

TDK-dolgozat

Zsoldos Péter

Elektromiográfiai jel alapú kézmozdulatfelismerés valós időben

EÖTVÖS LORÁND TUDOMÁNYEGYETEM

INFORMATIKAI KAR

MESTERSÉGES INTELLIGENCIA TANSZÉK



Szerző:

Zsoldos Péter

Programtervező informatikus BSc

2. évfolyam

Témavezető:

Dr. Nagy Balázs

egyetemi adjunktus

Budapest, 2023.12.03.

Tartalomjegyzék

1. Absztrakt	3
2. Bevezetés	4
3. Technológiai áttekintés	7
3.1. Myo karpánt	8
3.2. A kézmozdulatok	10
3.3. Mélytanulás	10
3.3.1. Perceptron	11
3.3.2. A neurális rétegek	12
3.3.3. Teljesen kapcsolt rétegek	12
3.3.4. GRU	13
3.4. Tensorflow	14
3.5. Kiértékelési metrikák	14
4. Megvalósítás és alkalmazás	16
4.1. Adathalmazok	16
4.2. Előfeldolgozás	18
4.2.1. Jel szétválasztás	19
4.2.2. Filterek	19
4.3. Használt modell	19
4.3.1. Architektúra	20
4.3.2. Hibafüggvény és optimalizáló algoritmus	21
4.3.3. Hiperparaméterek	22
5. Eredmények és konklúzió	23
5.1. Online adathalmaz	24
5.2. Pangestu adathalmaz	26
5.3. Megnyújtott időablak	28

5.3.1. Eredmények az online adathalmazon tanított modelltől	28
5.3.2. Eredmények a Pangestu adathalmazon tanított modelltől	30
5.4. Konklúzió	32
5.5. Kitekintés	33
Köszönetnyilvánítás	34
Irodalomjegyzék	34
Ábrajegyzék	38
Táblázatjegyzék	39

1. fejezet

Absztrakt

Az intuitív ember-gép interfészek iránti növekvő igény a kézmozdulat-felismerés területén való kutatás virágzásához vezetett, alkalmazási területei a virtuális valóságtól a protézisek egy egészen új generációjáig terjednek. Ez a dolgozat egy valós idejű kézmozdulat-osztályozó rendszer fejlesztésére fókuszál, amely az elektromiográfiában (EMG) ismert jeleket használja. Az EMG jelek rögzítik az izmok elektromos aktivitását a kontrakciók során, és közvetlen, valamint hatékony módot kínálnak a kézmozdulatok dekódolására.

A javasolt rendszer újszerű algoritmust alkalmaz az EMG jelekből történő jellemzőkivonáshoz, amely az időbeli és spektrális jellemzőket hangsúlyozza, hogy megragadja a kézmozdulatok árnyalt dinamikáját. A Myo karszalag, egy fogyasztói minőségű, EMG jelek rögzítésére tervezett hordozható eszköz, amely biztosítja az alkar izomaktivitásának pontos és valós idejű rögzítését. Ezen eszköz általi mérések segítségével a dolgozat megírása során, gépi tanulást alkalmazva, egy a kézmozdulatok változatos adathalmazán tanított robusztus osztályozási rendszert alkottam meg. A rendszer úgy lett megtervezve, hogy valós időben működjön, lehetővé téve a zökkenőmentes integrációt olyan alkalmazásokba, amelyek gyors és pontos gesztusfelismerést igényelnek.

A kísérleti eredmények magas pontosságot és alacsony késleltetést mutatnak, ami a rendszer hatékonyságát a való életben történő alkalmazásokban bizonyítja. A javasolt megközelítés sokoldalúsága miatt számos különböző területen történő felhasználásra alkalmas, beleértve az ember-gép interakciót, a virtuális valóságot és a segítség nyújtó technológiákat a motoros rendellenességgel élő személyek számára. Ez a kutatás hozzájárul a kézmozdulat-alapú interfészek fejlesztéséhez, javítva a felhasználói élményt és a felhasználók számára való elérhetőséget a különböző területek széles palettáján.

2. fejezet

Bevezetés

A gesztusfelismerés, egy gyorsan fejlődő és interdiszciplináris terület, amely az ember-számítógép interakció élvonalában áll. Elsődleges célja, hogy a szándékos kézmozdulatok dekódolásával zökkenőmentesen összekapcsolja a felhasználókat a gépekkel, és ezáltal kulcsfontosságú csatornaként szolgáljon az emberi kifejezés és a digitális parancsértelmezés között. Ez a bonyolult folyamat magában foglalja a gesztusok megkülönböztetését, kategorizálását, hatékonyan lefordítva az emberi kifejezéseket olyan parancsokká, amelyeket a digitális rendszerek képesek értelmezni és végrehajtani.

A gesztusfelismerő rendszerekben alkalmazott módszerek a legkülönbözőbb technológiákat foglalják magukba, a kamerától és érzékelőktől kezdve a jelen dolgozatban használt elektromiográfiáig (EMG [1]). Ez a kifinomult eszközkészlet megkönnyíti a kézmozdulatokban rejlő térbeli és időbeli jellemzők rögzítését és elemzését.

A gesztusfelismerő technológiában rejlő lehetőségek messze túlmutatnak a jelenlegi alkalmazásokon és a felhasználói felületek forradalmi átalakulását ígérik. E technológia elfogadásával egy természetesebb és intuitívabb interakciós paradigma felé nyitjuk meg az utat, amely alapvetően átalakítja az egyének digitális rendszerekkel való kapcsolatát. A gesztusfelismerés alkalmazhatóságát jól példázza a felhasználási területeinek széles skálája, amely a magával ragadó virtuális valóságélményektől [2] a fogyatékkal élőkre szabott, élvonalbeli segítő technológiáig terjed [3]. A gesztusfelismerés lényegében a lehetőségek széles tárházát nyitja meg, és számos területen javítja a felhasználói élményt és a hozzáférhetőséget. Hatása áthatja a különböző területeket, és olyan innovatív megoldásokat kínál, amelyek újradefiniálhatják a technológiával való interakcióinkat.

A gesztusokat (továbbiakban kézmozdulatok) EMG jelek alapján klasszifikáltam ebben a dolgozatban, annak kiemelkedő alkalmassága miatt erre a feladatra [1]. Az EMG [4]

az izomaktivitás elektromos dinamikáját vizsgáló kulcsfontosságú terület, amely mélyreható betekintést nyújt az idegek és az izmok közötti bonyolult kölcsönhatásba. Az EMG során elektródákat helyeznek a bőrre vagy finom tűket szúrnak az izmokba, hogy érzékeljék és elemezzék az izomösszehúzódás során keletkező elektromos jeleket. Az EMG értékes a neuromuszkuláris rendellenességek diagnosztizálásában, az idegműködés értékelésében és az izomaktivitás klinikai megfigyelésében. A felületi EMG kevésbé invazív, míg a tű EMG részletesebb információkat nyújt. A dolgozat megírásához használt adathalmazok kizárólag felületi EMG [5] jeleket tartalmaznak.

Az EMG-osztályozás terén jelentős előrelépések tapasztalhatók, azonban nem mentes a kihívásoktól és problémáktól. Az EMG jelfeldolgozás néhány kulcsfontosságú kihívása az alábbi:

1. Áthallás (Crosstalk):

Az áthallás [6] a szomszédos izmok interferenciájára utal. Ez az izomaktivitás félreértelmezéséhez vezethet, és ronthatja az osztályozó algoritmusok pontosságát.

2. Nem-stacionaritás:

Az EMG jelek nem stacionáriusak [7], ami azt jelenti, hogy statisztikai tulajdonságaik idővel változnak. Az olyan tényezők, mint a fáradtság, az elektródák elmozdulása és az izomaktivációs minták változásai hozzájárulnak a nem-stacionaritáshoz. Az osztályozási algoritmusok adaptálása e dinamikus jelleg kezeléséhez állandó kihívást jelent.

3. Zaj:

Az EMG-jelek érzékenyek a különböző típusú zajokra, beleértve az elektromos zajt [8] és a mozgási artefaktumokat. Ezek a zajok eltorzítják az adatokat és hibákat vihetnek be az osztályozási folyamatba, ezzel óriási problémákat okozva egy ennyire változatos természetű jelenség klasszifikálásánál, mint az izomaktivitás elektromos jele.

4. Tanító adatok korlátai:

A jó minőségű tanító adatok rendelkezésre állása döntő fontosságú a robusztus EMG osztályozási modellek kialakításához. A lemért alanyoknak széles felhasználói réteget kell lefedniük, hogy kellően általános megoldás születhessen. Nagyon fontos hogy változatos korú és nemű alanyok szerepeljenek a tanító halmazokban, hiszen ezen tulajdonságok erősen befolyásolják az izomaktivitást. Fontos továbbá

a megfelelően címkézett adathalmaz használata, hiszen a kellő pontossághoz megfelelően frekvenciált mérésekre van szükség, ami érzékennyé teszi a címkék helyességét az emberi reakcióidőre. A kézmozdulatok tranziens szakaszainak pontos definiálása szintén egy lényeges kérdés. Akkor végződött-e egy kézmozdulat, amikor megállításra kerül a mérés vagy fél másodperccel előtte, ami 5 ezredmásodperces mérésnél 100 hibás címkét eredményez.

5. Valós idejű feldolgozás:

Az EMG osztályozás számos alkalmazása, például a protézisek vagy az ember-számítógép interfészek valós idejű feldolgozást igényelnek [9]. Az alacsony késleltetésű, nagy sebességű osztályozás elérése a pontosság fenntartása mellett olyan technikai kihívás, amellyel a kutatók és mérnökök továbbra is foglalkoznak.

6. A felhasználó variabilitása:

Az egyének jelentős eltéréseket mutathatnak az izom anatómiájában, fiziológiájában és mozgásmintáiban. Olyan osztályozási modellek tervezése, amelyek képesek figyelembe venni ezt az alanyok közötti variabilitást, állandó kihívást jelent. E kihívások kezelése multidiszciplináris megközelítést igényel, amely magában foglalja a jelfeldolgozást, a gépi tanulás, a biomechanika és a klinikai ismereteket.

Az alább ajánlott modell ezen kihívásokkal hivatott megküzdeni, ezért már tervezésekor nagy figyelmet fordítottam az ismeretlen adathalmazon való teljesítményre [10] és a valós idejű használhatóságra.

3. fejezet

Technológiai áttekintés

A Dolopikos és tsai. [11] kutatás a kézmozdulatokat reprezentáló EMG adatok valós idejű osztályozására összpontosít, egy 15 alanyon gyűjtött adathalmaz segítségével. Ez az adathalmaz ennek a dolgozatnak a megírásához is fel lett használva. A tanulmány egy olyan kalibrációs eljárást mutat be, amely 67,87%-ról 91,93%-ra javítja az osztályozási pontosságot. A klasszikus gépi tanulási modellek, mint például a Random Forest és a Support Vector Machine, egyenként hozva döntést egy többségi szavazás során állítanak ki egy együttes döntést a kutatásban, amely módszerrel felülmúlják a Deep Neural Network (DNN) teljesítményét, de ez nem igaz a kalibráció nélküli valós idejű alkalmazásnál, ahol a DNN pontossága felülmúlja ennek a megközelítésnek a pontosságát. Az adathalmaz öt gesztust tartalmaz (ujjak széttárása, integetés-kifelé, integetés-befelé, ökölbe szorított kézfej és null, vagyis semleges, nyugalmi állapot), amelyeket egy Myo Armband segítségével mértek le. A kutatás hangsúlyozza a modellkalibráció fontosságát a valós idejű EMG osztályozáshoz, és vizsgálja az többségi szavazás alapján döntést hozó modellek használatát.

A tanulmány kitér a folyamatban lévő világjárvány valós idejű ellenőrzésre gyakorolt hatására is, és a szükséges berendezések elérhetetlensége miatt valós idejű emulációs rendszert mutat be.

Ez a dolgozat csak a mély neurális hálókkal fog foglalkozni a témája és azon ok miatt, hogy ebben a kapcsolódó munkában a többi mélytanulási modell együttes erejét is felülmúlta az említett típusú modell ismeretlen adathalmazon tesztelve, így alkalmasabbnak bizonyult a valós idejű használatra hiszen jobban kezeli a felhasználók variabilitását.

Továbbá elhagyásra kerül a null, vagyis semleges kézmozdulat, ugyanis ez könnyen helyettesíthető egy olyan egyszerű algoritmussal, ami csak akkor engedi a modellt dön-

tést hozni, ha a nyolc szenzorból érkező értékekből legalább pár egyszerre átlép egy előre beállított határt ezzel igazolva, hogy felismerendő kézmozdulat történt, nem csupán a teljes kart mozgatták, vagy zaj keveredett a jelbe. Alábbi 3.1 és 3.2 táblázatokból kiolvasható az említett kutatásban előállított legjobb modell teljesítménye. Az elektroencefalográfia (EEG) súlyok az alábbi táblázatban a Dolopikos és tsai. tanulmány egyik módszere, mely szerint egy szintén elektromos izomaktivitást mérő módszerre (EMG) támaszkodó klasszikációs modellnek a súlyait finomhangolják tovább. A tanulmányban olvasható ennek indoklása, mivel én nem használtam itt nem fejttem ki.

Modell	Pontosság
Random súlyok	67.00
EEG súlyok	68.01
EMG súlyok	68.61
EMG kalibrációs súlyok	70.63

3.1. táblázat. Dolopikos és tsai. eredményei mély neurális hálónál ismeretlen adathalmazra

	ököl	ujjak szét	intés be	intés ki	null
ököl	0.77	0.01	0.02	0.00	0.21
ujjak szét	0.00	0.14	0.29	0.25	0.32
intés ki	0.00	0.00	0.88	0.00	0.12
intés be	0.00	0.09	0.01	0.87	0.04
null	0.00	0.25	0.38	0.05	0.31

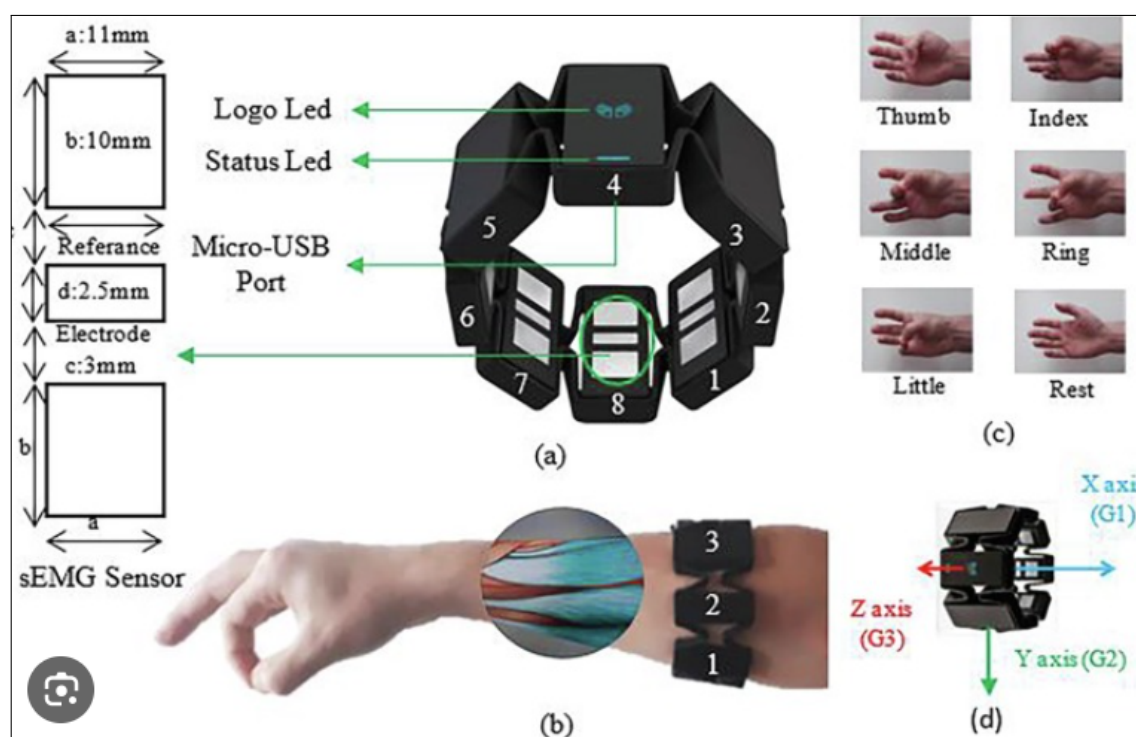
3.2. táblázat. Dolopikos és tsai. eredményeinek konfúziós mátrixa mély neurális hálónál ismeretlen adathalmazra

3.1. Myo karpánt

A Myo [12] karpánt, amely forradalmi újítás a viselhető technológiák terén, az EMG technológia zökkenőmentes integrációjával az ember-számítógép interakció élvonalába tartozik. A Thalmic Labs (ma már a North Inc. része) által kifejlesztett kompakt és stílusos eszköz az alkaron viselhető, és az izomösszehúzódások által generált elektromos jeleket rögzíti és értelmezi, lehetővé téve az elektronikus eszközök pontos, gesztusalapú vezérlését.

A Myo karpánt lényege, hogy a kialakításába ágyazott EMG-érzékelők kifinomult csoportját alkalmazza. Ezek az érzékelők érzékelik és elemzik az izomműködés által generált apró elektromos impulzusokat, és ezeket olyan adatsorokká alakítják át, amelyek a kéz és a kar konkrét mozgását írják le. Ez az innovatív megközelítés intuitív eszközt biztosít a felhasználók számára a különböző digitális platformokkal való interakcióhoz, az asztali számítógépektől és okostelefonoktól kezdve a virtuális valóságot alkalmazó szemüvegekig.

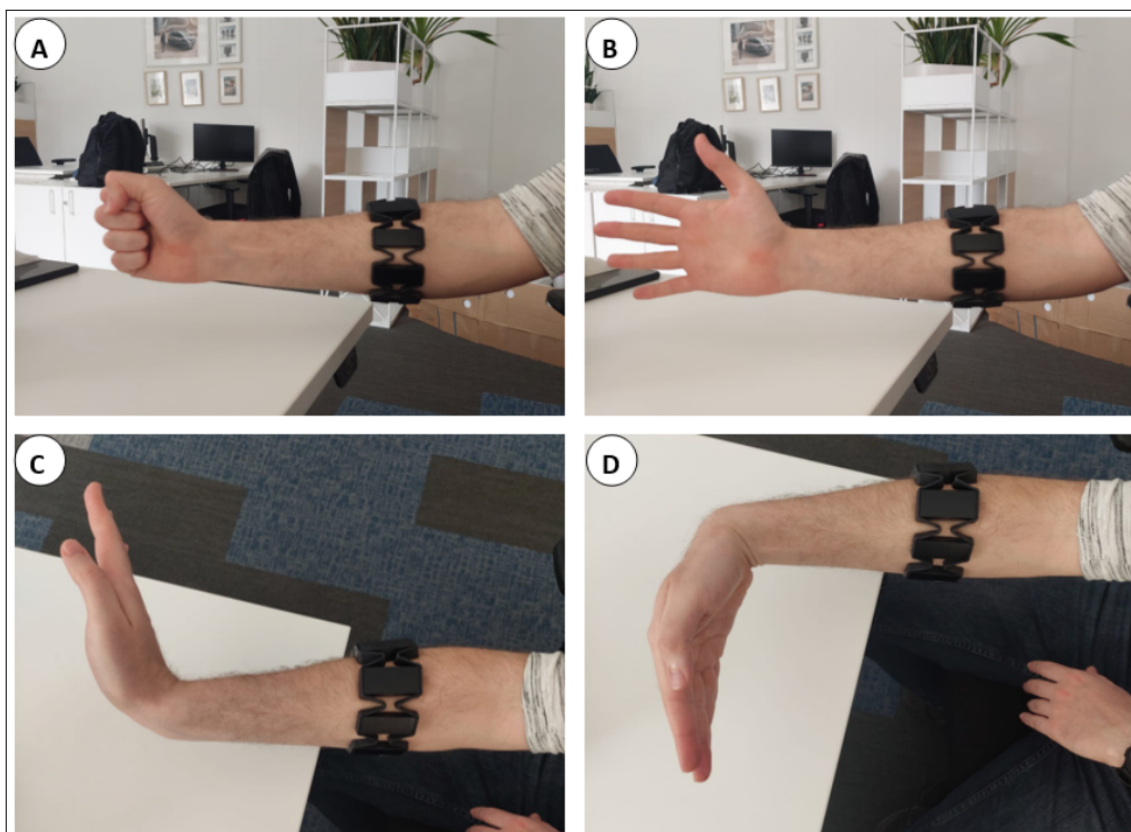
A Myo karpánt egyik kiemelkedő jellemzője a sokoldalúsága. Az eszköz nyolc izomjelek mérő szenzora elegendő információt biztosít a gesztusfelismeréshez, amely lehetővé teszi a felhasználók számára, hogy egyszerű kézmozdulatokkal hajtsanak végre parancsokat, átalakítva az ember technológiával való kapcsolódásának módját. Legyen szó alkalmazásokban való navigálásról, multimédiás lejátszás vezérléséről vagy akár drónok irányításáról [12]. A Myo karpánt a fogyasztói alkalmazásokon túl a legkülönbözőbb professzionális környezetekben is használható [3]. Az alábbi 3.1 ábrán látható a karpánt felhelyezése, méretei, valamint sokoldalúságát szemléltetve kézmozdulatok, melyeknek felismerésében kulcsfontosságú ebben a dolgozatban, illetve a beépített giroszkóp.



3.1. ábra. A Myo karpánt és tulajdonságai.

3.2. A kézmozdulatok

A problémák megoldására általam használt módszer erősen támaszkodik egy ezt megelőző kutatásra (Dolopikos és tsai. [11]) és annak tanító és tesztelő adathalmazát felhasználja, valamint egy alternatív megközelítést biztosít a kalibrációra, amely nem minden esetben megengedhető, ha valós idejű alkalmazásról van szó. Az említett kutatásban öt különböző kézmozdulat van különválasztva és ebből egyik a nyugalmi állapot. Ezt én is elegendőnek ítéltam, illetve a nyugalmi állapotot nem klasszifikáltam, mert az egyszerűen megoldható az EMG jel aktivációját figyelve, elér-e egy bizonyos határértéket legalább bizonyos számú szenzorokból érkező jel, ha igen volt kézmozdulat, ha nem akkor nyugalmi állapot van. A feladat során a Myo karpánt segítségével megkülönböztetett négy kézmozdulatot a 3.2. ábra szemlélteti.



3.2. ábra. A kézmozdulatok (A) öklbe szorított kézfej, (B) ujjak széttárása, (C) intés kifelé, (D) intés befelé

3.3. Mélytanulás

A mélytanulás lényege, hogy a rétegekbe szervezett neurális hálózatokra támaszkodik az információk hierarchikus módon történő feldolgozásához. Ezeknek a hálózatoknak a

mélysége lehetővé teszi az összetett jellemzők kinyerését a nyers adatokból, ami a gépi tanulás hatékony eszközei közé emeli.

A neurális hálózatok anatómiája utánozza az emberi agy szinaptikus kapcsolatait, ezzel érve el korábban nem látott sikereket nehéz problémák megoldásában [13]. A backpropagation által vezérelt tanítási folyamat lehetővé teszi, hogy ezek a hálózatok címkézett adathalmazokból tanuljanak, a súlyokat (a neuron kapcsolatokat leíró számokat) a hibák minimalizálása érdekében állítsák be, és növeljék a képességüket, hogy pontos előrejelzéseket készítsenek új, nem látott adatokra.

A mélytanulás azonban nem mentes a kihívásoktól. A hatalmas adathalmazok és számítási teljesítmény iránti kielégíthetetlen igény, a túlillesztés állandó problémája és a modellek értelmezhetősége folyamatos akadályokat jelent.

3.3.1. Perceptron

A perceptron a neurális hálózatok és a gépi tanulás egyik alapkoncepciója. Egyszerű bináris osztályozáshoz alkalmas elem, ami több bemenetet dolgoz fel azáltal, hogy súlyokat rendel hozzájuk, és a súlyozott összegük egy meghatározott küszöbértékkel eltolja. A kapott értéket egy függvénynek adja paraméterül, melynek kimenete meghatározza hogy a bináris klasszifikáció melyik osztályába esik a bemenet. A 3.1 matematikai képlet magában foglalja ezt a folyamatot. A perceptron tanítása a súlyok beállítását jelenti, amivel minimalizálja a hibákat az előrejelzésekben.

$$y = \sigma(wx + b) \quad (3.1)$$

ahol w a súlyok, x a bejövő adatok és b bias, pedig egy határ, ami azt kívánja módosítani, hogy tanulahtó legyen, milyen adatokra aktiválódik a perceptron, σ pedig csak egy aktivációs függvény, mely leírja, hogy milyen értéket kívánunk hozzárendelni az egyes súlyok és a bias által módosított inputokhoz. Míg egyetlen perceptron alapegységként működik, valódi jelentősége akkor bontakozik ki, amikor rétegekbe kombinálva összetettebb struktúrákat alkot, mint például a neurális hálózatokon belül a többrétegű perceptronok (MLP-k). Ezen összevonás révén a neurális hálózatok képessé válnak arra, hogy bonyolult mintákat és reprezentációkat tanuljanak az adatokból, és így képesek legyenek összetett problémák megoldására.

A perceptron, egyszerűsége ellenére, kulcsfontosságú szerepet játszott a modern gépi tanulás és a mélytanulás kialakításában. Előfutárként szolgált a fejlettebb neurális há-

lózati architektúrákhoz, bemutatva az evolúciós utat az alapvető bináris döntéshozataltól az összetett, adatvezérelt modellekig, amelyek a mai mesterséges intelligencia alapját képezik. A perceptron megértése kulcsfontosságú betekintést nyújt a neurális hálózatok képességeit vezérlő alapelvekbe, megnyitva az utat az intelligens rendszerekben rejlő lehetőségeket továbbra is újradefiniáló innovációk előtt.

3.3.2. A neurális rétegek

A neurális rétegek alkotják a mesterséges neurális hálózatok alapvető keretét, és az információ áramlásának és feldolgozásának csatornáiként szolgálnak. A bemeneti réteg dolgozza fel a bemeneti adatokat, míg a rejtett rétegek a súlyozott összegek és aktiválási függvények révén hierarchikus jellemzőket és összetett reprezentációkat vonnak ki. A kimeneti réteg állítja elő a végső eredményt vagy előrejelzést. A neuronok kapcsolatainak erősségét és reakciókészségét a súlyok és az előfeszítések szabályozzák.

A neurális rétegekben belüli, súlyozott összegeket és aktiválási függvényeket tartalmazó feldolgozás nemlinearitást hoz létre, és lehetővé teszi a hálózat számára, hogy bonyolult mintákat tanuljon. Fontos része, a neurális hálózatok tanulási mechanizmusának a súlyok és az bias-ok iteratív beállítása a tanítás során, amelyet a backpropagation segítségével hajtunk végre. Ez az algoritmus azt adja meg, hogy melyik súlyt, mennyivel kell állítanunk a pontosabb előrejelzéshez.

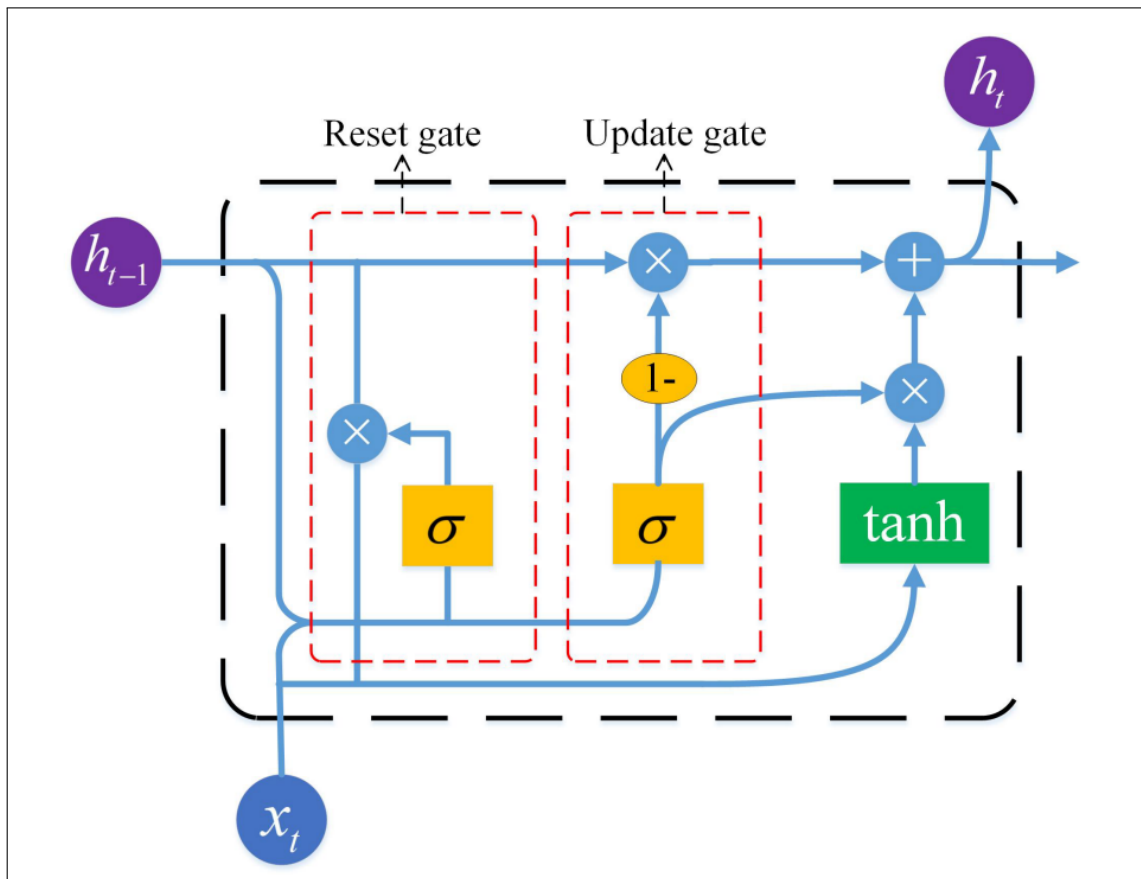
3.3.3. Teljesen kapcsolt rétegek

A teljesen kapcsolt rétegek (fully connected layers - FC) dinamikus és hatékony struktúra a mesterséges neurális hálózatok területén. Olyan perceptronokból állnak melyek kapcsolódnak a következő réteg perceptronjaihoz és az előzőiéhez, de semmiképpen sem egymáshoz, ezzel megakadályozva hogy egy absztrakciós szinten a beérkező adatok befolyásolják az absztrakciós folyamatot. A teljesen kapcsolt rétegek sokoldalúan alkalmazhatók a legkülönbözőbb alkalmazásokban, kiválóan teljesítenek a képfelismerésben, a természetes nyelvi feldolgozásban, a pénzügyi előrejelzésben és az egészségügyben. Az olyan kihívások azonban, mint a túlillesztés, továbbra is fennállnak, ami olyan fejlett technikák kifejlesztésére késztet, mint a regularizáció és több alternatív aktivációs függvény.

3.3.4. GRU

A GRU [14], mint a rekurrens neurális hálózatok egy formája, egy olyan kapus mechanizmust vezet be, amely szabályozza az információáramlást az egyes időlépések között. Előnye, többek között a gyorsabb konvergencia a tanítás során és az eltűnő gradiens problémájának enyhítése, ami alkalmassá teszi őt olyan alkalmazásokhoz, mint a természetes nyelvi feldolgozás, az idősor-előrejelzés és a beszédfelismerés. Hatékonysága ellenére olyan kihívások ismertek el, mint a modell értelmezhetősége és a hiperparaméterekre való érzékenység.

Összefoglalva, a GRU-k befolyásos szereplőként jelennek meg a neurális hálózatok dinamikus világában, folyamatos innovációt és áttörést ígérve a szekvenciális adatfeldolgozás területén. A 3.3 ábrán látható, hogyan helyezkednek el egymáshoz viszonyítva a kapuk, melyek az információ áramlását szabályozzák (Update gate, Reset gate).



3.3. ábra. Gated Recurrent Unit (GRU) architektúra [15]

3.4. Tensorflow

A TensorFlow [16] a Google Brain által kifejlesztett nyílt forráskódú gépi tanulási könyvtár, amelyet széles körben használnak gépi tanulási modellek implementálására. A rugalmasságáról, skálázhatóságáról és átfogó ökoszisztémájáról ismert TensorFlow változatos feladatokat támogat, a képfelismeréstől a természetes nyelvi feldolgozásig. A Keras-hoz hasonló magas szintű API-k egyszerűsítik a modellfejlesztést, és a TensorFlow népszerűségéhez hozzájárul az aktív közösség is. A keretrendszer funkciói közé tartozik a TensorBoard a vizualizációhoz, a TensorFlow Lite az edge eszközökhöz, és a TensorFlow.js a böngészőalkalmazásokhoz. Összességében a TensorFlow egy robusztus eszköz, amely a gépi tanulás területén dolgozó kutatók és fejlesztők számára nyújt számos lehetőséget. A kutatás során ezen könyvtár segítségével végeztem a modell definiálását és tanítását.

3.5. Kiértékelési metrikák

A modell értékeléséhez, a pontosságot (accuracy), a precizitást (precision) és a F1-score-t használtam. A pontosság az osztályozási eredmények általános helyességét méri a helyesen osztályozott minták arányának kiszámításával. Általános jelzést ad arról, hogy a modell mennyire pontosan jósolja meg az osztályokat. A precízió a modell által tett összes pozitív előrejelzésen belül az igaz pozitív előrejelzések arányát méri, és a pozitív előrejelzések helyességére összpontosít. Az F1-score azt hivatott jelezni, hogy ez a két metrika mennyire állnak közel egymáshoz. A 3.2-3.4 egyenletek az egyes metrikák számítását írják le.

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \quad (3.2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3.3)$$

$$F1_score = 2 \times \frac{Precision \times Accuracy}{Precision + Accuracy} \quad (3.4)$$

Igaz pozitív (True Positive, TP), ha a modell helyesen jósolja meg a pozitív osztályt. Hamis Pozitív (False Positive, FP) az, amikor a modell tévesen jósolja meg a pozitív osztályt, amikor a negatív osztályt kellett volna megjósolnia. Hamis negatív (False Negative, FN) az, amikor a modell tévesen a negatív osztályt jelzi előre, amikor a pozitív osztályt kellett volna. Igaz negatív (True Negative, TN), amikor a modell helyesen jelzi előre a negatív osztályt.

Az utolsó metrika pedig a konfúziós mátrix, ami gépi tanulás alapvető eszköze, amelyet kifejezetten az osztályozási modellek teljesítményének értékelésére terveztek. Olyan táblázatos ábrázolásként szolgál, amely rendszerezi a modell előrejelzéseinek eredményeit, és betekintést nyújt abba, hogy a modell mennyire jól különbözteti meg a különböző osztályokat egy adathalmazon belül. Ezek a metrikák az 5. fejezetben lesznek felhasználva.

4. fejezet

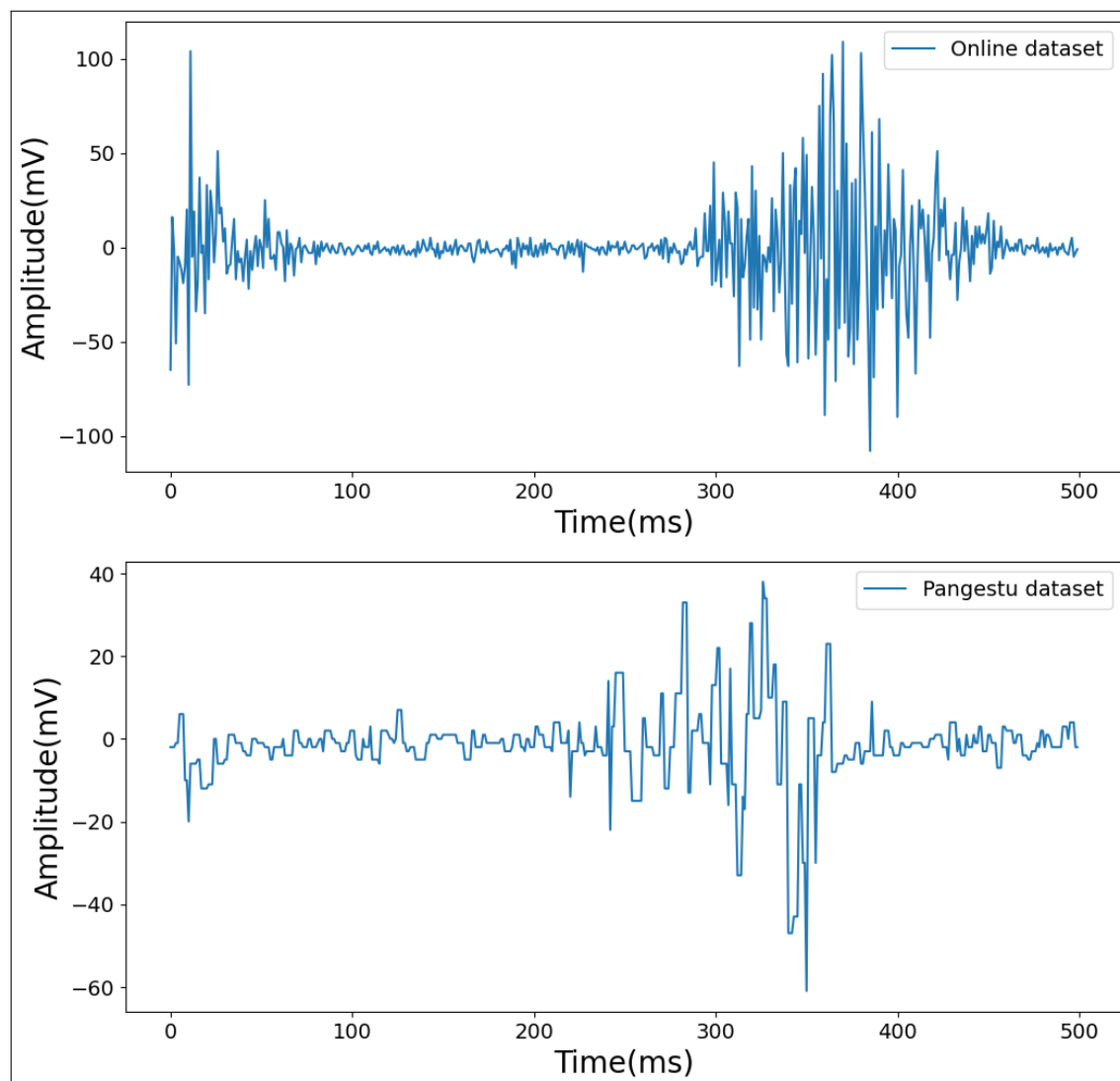
Megvalósítás és alkalmazás

A neurális hálózat megvalósítása kulcsfontosságú tényező, amely befolyásolja annak teljesítményét és hatékonyságát. A legfontosabb szempontok közé tartozik az architektúra megtervezése, a hiperparaméterek gondos beállítása, a tanítási adatok előfeldolgozása, a megfelelő súlyinicializálás és aktivációs függvények óvatos megválasztása, a hatékony tanítási időtartam és a megfelelő tanítási stratégia. A jól átgondolt megvalósítás elengedhetetlen ahhoz, hogy a neurális hálózatokban rejlő teljes potenciált kiaknázzuk a különböző területeken felmerülő összetett problémák megoldásában.

4.1. Adathalmazok

Ez a kutatás két adathalmazt használ: egy online adathalmazt (Dolopikos és tsai.[11]), amely a Kaggle online platformon lett megosztottva, és egy saját adathalmazt, melyet Wijaya Feki Pangestu gyűjtött, szintén Myo karpánttal (továbbiakban Pangestu adathalmaz). Ez a tanulmány kombinálja az online és a saját gyűjtésű adathalmazokat, hogy nagy mennyiségű, különböző forrásokból származó adatot szolgáltatson és ezáltal a modell robusztusabbá tételét támogassa. Ennek a kombinációnak formája, hogy külön-külön kerül tanításra rajtuk a modell és aztán mindkettő adathalmazra letesztelésre kerül. Bár ezekben az adatkészletekben sok a közös vonás, mint például ugyanazon gesztus gyűjtése, ugyanazon típusú karszalag használata és ugyanazon eljárás az adatgyűjtéshez, mégis vannak különbségek. A fő különbség az adatok minőségében van, az online adathalmaz jelminősége jobb mint az Pangestu adathalmazé. Az online adatkészlet jelének kisebb a zaja és ez a 4.1 ábrán jól látható, melyen két intés befelé mozdulat szerepel a két adathalmaz-

ból. A különbség a Myo karpánt konfigurációja, a kézen való állásának és a környezető tényezőknek tudható be.



4.1. ábra. Különbség a két adathalmaz között, a felső ábra mutatja az online adathalmazt, az alsó pedig a Wijaya Feki Pangestu által gyűjtöttet.

Az online adathalmazt Dolopikos és tsai.[11] gyűjtötték össze kísérletükhöz. Háromféle adathalmaz áll rendelkezésre: az első csak a nyers jeleket tartalmazza, a második egy MATLAB segítségével előfeldolgozott jelet tartalmaz, az utolsó pedig egy egyengetett jelet, amely az előfeldolgozott jel abszolút értékéből áll. Minden típusú adathalmaz 300 CSV fájlból (mintából) áll, 15 résztvevőtől (kilenc férfi és hat nő). Az adathalmazt a résztvevők EMG-jelének rögzítésével gyűjtöttük, miközben négy kézmozdulatot végeztünk: kézfej ökölbe szorítása, ujjak szétnyitása, befelé integetés és kifelé integetés. A CSV fájlban rögzített EMG jel a résztvevő által használt Myo karszalagból származik. Ebben a

kísérletben csak a előfeldolgozott adathalmazt használtam fel, hiszen ez volt megfelelően címkézve a modell tanításához.

A Pangestu adathalmaz következő leírása Wijaya Feki Pangestu szakdolgozatából származik: Az adatgyűjtés egy laptop-hoz csatlakoztatott Bluetooth-adapterrel ellátott Myo karpánt segítségével történt. A kapcsolatot a Myo Python API segítségével érjük el, amely lehetővé teszi a Python számára a karpánt elérését. Tizenegy résztvevő (9 férfi és két nő, 26 és 35 év közötti korú) vett részt a kísérletben. A résztvevők a karpántot a jobb alkarjukon viselték és 65 másodpercig ismételten egyetlen gesztust kellett végrehajtaniuk, és ezt minden egyes gesztus esetében ötször ismételték. A mérés első öt másodperce a kéz "nyugalmi helyzetéből" áll, a maradék 60 másodperc pedig a tényleges kézmozdulat. A Myo karpánt az alkar EMG-jét rögzíti. Az összegyűjtött EMG-adatokat 200 Hz-es frekvencián áramlottak, és minden egyes mérési munkamenetet egy `signal.csv` fájlba mentettünk. Minden `.csv` minta nyolc különböző érzékelőcsatorna jelét tartalmazza.

4.2. Előfeldolgozás

Az adatok előfeldolgozása kulcsfontosságú szerepet játszik a későbbi elemzések vagy gépi tanulási modellek pontosságának, megbízhatóságának és hatékonyságának biztosításában. Ez az előkészítő fázis magában foglalja a nyers adatok tisztítását, átalakítását és elemzésre alkalmas formátumba rendezését, olyan problémák kezelését, mint a hiányzó értékek, a kiugró értékek és az irreleváns információk. Az adatok előfeldolgozásának fontossága abban rejlik, hogy a zaj és az irreleváns jellemzők eltávolításával javítja az adathalmaz minőségét, mérsékli a torzításokat, és javítja a modellek teljesítményét. Ezenkívül megkönnyíti az eredmények jobb értelmezhetőségét, így azok értelmesebbek és jobban hasznosíthatók lesznek. Akár a statisztikai elemzés, akár a gépi tanulás vagy bármely adatvezérelt döntéshozatali folyamat összefüggésében, az adatok előfeldolgozásának jelentőségét nem lehet eléggé hangsúlyozni, ez teremti meg az alapját az értelmes meglátások levezetésének és a robusztus, pontos modellek felépítésének.

Fontos megjegyezni, hogy a módszerben nem kerül használat se gyors Fourier transzformációra, sem pedig normalizálásra. Míg előbbi kifejezetten az EMG klasszifikáció területén ajánlott a szakmai irodalom szerint [17][18], addig utóbbi általános eljárás a neuronhálók tanítása előtti előfeldolgozásban. Azonban ezen dolgozat írása közben mindkettő a modell teljesítményének visszavetésével járt ezért elhagyásuk mellett döntöttem. Előbbit elhagyhatónak ítéltam, hiszen egy elég erős modell maga is képes levonni

a számára fontos jellemzőket az adatokból, nincs szüksége egy előzetes függvény approximációra a gyors Fourier transzformáció által, utóbbi pedig esetünkben nem szükséges, hiszen azonos intervallumon belül, azonos mértékegységben, azonos típusú szenzorok mérik az adatokat.

4.2.1. Jel szétválasztás

Az adatelőfeldolgozási folyamat első lépése a jelek felosztása. A következőkben két részre osztjuk őket: a nulljelre és a gesztusjelre. A nulljel, ahol csak egy "nyugalmi helyzetű" gesztus van, és a gesztusjel, ahol a tényleges kézmozdulat történik. A nulljelet és a gesztusjelet könnyű szétválasztani a Pangestu adathalmaznál, mivel a válaszadót már utasították, hogy végezzen "nyugalmi mozgulatot" az első öt másodpercben, így minden mintánál az első öt másodpercet levághatjuk. Az online adathalmaz azonban bonyolultabb, mert Dolopikos és tsai. azt állítják tanulmányukban, hogy a reakcióidők, amivel a válaszadók reagáltak, a tényleges kézmozdulatra nézve eltérőek. Ennek a problémának a megoldására a MATLAB segítségével már előfeldolgozott adatkészleteket használtam kiértékelési alapként.

4.2.2. Filterek

Az adatok szűrésének fő célja, hogy kizárja az EMG jel zaját, mert amikor egy EMG jelet rögzítenek az izomból, különböző típusú zajok szennyezik azt [19]. A zajok az elektronikus berendezések és fiziológiai tényezők miatt érkeznek. Az EMG-jel jellemzői függenek az alany belső szerkezetétől, beleértve az egyedi bőrképződést, a véráramlási sebességet, a mért bőrhőmérséklet, a szöveti szerkezet (izom, zsír stb.), a mérési helytől, és így tovább. Ez hatással lehet a jellemzők kinyerésének eredményére, és így befolyásolhatják az EMG-jelek diagnózisát. A Butterworth sávszűrőt használtam a zaj csökkentésére a magas és alacsony frekvenciájú szűrők kombinálásával. A szűrők egy tartományon (20 Hz - 100 Hz) belüli frekvenciát engednek át.

4.3. Használt modell

A jó mély neurális hálózat (DNN) modell kiválasztása több okból is kulcsfontosságú a gépi tanulás és a mesterséges intelligencia területén. A jó DNN modell kiválasztásának fontosságát a következő szempontok alapján lehet felvázolni:

- **Teljesítmény:**

Minden gépi tanulási modell elsődleges célja, hogy jól teljesítsen a feladatban, amelyre tervezték. Egy jól megválasztott DNN modell jelentősen növelheti a rendszer teljesítményét az adatokban rejlő minták pontos megragadásával. Ez elengedhetetlen a nagy pontosság eléréséhez és a predikciók hibáinak minimalizálásához.

- **Általánosítás:**

Egy jó DNN modell nemcsak a tanítási adatokon hatékony, hanem jól általánosít a nem látott adatokra is. Az általánosítás kulcsfontosságú ahhoz, hogy a modell képes legyen pontos előrejelzéseket készíteni új, korábban nem látott példákra. A túlságosan összetett modellek jól teljesíthetnek a tanítási adatokon, de nem képesek általánosítani, ami gyenge teljesítményt eredményez az új adatokon.

- **Számítási hatékonyság:**

A DNN modell számítási hatékonysága fontos szempont, különösen a valós alkalmazásokban, ahol az erőforrások, mint például a memória, korlátozottak. A teljesítményt és a számítási költségeket egyensúlyban tartó modell kiválasztása alapvető fontosságú a gyakorlati alkalmazáshoz.

- **Robusztusság:**

Egy jó DNN-modellnek robusztusnak kell lennie a bemeneti adatok változásaira és ellenállónak a zajjal szemben. A robusztus modellek kevésbé hajlamosak a bemeneti adatok kisebb ingadozásai által okozott hibákra, így megbízhatóbbak a valós életben.

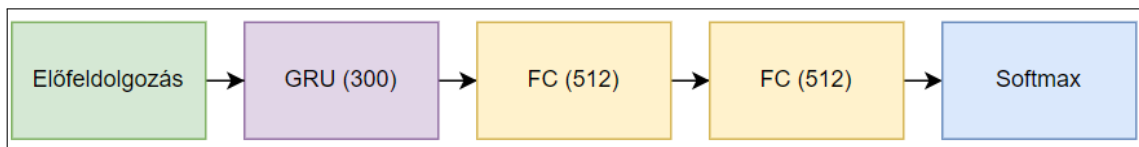
A következő lépésekben ezen feltételek beteljesítésére összpontosítottam.

4.3.1. Architektúra

A GRU rétegek és több fully connected feedforward rétegek fúziója robusztus és sokoldalú neurális hálózati architektúrát hoz létre, amely különösen alkalmas szekvenciális vagy idősoros adatokat tartalmazó feladatokra. A GRU-rétegek egyedülálló kapuadó mechanizmusukkal figyelemre méltó képességet nyújtanak a bonyolult szekvenciális függőségek megragadására. Kialakításuk olyan kihívásokat old meg, mint például az eltűnő gradiens probléma, lehetővé téve a modell számára, hogy hatékonyan megőrizze és felhasználja a múltbeli időlépésekből származó információkat. Ez az időbeli megértés döntő fontosságú olyan forgatókönyvekben, ahol az adatpontok sorrendjének jelentősége van.

A GRU-rétegek kiegészítéseként a teljesen kapcsolt rétegek beépítése az architektúra végére kettős célt szolgál. Először is, nemlinearitást visz be a modellbe, lehetővé téve a GRU-rétegekből tanult jellemzők összetett kapcsolatainak ábrázolását. Ez a nem lineáris leképezés különösen akkor előnyös, ha az adatok mögöttes mintái bonyolultak és a lineáris modellekkel nehéz megragadni őket. Másodszor, a teljesen kapcsolt rétegek jellemző-kivonóként működnek, finomítják és absztrahálják a GRU rétegek által megtanult reprezentációkat.

Az architektúra hierarchikus jellege megkönnyíti a hatékony jellemzőtanulást. A GRU rétegek a rövid és hosszú távú függőségek megragadására összpontosítanak, míg az teljesen kapcsolt rétegek finomítják ezeket a tanult jellemzőket, ami egy robusztus és általánosított modellt eredményez. Az ezen összetevők közötti szinergia fokozza a modell azon képességét, hogy a tanításból származó tudást a nem látott adatokra extrapolálja, így hatékony eszközzé válik olyan forgatókönyvekben, ahol az alkalmazkodóképesség és az általánosítás kulcsfontosságú. Ezt figyelembe véve a rétegek így követik egymást: 300 neuronból álló GRU réteg, 512 neuronból álló teljesen kapcsolt réteg majd még egy 512 neuronból álló teljesen kapcsolt réteg, mindegyik ReLU aktivációs függvényvel, amely 0-át rendel a nem pozitív számokhoz, a pozitívakhoz pedig önmagukat, és végül egy négy neuronból álló softmax (maximum valószínűséget kiválasztó) réteg, melynek neuronjai már az egyes osztályokba tartozás valószínűségét állítják elő. A modell architektúrája a 4.2 ábrán látható.



4.2. ábra. A szekvenciálisan összekötött model ábrázolása.

4.3.2. Hibafüggvény és optimalizáló algoritmus

A kereszt-entrópia hibafüggvény, széles körben használt a gépi tanulás osztályozási feladataihoz. A megjósolt valószínűségek és a valódi osztálycímkék közötti eltérést méri. Bináris osztályozás esetén a képlet a valós címkét (y) és az előre jelzett valószínűséget ($\hat{y}$) veszi figyelembe, míg többosztályos problémák esetén a valós osztálycímkék és az előre jelzett valószínűségek összegét veszi figyelembe az összes osztályra vonatkozóan. A tanítás során a cél ennek a veszteségnek a minimalizálása, arra ösztönözve a modellt, hogy a helyes osztályokhoz magas valószínűségeket rendeljen. A kereszt-entrópia hiba-

függvényt azért kedvelik, mert képes büntetni a magabiztos hibás előrejelzéseket. Ez a fontos tényező tette az osztályozási modellek tanításának szabványos választásává.

A Nadam optimalizáló, a Nesterov-gyorsított adaptív pillanatbecslés rövidítése, egy kifinomult optimalizációs algoritmus, amelyet neurális hálózatok tanítására terveztek. Ez az algoritmus zökkenőmentesen integrálja két kiemelkedő optimalizálási technika erősségeit: A Nesterov-gyorsított gradiens (NAG) és az adaptív pillanatbecslését (Adam). A NAG minimalizálja az oszcillációkat és fokozza a konvergenciát azáltal, hogy a paraméterfrissítéseket megelőzően vizsgálja a gradienst, míg az Adam adaptív tanulási rátákat alkalmaz az egyes paraméterekhez a korábbi gradiensek és a négyzetes gradiensek alapján. A Nadam e módszerek kombinálásával igyekszik egyensúlyt teremteni a stabilitás és a gyors konvergencia között a neurális hálózatok tanításában. Az eredmény egy olyan optimalizáló, amely hatékonyan kihasználja mind a NAG, mind az Adam legjobb tulajdonságait, hozzájárulva a hatékonyság és a teljesítmény javításához az optimalizálási folyamat során.

4.3.3. Hiperparaméterek

A mesterséges intelligenciában és a gépi tanulásban a hiperparaméterek olyan előre meghatározott beállítások, amelyek befolyásolják a tanítási folyamatot. A legfontosabb hiperparaméterek közé tartozik a tanulási sebesség, amely az optimalizálás lépésméretét szabályozza, a neurális hálózatok architektúrája, beleértve a rétegek és egységek számát, a Batch méret, amely meghatározza az egyes iterációkban feldolgozott tanítási példák számát, az tanítási iterációk száma, amely a tanítási adathalmaz teljes átfutását jelenti. A hiperparaméterek hangolása döntő fontosságú a modell teljesítményének optimalizálásához, ami kísérletezést és validálást igényel az adott feladathoz leghatékonyabb beállítások megtalálásához. A használt paraméterek a 4.1-es ábrában találhatók.

4.1. táblázat. A modell hiperparaméterei

Név	Érték
Batch méret	256
Első rejtett réteg mérete	512
Második rejtett réteg mérete	512
Tanítási iterációk száma	10

5. fejezet

Eredmények és konklúzió

Az EMG jelek osztályozásának fejlesztésére tett erőfeszítéseim során az új neurális hálózati modell megvalósítása változást hozott, és számos szempontból felülmúlta elődje képességeit. A modell sikere a tényezők stratégiai kombinációjának tulajdonítható.

A modell kivételes teljesítménye jelentős ígéretet hordoz magában, különösen az olyan alkalmazásokban, mint a teleoperáció és a protézis technológiák, ahol az EMG-jelek pontos osztályozása kritikus fontosságú a zökkenőmentes és érzékeny vezérlőrendszerekhez. A korábbi modellt felülmúlva a megvalósításunk nemcsak kiváló pontosságot mutat be, hanem olyan hatékonysági és megbízhatósági szintet is bevezet, amely jelentősen növeli az EMG-alapú vezérlőrendszerek gyakorlati alkalmazhatóságát. A modell sikere az EMG-jelek bonyolult változásainak kezelésében kiemeli a modell lehetséges hatásait a valós forgatókönyvekre. A távműködtetésben, ahol a pontos jelosztályozás kiemelkedően fontos a távoli eszközök vezérléséhez, a modell példátlan pontosságot kínál. Hasonlóképpen, a protézisek területén, ahol az EMG jelek árnyalt értelmezése kulcsfontosságú a művégtagok természetes és intuitív irányításának biztosításához, a megvalósításom új megoldást kínál.

Ez az neurális háló nemcsak a pontosság határait feszegeti az EMG-jelek osztályozásában, hanem új távlatokat nyit az ilyen technológiák mindennapi életbe való zökkenőmentes integrálása előtt is. Ahogy tovább finomítjuk és bővítjük ezt az innovatív megközelítést, egyre ígéretesebbé válnak a pontosabb és érzékenyebb EMG-jelfeldolgozáson keresztül a teleoperáció és a protézis technológiák lehetőségei.

5.1. Online adathalmaz

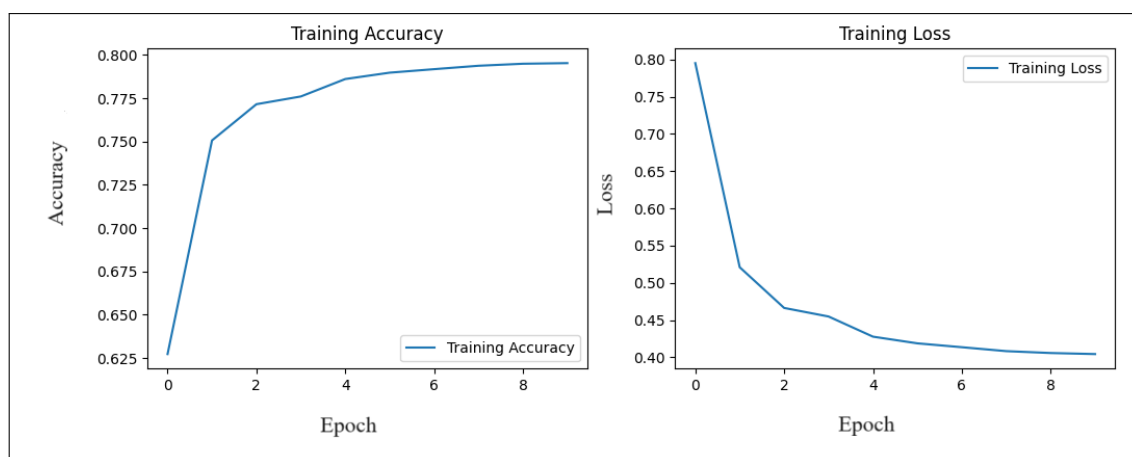
A szigorú tanítás révén a mesterséges intelligencia modell az ismeretlen adathalmazból származó adatokon kivételes klasszifikáló képességéről tett tanúbizonyságot. Ez az eredmény a modell alkalmazkodóképességét és általánosítási képességét mutatja be. A siker a fejlett algoritmusok és egy átfogó adathalmaz aprólékos kombinációjának tulajdonítható, amelynek eredményeképpen egy olyan modell jött létre, amely mind előrejelzési pontosságban, mind robusztus teljesítményben kiemelkedő a forgatókönyvek széles skáláján.

A modellt az online adathalmaz első 10 emberén tanítottam és az utolsó ötön teszteltem illetve az ismeretlen adathalmaz a Pangestu adathalmaz 240 darab 1 másodperces időablakot tartalmazó teszthalmaza. A ismeretlen adathalmaz esetében a pontosság 80,42%, a precízió 80,98%, és az F1-érték 80,59%. Ez azt jelenti, hogy az ismert adathalmazon végzett osztályozás során az esetek 80,42%-át sikerült helyesen osztályozni, a pozitív osztályra vonatkozóan pedig a jósolt értékek 80,98%-a valóban pozitív volt. Az F1-érték 80,59%-os értéke azt mutatja, hogy az osztályozó modell a precízió és a pontosság értékei között jó egyensúlyt talált. Az "Ismert adathalmaz" esetében a pontosság 82,92%, a precízió 84,50%, és az F1-érték 83,18%. Ez azt jelenti, hogy az ismeretlen adathalmazon végzett osztályozás során az esetek 82,92%-át sikerült helyesen osztályozni, a pozitív osztályra vonatkozóan pedig a jósolt értékek 84,50%-a valóban pozitív volt. Az F1-érték 83,18%-os értéke ismét azt jelzi, hogy a modell jól teljesített az egyensúly megtalálásában a precízió és a pontosság között.

Ezek az értékek együttesen nyújtanak átfogó képet arról, hogy a klasszifikációs modell milyen jól teljesít mind a ismert, mind az ismeretlen adathalmazon. A 5.1 táblázat, valamint a 5.1, a 5.2 és a 5.3 ábrák együttesen festenek egy átfogó képet a modell teljesítményéről. A 5.1 ábrán jól kivehető a pontosság növekedése majd 80% körüli megállása. A költség függvény pedig hasonlóan viselkedik ám növekedés helyett csökken, ahogyan elvárható, hiszen csökken a predikciók és a valós címkék közti eltérés és 0.4 körülre áll le a meredek csökkenés.

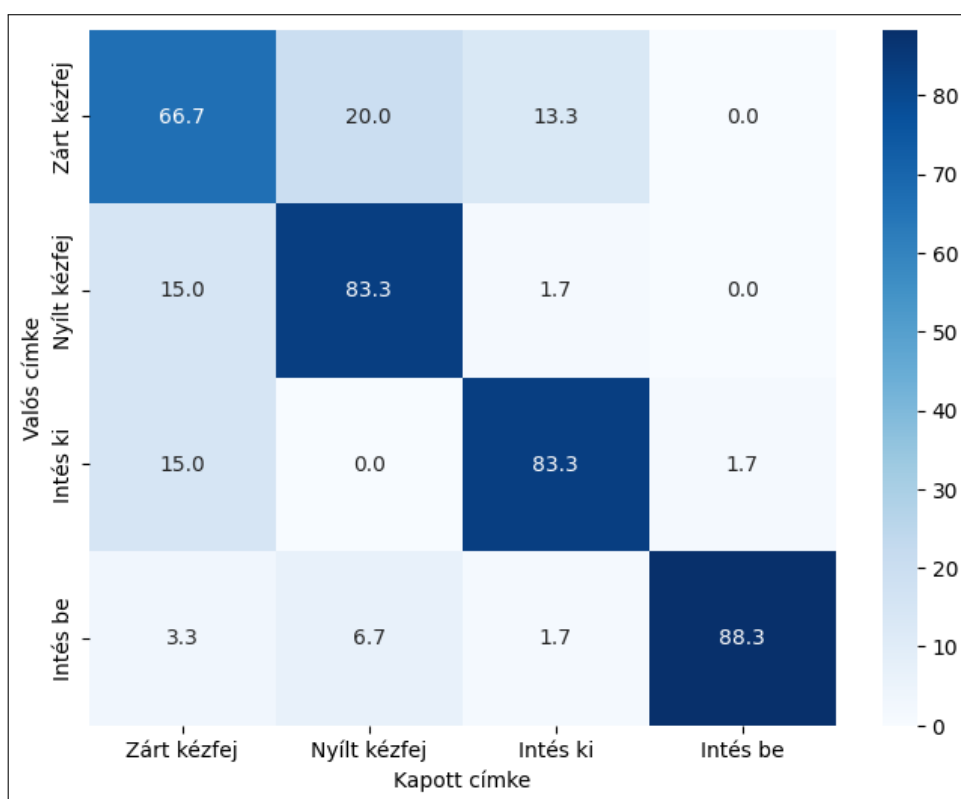
5.1. táblázat. online adathalmazon tanított modell eredményei az ismert adathalmazon és az ismeretlen adathalmazon.

	Accuracy (%)	Precision (%)	F1-score (%)
Ismert adathalmaz ismeretlen adat	82.92	84.50	83.18
Ismeretlen adathalmaz	80.42	80.98	80.59

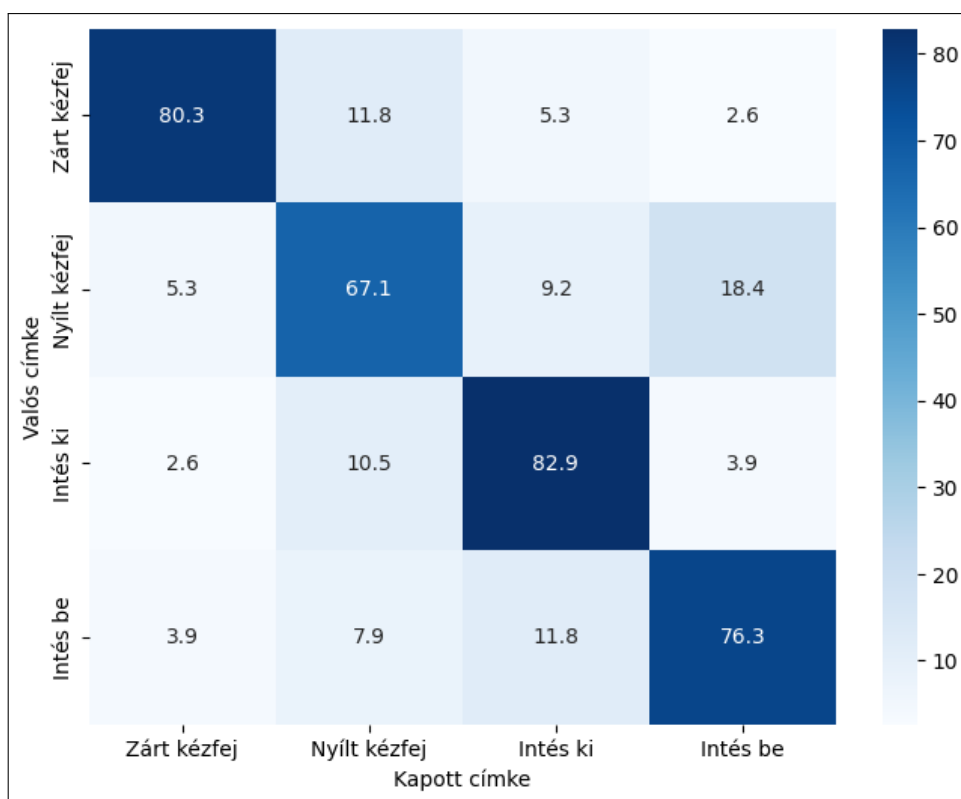


5.1. ábra. Költség és pontosság a Pangestu adathalmazon

A konfúziós mátrixok mutatják, hogy a modell minden osztályt hasonló sikerességgel talál el, és az ezt megelőző kutatásokhoz képest lenyűgöző egybehangzással teszi ezt. A kapott címke a modell az adatokon való használatával lett elkészítve, vagyis a kapott címke a predikció.



5.2. ábra. Konfúziós mátrix az eredményekhez a Pangestu adathalmazon (ismeretlen adat ismert személytől)



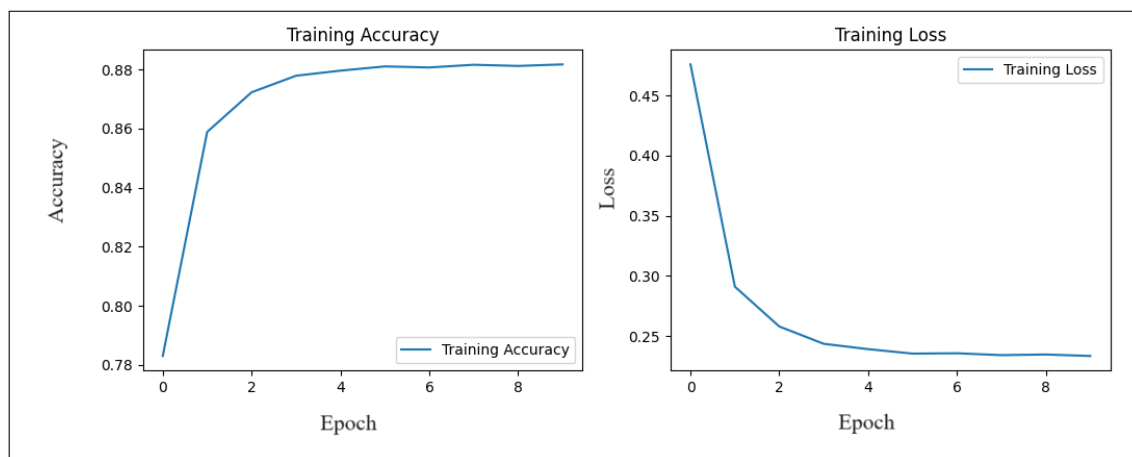
5.3. ábra. Konfúziós mátrix az eredményekhez az online adathalmazon

5.2. Pangestu adathalmaz

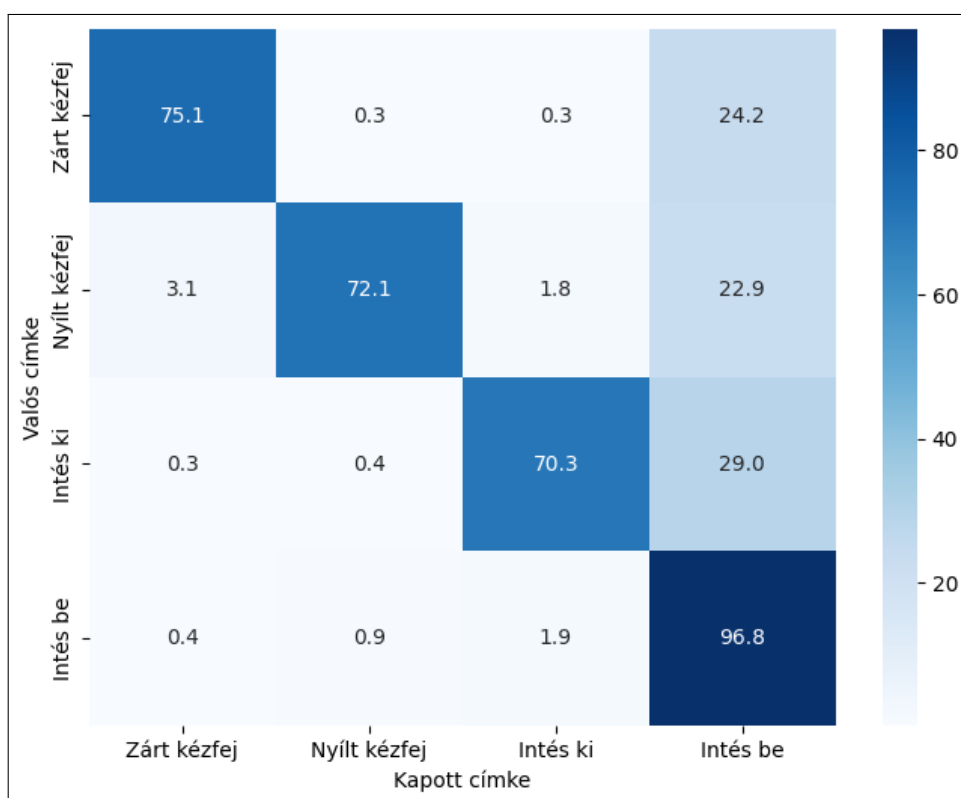
A Pangestu adathalmaz meg lett osztva két halmazra: 90% a tanításhoz, 10% a teszteléshez, és mindkettő előfeldolgozásra került. Az alábbi ábrák mutatják a költség és a pontosság alakulását a tanítás során. Az ismeretlen adathalmaz az online adathalmaz teljes előfeldolgozott állománya. Az elért eredmények alapján a modell igen ígéretes teljesítményt mutat mind az ismert, mind az ismeretlen adathalmazon. Az "accuracy": az ismert adathalmazon 84.05%, az ismeretlen adathalmazon pedig 78.94%. A pontosság az ismert adathalmazon 77.08%, az ismeretlen adathalmazon pedig 78.95% a precízió értéke. Ez azt jelzi, hogy a modell hatékonyan képes az osztályokba sorolni az adatokat. Az F1-score egy olyan metrika, amely az pontosság és a precízió súlyozott átlaga. Az ismert adathalmazon elért 78.28%, az ismeretlen adathalmazon elért 78.86% F1-score értékek azt sugallják, hogy a modell jól egyensúlyozza az pontosság és a precízió teljesítményét. Ezen eredmények alapján elmondható, hogy a modell megbízhatóan osztályozza az ismert és ismeretlen adatokat is, és alkalmas lehet a hasonló alkalmazásokban való felhasználásra. A 5.2 táblázat, valamint a 5.4, a 5.5 és a 5.6 ábrák együttesen festenek egy átfogó képet a modell teljesítményéről.

5.2. táblázat. Pangestu adathalmazon tanított modell eredményei az ismert adathalmazon és az ismeretlen adathalmazon

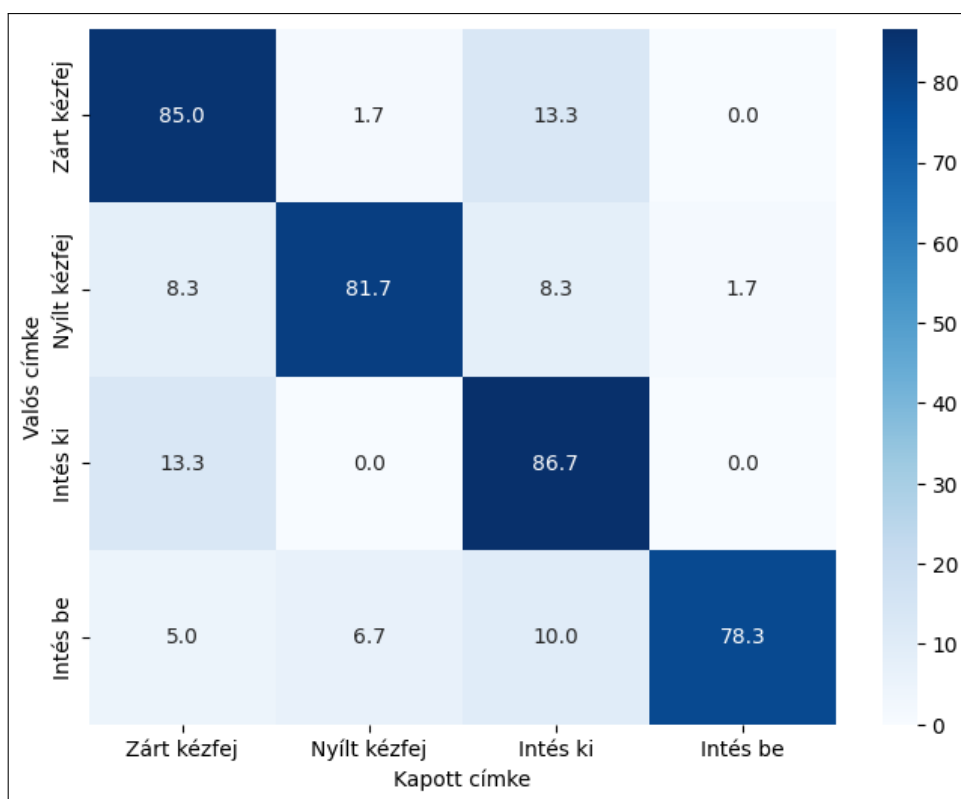
	Accuracy (%)	Precision (%)	F1-score (%)
Ismert adathalmaz ismeretlen adat	84.05	77.08	78.28
Ismeretlen adathalmaz	78.94	78.95	78.86



5.4. ábra. Költség és pontosság a Pangestu adathalmazon



5.5. ábra. Konfúziós mátrix az eredményekhez a Pangestu adathalmazon (ismeretlen adat ismert személytől)



5.6. ábra. Konfúziós mátrix az eredményekhez az online adathalmazon

5.3. Megnyújtott időablak

Az időablak kiterjesztése a GRU + teljesen összekötött rétegek architektúra jelentős előnyökkel jár, mivel növeli a modell képességét a hosszú távú függőségek megragadására, gazdagabb kontextuális információt nyújt, javítja a mintafelismerést, mélyebb jellemzőtanulást tesz lehetővé, alkalmazkodik a változó szekvenciahosszúságokhoz, hozzájárul a stabil tanításhoz, és végül pontosabb előrejelzésekhez vezet. Ezen előnyök és az olyan szempontok, mint a számítási komplexitás egyensúlyban tartása döntő fontosságú a hálózat teljesítményének optimalizálásához a különböző alkalmazásokban.

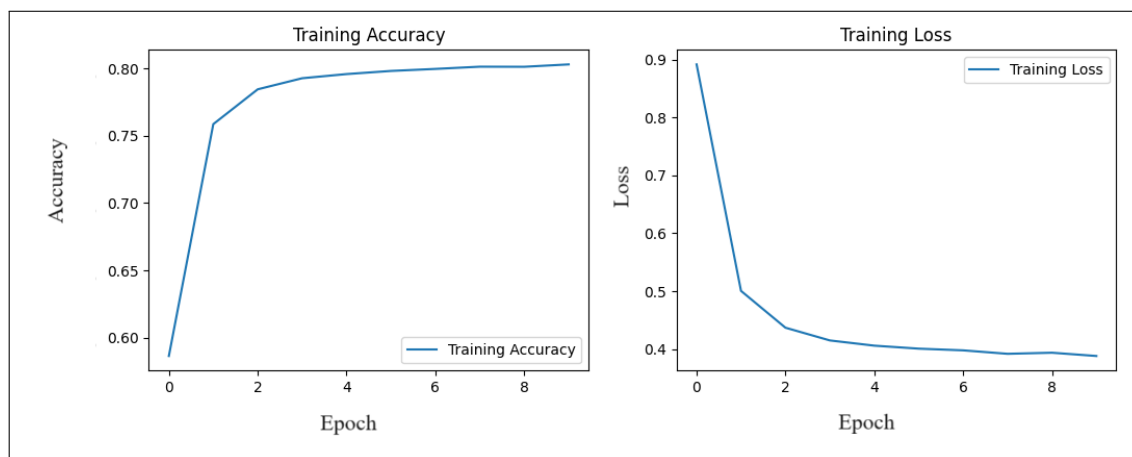
Ezen okoknál fogva a használt egy másodperces időablak mellett letesztelésre került ugyanaz a modell két másodperces időablakkal, hiszen nagyobb időegység alatt átfogóbb képet kaphat a mozdulatokról. A tesztelés menete ugyanaz.

5.3.1. Eredmények az online adathalmazon tanított modelltől

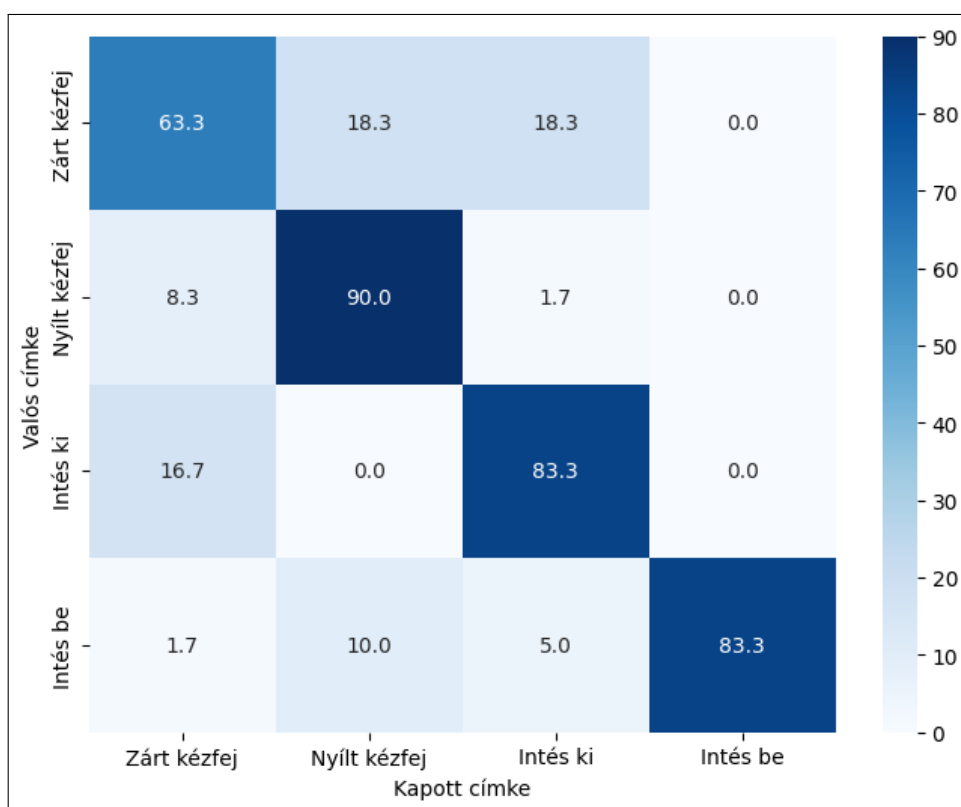
A 5.3 táblázat, valamint a 5.7, a 5.8 és a 5.9 ábrák együttesen festenek egy átfogó képet a modell teljesítményéről.

5.3. táblázat. Online adathalmazon tanított modell eredményei az ismert adathalmazon és az ismeretlen adathalmazon

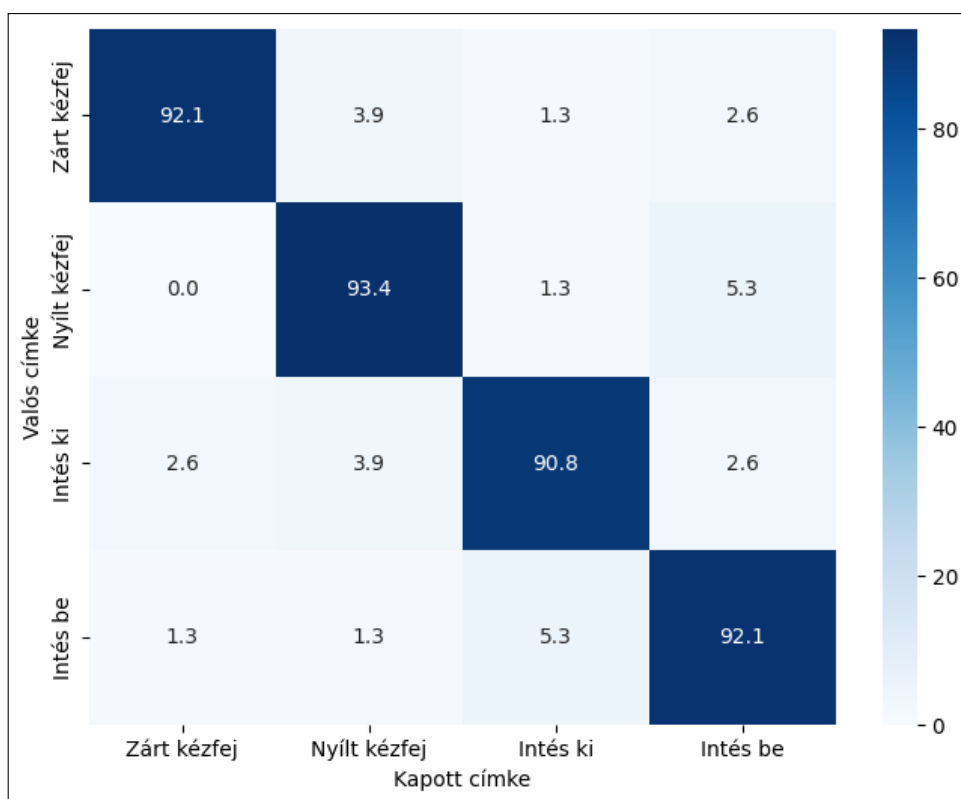
	Accuracy (%)	Precision (%)	F1-score (%)
Ismert adathalmaz ismeretlen adat	92.11	92.16	92.12
Ismeretlen adathalmaz	80.00	80.84	80.00



5.7. ábra. Költség és pontosság a Pangestu adathalmazon



5.8. ábra. Konfúziós mátrix az eredményekhez a Pangestu adathalmazon (ismeretlen adat ismert személytől)



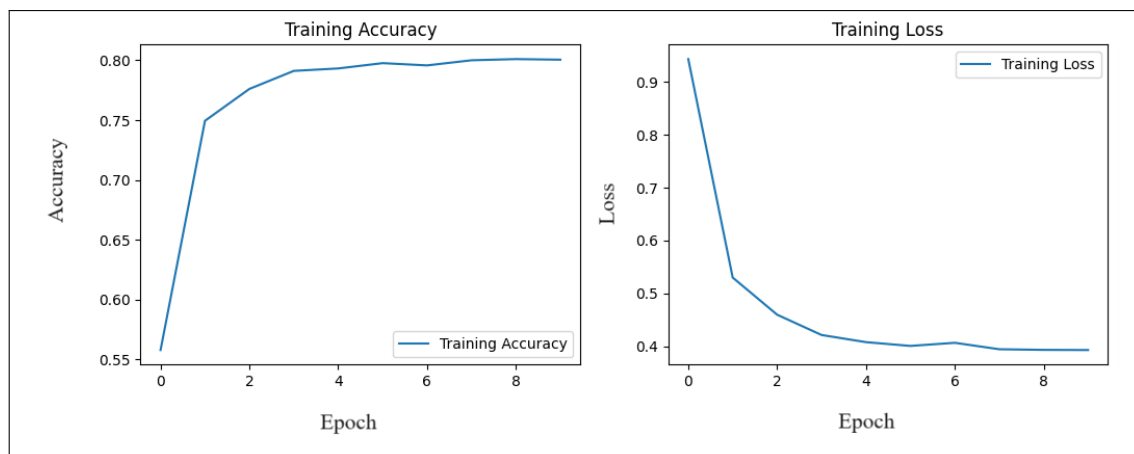
5.9. ábra. Konfúziós mátrix az eredményekhez az online adathalmazon

5.3.2. Eredmények a Pangestu adathalmazon tanított modellről

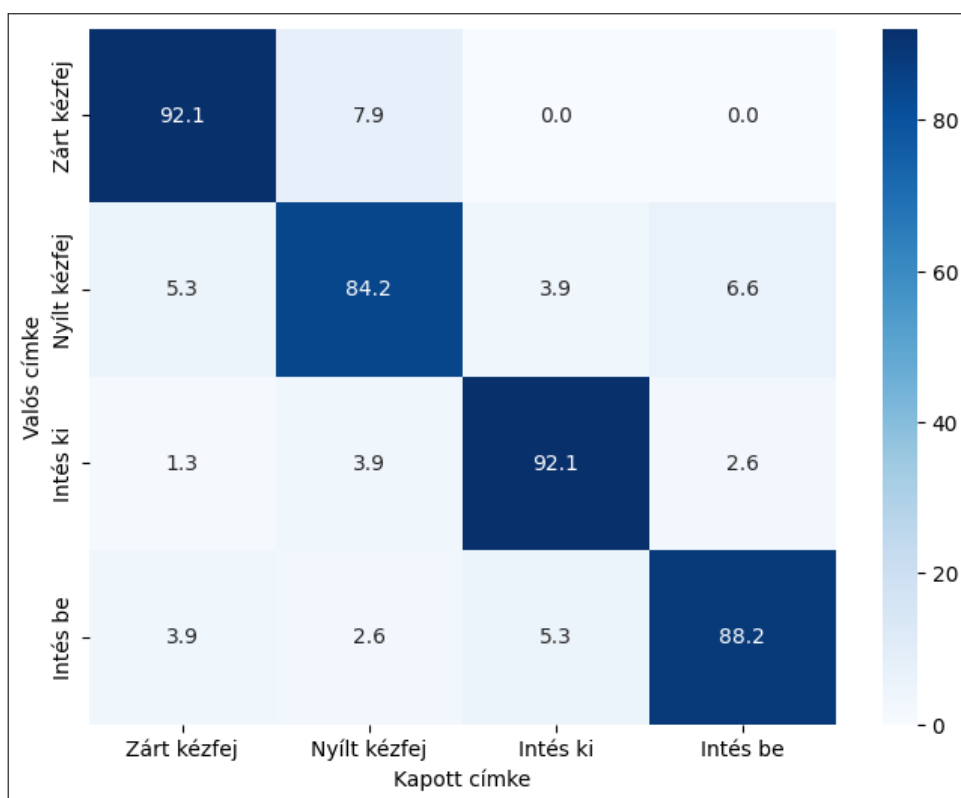
A 5.4 táblázat, valamint a 5.10, a 5.11 és a 5.12 ábrák együttesen festenek egy átfogó képet a modell teljesítményéről.

	Accuracy (%)	Precision (%)	F1-score (%)
Ismert adathalmaz ismeretlen adat	82.50	82.96	82.44
Ismeretlen adathalmaz	89.14	89.13	89.13

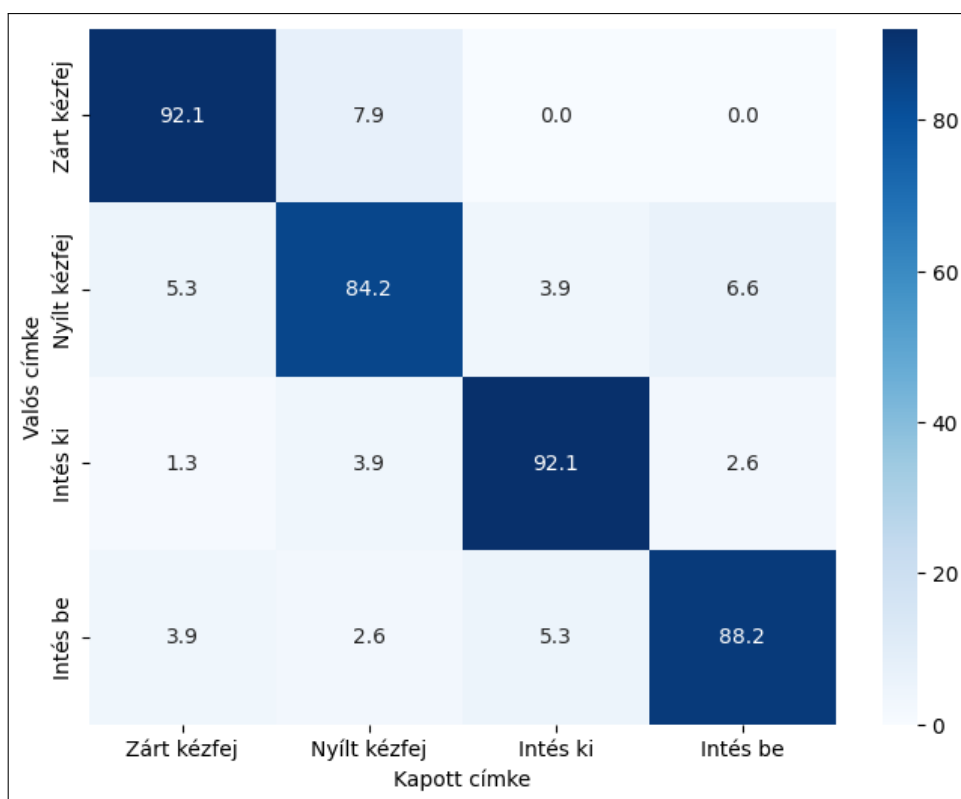
5.4. táblázat. Pangestu adathalmazon tanított modell eredményei az ismert adathalmazon és az ismeretlen adathalmazon



5.10. ábra. Költség és pontosság a Pangestu adathalmazon



5.11. ábra. Konfúziós mátrix az eredményekhez a Pangestu adathalmazon



5.12. ábra. Konfúziós mátrix az eredményekhez az online adathalmazon

5.4. Konklúzió

Az elődjét 12,29%-kal (1 másodperces időablakkal) felülmúló pontosságú új neurális háló bemutatása úttörő eredmény. Ez a jelentős javulás messzemenő hatással van az iparágakra, mivel csökkenti a hibákat az olyan kritikus alkalmazásokban, mint az egészségügy és a pénzügy. A megnövekedett pontosság megnyitja az utat a bonyolult adathalmazokban való navigálás, a finom minták feltárása és az AI-kutatás innovációjának előmozdítása előtt. Üzleti szempontból a 12,29%-os nyereség jelentős versenyelőnyt jelent, újradefiniálja a felhasználói élményt, és nagyobb bizalmat ébreszt az AI-megoldásokban. Ez az eredmény paradigmaváltást jelent, magasabb mércét állít a jövőbeli fejlesztések számára, és bemutatja a gépi tanulás folyamatos fejlesztésének átalakító erejét.

Az új modell általános pontosságának 12,29%-os javulása még figyelemre méltóbb, ha a konfúziós mátrixokat vizsgáljuk. Míg az átfogó pontossági mérőszám kezdetben egyértelműnek tűnik a kategóriák között, a közelebbi vizsgálat a modell pontosságának árnyalt finomodását tárja fel. Például egy orvosi diagnosztikai foratókönyvben, ahol az előző modell 70%-os pontosságot ért el egy adott állapotra vonatkozóan, az új modell felülmúlja annak pontosságát és növeli az általános pontosságot. A konfúziós mátrixok azonban

a téves pozitív és negatív eredmények számának jelentős csökkenését mutatják, ami a modell megnövekedett érzékenységét és specifikusságát bizonyítja.

Láthatóan jobban teljesít már ismert adathalmazból származó adatokra, de ilyen kis különbséget előállítani ebben a dolgozatban sikerül először a két kategória között, ami nagyszerű hír az olyan alkalmazások esetére, ahol minden pillanat fontos és az újratartás, illetve a kalibrációs mérések fontos időt rabolnak, vagy az alany nem képes az együttműködésre (pl.: Sürgősségi ellátás).

Az használt mércékre legjobb eredményt elérő eset a két másodperces időablakkal rendelkező online adathalmazon tanított modell, amely az elképesztő 92.11%-os pontossággal rendelkezik, és egy mozdulatra se megy ez az érték 90% alá. Így annak valószínűsége, hogy kétszer egymás után rosszul döntsön kevesebb, mint 1%.

Ez az árnyalt javulás különösen fontos az olyan alkalmazásokban, mint az autonóm rendszerek, a csalások felderítése és az orvosi diagnózisok, ahol a téves osztályozásnak mélyreható következményei vannak. A kategória-specifikus mérőszámok elemzésével mélyebb betekintést nyerhetünk a modell javított pontosságába, bemutatva, hogy a modell képes megbízhatóbb és megalapozottabb döntéseket hozni a különböző kategóriákon belül. Összefoglalva, a 12.29%-os javulás nemcsak számszerű előrelépést jelent, hanem jelentős lépést is a fokozott megbízhatóság és pontosság felé a kritikus forgatókönyvekben.

5.5. Kitekintés

A valós idejű teszteléshez a modell egy teleoperációs alkalmazásba való integrációját tűzöm ki célul, melyet Wijaya Feki Pangestu osztott meg a Github-on. Egy uArm Swift Pro robotkar mozgatását fel, le jobbra és balra az alkar mozgatásával valósítanám meg a Myo karpántba épített giroszkóp segítségével. A robot előre-hátra mozgását, illetve a fogás és elengedés gesztusát a modell segítségével implementálnám, négy megfelelő kézmozdulat klasszifikációjára támaszkodva. Ez megfelelően kihasználná a robot 4 szabadsági fokát illetve a kar végén elhelyezkedő pneumatikus endeffektort. Ez a Kísérlet rávilágíthat a használhatóságának hatáira, illetve előre nem látott problémákra. Mindenképpen fontosnak tartom, hogy a modell helytállását az egyes felhasználási területek, külön-külön vizsgálat bizonyítsa.

Köszönetnyilvánítás

Szeretném megragadni a lehetőséget, hogy köszönetet nyilvánítsak a témavezetőmnek, Dr. Nagy Balázsnak, minden segítségéért és az egyetemnek, hogy támogatta mind felszereléssel mind pedig anyagilag ezen dolgozat elkészültét. Illetve köszönet illeti még Wijaya Feki Pangestu-t, mert az ő munkája nélkül ez a kutatás nem tudta volna megfelelően bizonyítani eredményeit.

Irodalomjegyzék

- [1] Mingde Zheng, Michael S Crouch és Michael S Eggleston. „Surface electromyography as a natural human–machine interface: a review”. *IEEE Sensors Journal* 22.10 (2022), 9198–9214. old.
- [2] J. Weissmann és R. Salomon. „Gesture recognition for virtual reality applications using data gloves and neural networks”. *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*. 3. köt. 1999, 2043–2046 vol.3. DOI: 10.1109/IJCNN.1999.832699.
- [3] S Masson és tsai. „Integrating Myo armband for the control of myoelectric upper limb prosthesis”. *Proceedings of the XXV Congresso Brasileiro de Engenharia Biomédica*. 2016.
- [4] K R Mills. „The basics of electromyography”. *Journal of Neurology, Neurosurgery & Psychiatry* 76.suppl 2 (2005), ii32–ii35. old. ISSN: 0022-3050. DOI: 10.1136/jnnp.2005.069211. eprint: https://jnnp.bmj.com/content/76/suppl_2/ii32.full.pdf. URL: https://jnnp.bmj.com/content/76/suppl_2/ii32.
- [5] Cengiz Tepe és Mehmet Demir. „Detection and Classification of Muscle Activation in EMG Data Acquired by Myo Armband”. *European Journal of Science and Technology* (2020. aug.), 178–183. old. DOI: 10.31590/ejosat.779660.
- [6] Luca Mesin. „Crosstalk in surface electromyogram: literature review and some insights”. *Physical and Engineering Sciences in Medicine* 43 (2020. máj.). DOI: 10.1007/s13246-020-00868-1.
- [7] Hritam Basak és tsai. *SVM and ANN based Classification of EMG signals by using PCA and LDA*. 2021. arXiv: 2110.15279 [eess.SP].
- [8] Mamun Bin Ibne Reaz, M Sazzad Hussain és Faisal Mohd-Yasin. „Techniques of EMG signal analysis: detection, processing, classification and applications”. *Biological procedures online* 8 (2006), 11–35. old.

- [9] Reza Bagherian Azhiri, Mohammad Esmaeili és Mehrdad Nourani. *Real-Time EMG Signal Classification via Recurrent Neural Networks*. 2021. arXiv: 2109 . 05674 [eess.SP].
- [10] Jose Guillermo Colli Alfaro és Ana Luisa Trejos. „User-Independent Hand Gesture Recognition Classification Models Using Sensor Fusion”. *Sensors* 22.4 (2022). ISSN: 1424-8220. DOI: 10.3390/s22041321. URL: <https://www.mdpi.com/1424-8220/22/4/1321>.
- [11] Christos Dolopikos és tsai. „Electromyography signal-based gesture recognition for human-machine interaction in real-time through model calibration”. *Advances in Information and Communication: Proceedings of the 2021 Future of Information and Communication Conference (FICC), Volume 2*. Springer. 2021, 898–914. old.
- [12] Paolo Visconti és tsai. „Technical Features and Functionalities of Myo Armband: An Overview on Related Literature and Advanced Applications of Myoelectric Armbands Mainly Focused on Arm Prostheses”. *International Journal on Smart Sensing and Intelligent Systems* 11 (2018. jún.), 1–25. old. DOI: 10.21307/ijssis-2018-005.
- [13] Frank Rosenblatt és tsai. *Principles of neurodynamics: Perceptrons and the theory of brain mechanisms*. 55. köt. Spartan books Washington, DC, 1962.
- [14] Kyunghyun Cho és tsai. „Learning phrase representations using RNN encoder-decoder for statistical machine translation”. *arXiv preprint arXiv:1406.1078* (2014).
- [15] Jing Chen és tsai. „Low-Cost and Device-Free Human Activity Recognition Based on Hierarchical Learning Model”. *Sensors* 21 (2021. márc.), 2359. old. DOI: 10.3390/s21072359.
- [16] Bo Pang, Erik Nijkamp és Ying Nian Wu. „Deep learning with tensorflow: A review”. *Journal of Educational and Behavioral Statistics* 45.2 (2020), 227–248. old.
- [17] Nuradebah Burhan, Rozaimi Ghazali és tsai. „Feature extraction of surface electromyography (sEMG) and signal processing technique in wavelet transform: A review”. *2016 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*. IEEE. 2016, 141–146. old.

- [18] Sachin Sharma és tsai. „Techniques for feature extraction from EMG signal”. *International Journal of Advanced Research in Computer Science and Software Engineering* 2.1 (2012).
- [19] Nienke Willigenburg és tsai. „Removing ECG contamination from EMG recordings: A comparison of ICA-based and other filtering procedures”. *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology* 22 (2012. jan.), 485–93. old. DOI: 10.1016/j.jelekin.2012.01.001.

Ábrák jegyzéke

3.1. A Myo karpánt és tulajdonságai.	9
3.2. A kézmozdulatok (A) ökölbe szorított kézfej, (B) ujjak széttárása, (C) intés kifelé, (D) intés befelé	10
3.3. Gated Recurrent Unit (GRU) architektúra [15]	13
4.1. Különbség a két adathalmaz között, a felső ábra mutatja az online adathalmazt, az alsó pedig a Wijaya Feki Pangestu által gyűjtöttet.	17
4.2. A szekvenciálisan összekötött model ábrázolása.	21
5.1. Költség és pontosság a Pangestu adathalmazon	25
5.2. Konfúziós mátrix az eredményekhez a Pangestu adathalmazon (ismeretlen adat ismert személytől)	25
5.3. Konfúziós mátrix az eredményekhez az online adathalmazon	26
5.4. Költség és pontosság a Pangestu adathalmazon	27
5.5. Konfúziós mátrix az eredményekhez a Pangestu adathalmazon (ismeretlen adat ismert személytől)	27
5.6. Konfúziós mátrix az eredményekhez az online adathalmazon	28
5.7. Költség és pontosság a Pangestu adathalmazon	29
5.8. Konfúziós mátrix az eredményekhez a Pangestu adathalmazon (ismeretlen adat ismert személytől)	29
5.9. Konfúziós mátrix az eredményekhez az online adathalmazon	30
5.10. Költség és pontosság a Pangestu adathalmazon	31
5.11. Konfúziós mátrix az eredményekhez a Pangestu adathalmazon	31
5.12. Konfúziós mátrix az eredményekhez az online adathalmazon	32

Táblázatok jegyzéke

3.1. Dolopikos és tsai. Eredmények	8
3.2. Dolopikos és tsai. konfúziós mátrix	8
4.1. A modell hiperparaméterei	22
5.1. online adathalmazon tanított modell eredményei az ismert adathalmazon és az ismeretlen atadhalmazon	24
5.2. Pangestu adathalmazon tanított modell eredményei az ismert adathalma- zon és az ismeretlen adathalmazon	27
5.3. Online adathalmazon tanított modell eredményei az ismert adathalmazon és az ismeretlen atadhalmazon	29
5.4. Pangestu adathalmazon tanított modell eredményei az ismert adathalma- zon és az ismeretlen atadhalmazon	30