доведен до димензии (batch, 256)) може да се конкатенира со дополнително изведените статистички карактеристики од прозорецот доколку се користи опцијата use_features. По поврзувањето, резултатниот вектор се подава во низа полносврзни (FC) слоеви: првиот FC слој ја намалува димензијата на 256, потоа следат два слоја

со ReLU активација и Dropout (пр. 0.4 и 0.3), а на крај се наоѓа излезен слој со онолку број на неврони колку што има класи на гестови.

Тренирање на моделот и евалуација

Тренирањето на моделот се врши преку поделба на дадениот датасет на тренинг и тест сет, а како резултат на ова добиваме објективна проценка на неговите перформанси. За тренирање се користи AdamW оптимизатор. Пред почеток на тренирањето, се оценува распределбата на примероците по класи. Поради нерамномерна распределба, на секоја класа и се пресметуваат class weights обратно пропорционални на бројот на примери, по што се аплицира WeightedRandomSampler. Ова обезбедува да се земаат примероци со веројатност пропорционална на нивната тежина при создавање на секој батч, со што се избегнува доминација на честите класи во учењето. За функција на губење се користи модификација на крос-ентропијата (Focal Loss) која дополнително ги балансира тешките примероци.

Тренирањето се одвива во 50 епохи со можност да се запре рано доколку точноста на валидационата партиција пред неколку последователни епохи. По секоја епоха се пресметува точноста на тренинг и тест сетот. Моделот е обучен на податоци од сесии 1 и 2, а тестиран на податоци од сесија 3. За евалуација искористена е метриката асигуасу (точност) која се дефинира како однос на точно класифицираните примери кон вкупниот број примери. Дополнително се користат и confusion matrix и classification report кој содржи содржи точност, прецизност, recall и F1 мера за секој гест.

Во финалните резултати, моделот постигна тест-точност од околу 0.72(72%), што значи дека приближно 72% од прозорците се правилно класифицирани. Овие резултати се споредуваат со очекувањата од литературата за слични задачи на препознавање на гестови од ЕМГ, при што добиениот резултат укажува на корисноста на презентираната методологија. Резултатите покажуваат дека се постигнува задоволителна точност во класификацијата, иако постојат одредени гестови кои се потешко се разликуваат токму поради сличноста на мускулната активност.

Споредба на модели

Во рамките на оваа методологија беа развиени два модела, нашиот финален модел LightweightCNN – длабока невронска мрежа со 1D конволутивни слоеви, но исто така беше развиен прв модел во првиот обид – класичен пристап на извлекување на рачно дефинирани карактеристики и повеќеслоен перцептрон(MLP). Првиот модел следи традиционален пристап за анализа на биомедицински сигнали. Тој користи седум карактеристики по еден канал, така што имаме вкупно 196 карактеристики (28 канали x 7 карактеристики). Неговата архитектура е отсликана преку повеќеслоен перцептрон со четири скриени слоја, користејќи ReLU активација, BatchNorm и Dropout за регуларизација. Од друга страна вториот(финалниот) модел претставува хибриден иновативен пристап со двоен влез:

- прв влез : сурови ЕМГ сигнали
- втор влез : рачно дефинирани карактеристики

Предностите на двоениот влез се :

- комплементарност : CNN гранката кај овој модел открива комплексни темпорални шеми, додека рачните карактеристики обезбедуваат стабилни статистички дескриптори
- робусност : доколку една гранка не успее да открие релевантни шеми, следната гранка може да компензира
- мултискална анализа : дилатираните конволуции заземаат различни временски скали

Перформанси – резултати на тестирање

Модел	Пристап	Accuracy	F1-score (weighted avg)	Recall (weighted avg)
Прв модел (MLP)	Per-window (raw)	0.561	0.560	0.561
	Per-window (smoothed)	0.612	0.605	0.612
Втор модел (CNN/SVM)	Per-window (raw)	0.704	0.699	0.704
	Per-window (smoothed)	0.722	0.717	0.722

Резултатите јасно потврдуваат дека **вториот модел** има значителна предност во однос на **првиот модел**. Кај првиот модел точноста се движи околу **56%–61%**, додека финалниот модел постигнува **70%–72%**. Сличен тренд се забележува и кај F1-score и Recall, каде вториот модел доследно остварува повисоки вредности.

Оваа споредба несомнено покажува дека хибридниот модел значително ги надминува перформансите на првиот модел. Имено оваа предност се должи на неколку клучни фактори :

- автоматско учење на карактеристики - CNN компонентата открива комплексни темпорални шеми кои не се заземени со традиционални статистички мерки
- мултимодален пристап – комбинацијата на рачно дефинирани и автоматски карактеристики обезбедува робусност
- ефикасна архитектура - употребата на depthwise separable конволуции овозможува длабока мрежа без прекумерно зголемување на комплексноста

Добиените резултати укажуваат дека идните истражувања подобро е да се насочат кон хибридни пристапи кои ќе бидат еден вид мост-спој на автоматското учење со доменското знаење интегрирано во рачно дефинирани карактеристики.

Заклучок

Од сите досегашни направени чекори во оваа методологија полека, но сигурно доаѓаме до јасен заклучок дека со внимателно претпроцесирање на податоците и примена на хибридниот модел ефикасно и успешно може да се реши задачата за класификација на ЕМГ сигнали. Мостот изграден како спој меѓу традиционалните карактеристики со длабоката LightweightCNN архитектура донесе висока точност, стабилност и задоволност во резултатите, што уште еднаш го потврдува потенцијалот на овој пристап. Постигнувањето на точност над 70% ја демонстрира променливоста на методологијата во реални услови, а она што е добиено е само една солидна основа за идни подобрувања. Односно, оставен е простор за понатамошни подобри решенија преку: зголемување на бројот на учесници и сесии за подобрена генерализација и користење на крос-валидација за пообјективна евалуација. Како што многумина велат: „Секое истражување не претставува крајна точка, туку темел врз кој се градат идните откритија“.