



Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar

Torma Szabolcs

EEG JELEK FELDOLGOZÁSA DEEP LEARNING SEGÍTSÉGÉVEL

KONZULENS

Dr. Szegletes Luca

BUDAPEST, 2019

Tartalomjegyzék

1 Bevezetés.....	5
2 EEG jelek, rendszerek és agyhullámok.....	6
2.1 Agyi aktivitás.....	6
2.1.1 Jelek keletkezése.....	6
2.1.2 Alap hullámtípusok.....	7
2.2 EEG rendszerek felépítése és alkalmazásuk.....	8
2.2.1 Elektródák.....	8
2.2.2 Lehetséges problémák a jelrögzítés során.....	9
2.2.3 Jelfeldolgozás.....	10
2.3 Evoked Potential (EP).....	11
2.3.1 Event-Related Potential (ERP).....	11
2.3.2 Egy rejtélyes, de lényeges ERP, a P300.....	12
3 Brain-Computer Interface rendszerek.....	18
3.1 BCI rendszerek osztályozása és felépítése.....	18
3.1.1 Rendszerek osztályozása.....	19
3.1.2 Rendszerek felépítése.....	20
3.2 Rendszerek alkalmazhatósága.....	20
3.3 Létező rendszerek bemutatása.....	22
3.3.1 P300 speller.....	22
3.3.2 Exoskeleton irányítása vezeték nélküli ECoG alapú BCI rendszerrel.....	26
4 Single-Trial klasszifikáció Deep Learning segítségével.....	29
4.1 Konvolúciós neurális hálók.....	30
4.1.1 Rövid elmélet a konvolúciós neurális hálók működéséről.....	30
4.1.2 Aktivációs függvények.....	31
4.1.3 Gradiens-alapú tanulás.....	32
4.2 P300 detekció 1D CNN-el.....	32
4.2.1 Adathalmaz.....	33
4.2.2 Architektúra.....	34
4.2.3 Hyperparameter tuning.....	36
4.2.4 Eredmények.....	37

5 Konklúzió és jövőbeli tervek.....	39
--	-----------

Abstract

Electroencephalographic (EEG) signals, which are recorded through the skull, are widely used in medicine to make various diagnoses. However, beyond diagnosis, the use of EEG signals has provided a means for Brain-Computer Interface (BCI) systems which with a user can communicate without any speech or gesture. These tools include the speller created by Farwell and Donchin (1988), which allows individuals with closed syndrome to communicate independently with the outside world. The difficulty encountered in processing EEG signals is that they are loaded by noise, so averaging the results of multiple generated impulses is necessary to allow the user to transmit information with sufficient accuracy, which, however, limits the speed of the process.

The use of neural networks has provided innovative and effective solutions in many areas (eg image processing, chat bots). This technology is also frequently used in medicine because of its ability to solve high complexity tasks. Most commonly, the appearance of lesions and diseases is searched in magnetic resonance imaging (MRI), computer tomography (CT) images using two-dimensional convolutional networks. The use of one-dimensional convolutional networks is not yet as widespread as two-dimensional ones.

In this paper, I investigate the state-of-the-art EEG and P300 based BCI systems, and attempt to classify the P300 signal using a convolutional neural network. My goal is to build a one-dimensional convolutional neural network architecture with respect to the type of data that is capable of processing cerebral signals with sufficient accuracy, and is not common in the literature. For teaching and testing of the model, I use the BCI Competition 3 2 dataset, which was recorded using a tool similar to the speller created by Farwell and Donchin.

1 Bevezetés

A koponyán keresztül rögzített elektroencefalográfiai (EEG) jeleket az orvostudomány széleskörűen alkalmazza különböző diagnózisok felállítására. Azonban diagnózisok meghatározásán túl az EEG jelek felhasználásával olyan Brain-Computer Interface (BCI) rendszerek közé tartozó eszközök kerültek megvalósításra, amelyek beszéd és gesztusok nélküli kommunikációra adnak lehetőséget. Ezek közé az eszközök közé tartozik a Farwell és Donchin (1988) által létrehozott betűző is, amely segítségével bezárt-szindrómás személyek is képesek önállóan kommunikálni a külvilággal. Az EEG jelek feldolgozása során felmerülő nehézség, hogy a jelek zajosak, így több generált impulzus mérési eredményének átlagolása szükséges ahhoz, hogy a felhasználó megfelelő pontossággal tudjon információt közvetíteni, amely azonban a folyamat sebességét limitálja.

Neurális hálók alkalmazása számos területen (pl. képfeldolgozás, chatbotok) hozott újszerű és hatékony megoldásokat. Ennek a technológiának a használatára az orvostudományban is gyakran sor kerül, mivel nagy komplexitású feladatok megoldására képes. Legelterjedtebben mágnesesrezonancia-képkeltő eszköz (MRI), illetve számítógépes tomográfia (CT) által létrehozott képeken megjelenő elváltozások vagy betegségek elemzését végzik kétdimenziós konvolúciós hálók segítségével. Az egydimenziós konvolúciós hálók használata egyelőre nem olyan mértékben elterjedt, mint a kétdimenziósaké.

Jelen dolgozatomban azt kutatom fel, hogy milyen state-of-the-art EEG, illetve P300 alapú BCI rendszerek léteznek, továbbá kísérletet teszek P300 jelklasszifikációra konvolúciós neurális háló segítségével. Célom, hogy az adat típusára való tekintettel olyan egydimenziós konvolúciós neurális háló architektúrát építsek, amely megfelelő pontossággal képes az agyi jelek feldolgozására és a szakirodalomban sem fordul még elő gyakori mértékben. A háló tanításához és teszteléséhez a BCI Competition 3 2-es adathalmazt használom, amelyet Farwell és Donchin által létrehozott betűzőhöz hasonló eszköz alkalmazásával vettek fel.

2 EEG jelek, rendszerek és agyhullámok

Az agyi neuronok aktiválódása során elektromos potenciálok keletkeznek. [6] Az agyi kéregben keletkező jelek és hullámok mérésére legelterjedtebben használt technológiák között szerepel az elektroencefalográfia (EEG). Az EEG alkalmazása lehetőséget ad különböző diagnózisok felállítására, kutatásokra az emberi agy jobb megértésére, továbbá Brain-Computer Interface (BCI) rendszerek fejlesztésére, amelyek segítségével mozgásukban vagy kommunikációjukban korlátozott emberek életminősége javítható. Az olyan képalkotó rendszerek, mint a számítógépes tomográfia (CT) vagy a funkcionális mágneses rezonanciavizsgálat (fMRI) nagyobb térbeli felbontásban képes feltérképezni az agyat, mint az EEG, azonban az EEG milliszekundumos nagyságrendben tud információt szolgáltatni az agyban folyó folyamatokról. [7] Ritmusok - pl. a μ vagy a beta -, illetve kiváltott potenciál típusok – evoked potential (EP) – idő- és frekvenciabeli karakterisztikái mind jól megfigyelhetők EEG használata során. További előnye az EEG alkalmazásának, hogy egyszerű felépítésű, olcsóbb, mint a fentebb említett technológiák és kevésbé korlátozza a felhasználó mozgását. [7][9]

2.1 Agyi aktivitás

Az agyban keletkező elektromos áramokat Richard Caton fel fedezte fel 1875-ben. Később, 1924-ben Hans Berger publikálta megfigyeléseit arra vonatkozóan, hogy a különböző emberi tevékenységek – pl. alvás, mozgás -, érzelmi állapotok, illetve betegségek különböző agyi aktivitást generáltak. Kísérleteit a jelek felerősítésével végezte el, amelyhez általános rádiókészüléket használt. Adrian és Matthews 1934-ben megerősítették az agyhullámok létezésének elméletét, továbbá először figyelték meg az egyik – közhiedelemben is – legtöbbször emlegetett alfa hullám jelenlétét. [6]

2.1.1 Jelek keletkezése

Az emberi agy fő részegységeit a kisagy, a nagyagy, diencephalon és az agytörzs alkotja. A nagyagyat négy lebeny jobb- és baloldali párosa alkotja, amelyek sorra a frontális, parietális, nyakszirti, és temporális lebenyek. A nagyagy jobb és bal agyféltekéjét a sűrűn kanyargós agykéreg borítja be, amelyet körülbelül 100 milliárd

idegsejt, másképpen neuron együttes hálóját alkot. Az agykéreg a központi idegrendszer egyik legjelentősebb egysége. Az EEG mérések során a legtöbb rögzített jel az agykéregből származik, mivel az agy ezen része helyezkedik el a koponya felszínéhez a legközelebb. Elektromos jelek a neuron aktiválódásakor keletkeznek. Az aktiválódás során a neuron aktivációs potenciált hoz létre, azaz egy impulzust, amely az axonon és a szinapszisukon keresztül jut el a másik sejtesthez. Az axon és a szinapszisuk kapcsolják össze a sejtesteket. Amíg az egy agysejthez tartozó szinapszisuk száma folyamatosan nő, addig a neuron száma csökken az ember életkorának növekedésével. Az EEG-vel mérhető jelek értéke általános esetben – pl. nincs epilepsziás rohama a páciensnek - kb. 0.5-100 mikrovoltos tartományban változik és szinuszos jelalakokkal rendelkezik. [6][7][9]

2.1.2 Alap hullámtípusok

A jelek frekvenciája és amplitúdója alapján öt alap agyhullámtípus megkülönböztetése lehetséges. Az eltérő hullámtípusok az agy különböző állapotában, annak különböző területein mérhetők (az egyes területek eltérő típusú hullámokat bocsátanak ki). Az öt alap hullámfajta a gamma, beta, alpha, delta és theta.

A gamma hullámot 30Hz-nél nagyobb frekvencia jellemzi, az öt hullámtípus közül ez rendelkezik a legkisebb amplitúdóval. Ébrenléti állapotban, főképpen éber figyelem és motoros funkciók használata során mérhető. Olyan szabályos ritmikusság lelhető fel a hullám jelalakjában, amely például az alpha vagy a theta esetében nem.

A beta hullám frekvenciatartománya 13-tól 30 Hz-ig terjed, emellett a gamma hullámhoz hasonlóan alacsony amplitúdóval és szabálytalan jelalakokkal rendelkezik. A beta hullám a frontális és a temporális lebeny területén mérhető, általános ébrenléti állapotban.

Az alpha hullámtípus 8-13 Hz között szabályosan és ritmikusan oszcillál, alacsony amplitúdóval. A hullám jelenléte a hátsó, parietális, illetve a nyakszirti lebenyek területén mérhető a legjelentősebben. Ezekben a területeken nyugodt állapotban jelenik meg a hullám, egyes egyénekénél már a szem becsukása is elég az előhíváshoz.

A theta hullám - amelynek frekvenciája 4-8 Hz-ig terjedhet - a másik négy hullámhoz képest közepesen nagy amplitúdóval rendelkezik. Az említett hullám normál alvási állapotban, a nyakszirti és a parietális lebenyek régiójában lelhető fel.

A delta hullám az alvás nonrapid eye movement (NREM) fázisában jelenik meg: minél mélyebb az alvás, annál nagyobb mértékben figyelhető meg a jelenléte. Ebben az állapotban a hullám nagy amplitúdóval, 0.5-4 Hz között oszcillál. [6][7][9]

Az agyban keletkező hullámok és jelek vizsgálata során a mért információt az agy adott területéhez és a vizsgált egyénhez mérten értékelik ki. Minden személy egyedi hullám- és jelkarakterisztikával rendelkezik. A személyiség, a mentális vagy a fizikai állapot mind a hullámokat befolyásoló tényezők közé tartoznak, így a jelek megfigyelésével fizikai és pszichológiai betegségek egyaránt diagnosztizálhatók. Például az agykéreg által generált, a normál frekvenciánál alacsonyabban oszcilláló alpha hullám tumorra vagy trombózisra, míg egy általánosan magasabb feszültségű csúccsal, azaz tüskével vagy ugrással rendelkező jel epilepsziára utalhat. [6][9]

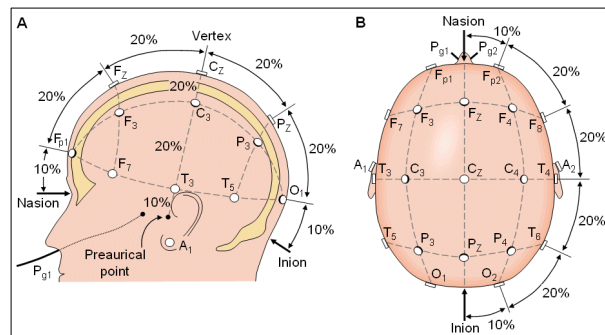
2.2 EEG rendszerek felépítése és alkalmazásuk

Az agyhullámok és egyéb, agyban keletkező jelek felvételére gyakran alkalmazott technológia az elektroencefalográfia (EEG). A technológia segítségével a fej felszínén keresztül [ms]-os időbeli felbontásban lehetséges a többnyire agykéregbeli potenciálváltozások rögzítése. A fejbe történő behatolás alapján háromféle EEG eljárás különböztethető meg: invazív, félig invazív és non-invazív. Az agykéreg elektrophysiológiai jeleit vezető elektródák érzékelik, elektromos jellé alakítják és továbbítják az erősítő egységnek. Az erősítés után a jelek digitalizálásra kerülnek, hogy a jelrögzítési folyamat végén elhelyezkedő számítógép el tudja végezni a szükséges jelfeldolgozást. [6]

2.2.1 Elektródák

Az EEG három eljárás típusa alapján az elektródák elhelyezése történhet a fej felszínén (non-invazív), a koponyába, de nem az agyba (félig invazív) vagy a koponya alá (invazív). Leggyakrabban a non-invazív módszer kerül használatra, mivel megfelelő anyagú és minőségű elektródákkal jó eredménnyel lehet mérést végezni, továbbá sokkal kisebb korlátozottságot jelent a vizsgált alany számára. A non-invazív eljárás az alany számára teljes mértékben fájdalommentes, az elektródák többnyire sapkába fűzve kerülnek a fejre. Egyes esetekben vezető anyag felkenésével biztosítják, illetve javítják az elektróda és a fejbőr közti kontaktust. Az elektródák lokalizációit az International Federation in Electroencephalography and Clinical Neurophysiology intézet 1958-ban

általánosított 10-20-as – napjainkban néhol már 10-10-es – szabványelrendezés határozza meg. Az elrendezés négy referenciapontjával rendre a nasion, a fültőmirigyi pontok és az inion szolgálnak. A 10-20-as elrendezés százalékban kifejezett résztávolságokat jelent a fülek és az orr között. Az elektródák az agy adott régióinak megfelelően kerülnek elnevezésre. Minden elektróda egy vagy két betűjelet és egy számot, a középvonali pozíció esetén a szám helyett „z” betűt kap. A betű címke jelzi a lebenyi területet, a számok pedig paritásuk szerint kerülnek kiosztásra; a jobb agyféltekén lévő elektródák kapják a páros számokat, míg a baloldalin lévők a páratlanokat (a számozás lehet fordítva is, és fontos az, hogy az egyik agyféltekén csak páros, míg a másikon csak páratlan számok szerepeljenek). A használt betűk F, C, T, P, O, jelentésük pedig rendre frontális, középső, temporális, posteriori és nyakszirti lebenyi pont. Az elrendezésben meghatározott pozíciók mindegyike az agy valamely fontos funkciót ellátó területéhez áll közel. Példák a területekre: Fz-hez közel az elhatározásért és motivációért, a C3, C4 és Cz címkék által jelölt területek az érzékelésért és motoros funkciókért, P3, P4 és Pz által meghatározott régióban pedig az észleléssel és megkülönböztetéssel kapcsolatos funkcióért felelős pontok helyezkednek el. [6][9]



Abra 1: Elektródák elhelyezkedése 10-20-as szabványrendszer szerint [20]

Jelrögzítő elektródák mellett referencia elektródák használata is szükséges lehet. Referencia elektródának alkalmasak egyes jelrögzítő elektródák is (pl.: Cz), de gyakran elkülönülten kerülnek fel valamely fejen lévő testrésze. A választott testrész lehet például az orrhegy, a fül vagy az ellenoldali-fül. [6]

2.2.2 Lehetséges problémák a jelrögzítés során

Habár az EEG több más technológiához képest – pl.: fMRI, MRI vagy CT – kevésbé komplex, a jelrögzítés és jelfeldolgozás koránt sem akadálymentes feladat. A

mérések során gyakori problémát okoznak a jelben megjelenő különböző melléktermékek. Ezek a melléktermékek lehetnek például a jelet felvevő és feldolgozó rendszerben keletkező zajok, zavarok, továbbá felmerülhetnek olyan alanyfüggő problémák is, mint például a pislogás, önkéntelen megmozdulás vagy a pace-maker léte.

A melléktermékek forrása a használt berendezés egységei és/vagy a páciens lehet. A páciensről származó zavarok (a pislogás vagy egy testrész megmozdulása) kiküszöbölését a zavarokat okozó tényezők megfigyelésével teszik. Gyakran alkalmazott megoldás, hogy elektro-okulográfiával (EOG) rögzítésre kerülnek a szem, és annak környékéről származó jelek, amelyek felhasználásával később kiszűrhetők azok a szekvenciák, amelyek rögzítése során a zavar bekövetkezett. Technikai zavarok ellen is jól lehet védekezni. Az elektródákon nem javasolt, hogy 5 kOhm-nál nagyobb impedancia essen, különben disztorziók jelenhetnek meg a rögzített jelekben. Az 50/60 Hz-es hálózat okozta zaj ellen az elektródák kábeleinek rövidítésével lehet védekezni. További technikai hiba lehet a kábel mozgatása, a kontaktus hiba vagy az impedancia ingadozás. [6]

2.2.3 Jelfeldolgozás

Az elektródák által rögzített jelek először erősítésen, majd jelszűrésen és végül digitalizáláson esnek át. Többnyire a jel ezután újabb szűrésen esik át, ekkor már FIR vagy IIR szűrők alkalmazásával. A digitalizált jeltől a feldolgozó számítógép valamilyen algoritmus alkalmazásával szelektálja a releváns komponenseket és az ún. feature-öket. A principal component analysis (PCA), a independent component analysis (ICA), a power spectrum analysis (PSA) azok közé az algoritmusok közé tartoznak, amelyek a jel komponenseinek szétválasztásával kiszűrik az esetlegesen jelben maradt melléktermékeket, illetve felismerik az olyan keresett jelkomponenseket, mint például a P300. A rögzített jelek analízise időben és/vagy frekvencia spektrumban történik. Az analízis során készíthetők olyan ábrák, amelyek segítséget tudnak nyújtani a jelek agyban elhelyezkedő forrásának megtalálásában vagy a jelkomponens szemmel való felismerésében. [6][9]

2.3 Evoked Potential (EP)

A neuronok belső vagy külső ingerrel történő stimulációjának hatására keletkező feszültségváltozás a kiváltott potenciál, angolul evoked potential (EP). Az inger lehet külső: a periférikus érzékszervek stimulálása vizuálisan vagy hangilag; illetve belső: epilepsziás roham. Az EP szinonimájaként az irodalomban gyakran jelenik meg az esemény-releváns potenciál, azaz event-related potential (ERP) elnevezés is: mérések és vizsgálatok esetén a stimulust - melyre a jel kiváltódik - egy előre meghatározott esemény (event) váltja ki. [14][16]

2.3.1 Event-Related Potential (ERP)

Az esemény releváns potenciálok (ERP) általában 1 és 100 mikrovolt körüli jelek – kivételek között van például az epilepszia által ingerelt ERP, amelynek csúcscértéke elérheti az 1000 mikrovoltot is. Időtartománybeli terjedelemben pár 10 ms-tól, több 100 ms-ig is tarthat a hosszuk. Az ERP jelek többnyire kis amplitúdóval rendelkeznek, ezért gyakran mérésükhöz több jeltörzítési iteráció – szakirodalomban epoch – szükséges a jel komponenseinek kinyeréséhez. Az epoch-ok során az ERP-t kiváltó inger és a kiváltott jel rögzítése ismétlésre kerül, majd az így kapott eredmények átlagolásából áll elő az kívánt eredmény. Az átlagolás során kiesnek az olyan komponensek, amelyek eltakarják az ERP jelet (pl.: zaj). Az EEG mellett mágnesesrezonancia-képzéssel (MR) és számítógépes-tomográfiával (CT) is rögzíthetők ERP-k. Az EEG-vel a folyamat időbeli tartamáról és a kiváltott jel alakjáról, míg a másik kettővel a kiváltott jel térbeli lokalizációjáról lehet részletes információt szerezni. [13][14]

Az agyban történt elváltozásokra, betegségekre vagy sérülésekre vonatkozóan diagnózisok állíthatók fel az ERP-ök figyelésével és rögzítésével. Agyműtétek során ellenőrzésképpen használják fel a jeleket, ugyanis az operáció közben az orvosok stimulálják és figyelik az érintett területek épségét, illetve megfelelő működését. Napjainkban az emberi memóriával, a figyelemmel és az információ-feldolgozással, továbbá a mentális rendellenességekkel foglalkozó kutatások hasznosítják az olyan komponensekkel rendelkező ERP jeleket, mint a P300, N200 vagy a P50.

Megjegyzés: a kiváltott potenciálok (EP) a szakirodalomban gyakran kapnak egy „előtagot”, amely azt a területet jelzi, amely ingerlésre, ezáltal pedig vizsgálatra kerül.

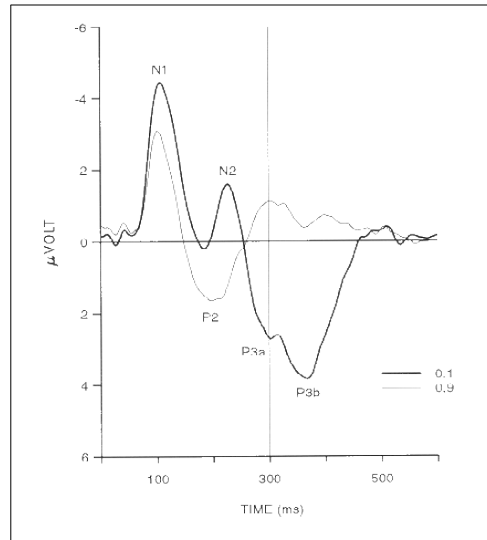
Ilyen EP-ök pl. a Visual Evoked Potential (VEP) – pl.: látókéreg területének vizsgálata - vagy a Somatosensory Evoked Potential (SSEP).[9]

2.3.2 Egy rejtélyes, de lényeges ERP, a P300

Az Event-Related Potential (ERD) jelek csoportjába tartozik a P300 a jel. Habár a jel jelenlétére már az 1960-as években felfigyeltek, a mai napig vannak megválaszolatlan kérdések vele kapcsolatban. Olyan kérdésekben nincs még egyetértés, mint hogy hol található a jel forrása az agyon belül, vagy hogy pontosan milyen folyamatnak a része és milyen célt szolgál. Ezen tények ellenére olyan P300 alapú Brain-Computer Interface (BCI) rendszerek fejlesztése vált lehetővé, amelyek a szórakoztató ipartól kezdve egészen a különböző orvosi alkalmazási területekig hasznosíthatók. Orvosi területen például kommunikációjukban vagy mozgásukban korlátozott embereknek jelentenek segítséget, a szórakoztató iparban videójátékokhoz és kiterjesztett valóság fejlesztéséhez kerül alkalmazásra. [7][13]

Event-Related Potential (ERP)-nek hívjuk azokat a jeleket, amelyek a neuronok külső vagy belső stimulusra való aktiválódása során keletkeznek [6][11]. A pozitív potenciálváltozással járó P300 ERP azt az információt hordozza magában, hogy az egyén egy feladatspecifikus, számára fontos vizuális vagy hangbeli ingerre észlelt és ismert fel. A jel nevében szereplő 300 az inger és a jel csúcsa között eltelt időt, azaz késleltetést jelzi milliszekundumban. Egyes irodalmakban P3 hullámként is említik, amely elnevezés onnan ered, hogy ez a harmadik fő pozitív csúccsal rendelkező komponense a kései szenzoros kiváltott potenciálnak. [14] A jelet két fő komponens alkotja: P3a és a P3b, utóbbi tekintendő a P300 csúcsának. A P3b-t követő lassú hullámot egyes irodalmak külön komponensként kezelik, mások pedig a P3b részeként tekintik azt. A komponensek az agy több területén mérhetőek, míg a P3a inkább a frontális (Fz) területen, addig a P3b, illetve az utána következő lassú hullám a parietális - falis - lebeny (Pz, P3, P4) régiójában figyelhető meg. A P300-hoz kötődő kutatások során arra is fény derült, hogy a két komponens különbözően viselkedik, különböző stimulusokra. [13][14] A P3a megelőzi a P3b-t, csúcsa kb. 250 ms-nál jelenik meg. Mérések szerint a jel amplitúdójára nincs kihatással az, hogy éppen milyen mértékben figyelt az alany a stimulusra. A P3a megjelenése azt mutatja, hogy az alany számára egy esetlegesen hasznos információt hordozó inger érte, ám az még nem került feldolgozásra. A P3b, amely standard körülmények között kb. 300-350 ms-nál

csúcsosodik erősen függ mind a környezeti, technikai paramétereiktől, mind a vizsgált alanytól.

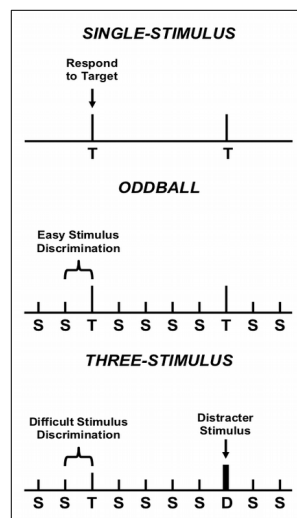


2. Abra: A P300 komponensei: P3a és a P3b [21]

A P300 ERD pontos forrására a mai napig nincs egyértelmű válasz. A felfedezése óta történő vizsgálatok, mérések és kutatások során azonban egyes régiók kiemelésre kerültek. A P300 hullám, az inger aktiválódása után először az agy frontális területén jelenik meg – a P3a komponens az, ami a legjelentősebben mérhető ezen a régión –, majd kis késleltetéssel halad tovább a középvonali – Fz, Cz, Pz – elektródákon keresztül a parietális területen elhelyezkedő elektródák fele. Knight [13] vizsgálatai szerint, azon alanyok, akik sérült vagy normálistól eltérő frontális lebennyel rendelkeztek kisebb P3a amplitúdót produkáltak a általánosnál. Fokális hippocampuszbeli elváltozással rendelkező alanyok szintén hasonló eredményt szolgáltattak. A P3b komponensre nem volt negatív hatással az előző két sérülés, azonban a temporális-parietális csomópont integritásának hiánya jelentősen befolyásolta a parietális területen. A továbbiakban a P300 jelet, mint egészét befolyásoló tényezők kerülnek bemutatásra. [14]

P300 jel kiváltásához a legalkalmasabb módszernek az oddball (kakukktójás) paradigma bizonyult. Az oddball vizsgálat során a vizsgált alanyt egy előre meghatározott eseményt, azaz célingert kell észlelnie - pl. sok betű közül egy adott betű felvillanása, különböző vagy azonosan mély hangok között egy magas hang felismerése –, azonban ez a releváns esemény vagy stimulus ritkábban fordul elő, mint a többi, háttér esemény. Az oddball paradigma többféleképpen is alkalmazható. Egyik verziójában, nincs háttéringer, a releváns stimulus magában jelenik meg bizonyos időközönként egyszer. A leggyakoribb verziójában a háttér stimulusok között jelenik meg a felismerendő inger –

pl. mély hangok között a magas –, emellett pedig a három-inger paradigma is sok helyen megjelenik. A három-inger esetében a háttérben nemcsak egy, az alanynak felismerendő inger jelenik meg, hanem egy másik ún. figyelemelvonó inger is. Ennek szerepe a következőkben látszani fog. A paradigmát alkalmazó mérések során a háttér és a releváns ingerek jellegének és gyakoriságának módosításával változó amplitúdóval és eltérő késleltetéssel rendelkező P300 jelek ingerelhetők. Nem csak technikai paraméterektől, hanem a vizsgált személytől is függ a P300 karakterisztikája. Olyan egyénfüggő paraméterek ezek, mint pl. a fáradtság, mentális állapot, fizikai vagy pszichológiai egészség, de már a figyelem mértéke is erősen befolyásolja a mérés eredményét. Példa Duncan-Johnson és Donchin (1977) mérése, amelyen arra kérték a pácienseiket, hogy két hang randomizált váltakozásának sorrendjéből álló szekvencia hallgatása közben az egyik hangot számolják. Az ERP rögzítésének során arra lettek figyelmesek, hogy azok a hangok, amelyeket az alanyaiknak számolniuk kellett valamelyest nagyobb amplitúdóval rendelkező jelet generáltak, továbbá a hang előfordulásának valószínűsége is kihatással volt a csúcserték alakulására. [2][14]



3. Abra: Az oddball paradigma 3 típusa [13]

A célstimulushoz kötődő valószínűség nagy mértékben befolyásolja az ingerelt P300 amplitúdóját. Minél valószínűtlenebb a célinger, annál nagyobb a kiváltott jel amplitúdója. Egy ideig ezt a valószínűséget a stimulus valószínűségének nyilvánították, azonban 1981-ben Fitzgerald és Picton egy új nézetbeli változtatási javaslatot tettek, amelyet később Polich (1990) és Gonsalvez kísérletei megerősítettek. Az új elmélet szerint nem pontosan a stimulus előfordulásának valószínűsége, hanem az egy rövid időintervallumon belül előforduló célingerek között eltelt idő – másképp időbeli valószínűség - az, amely a P300 amplitúdójának nagyságát befolyásolja. Fitzgerald és

Picton változatlanul, 0.2-n hagyták a stimulus megjelenésének valószínűségét, ám a célstimulusok között eltelt idő, azaz az interstimulus intervall (ISI) 0.5 s-ról 4 s-ra növelték, azaz csökkent az időbeli valószínűsége az ingernek. A P300 amplitúdója monoton nőtt az időbeli valószínűség csökkenésével. Pochkin – később Gonsalvezzel együtt – 1990 és 2002-beni eredményei megerősítették az elméletet. Méréseik során az ISI 10s fölé növelése alatt, egyes esetekben nem változtattak a stimulus valószínűségén, míg másokban növelték azt. Az összes mérés eredménye arra vezetett, hogy kb. 10 s-on felül már nem nő, sőt csökkenhet a P300 amplitúdója, azaz a jel időbeli korláttal rendelkezik és habár fontos paraméter az inger valószínűsége, az időbeli valószínűség a domináns paraméter a kettő közül. Ezen megfigyelések olyan későbbi elméletek felállítását eredményezték, amelyek szerint a P300 a memóriához köthető. [14]

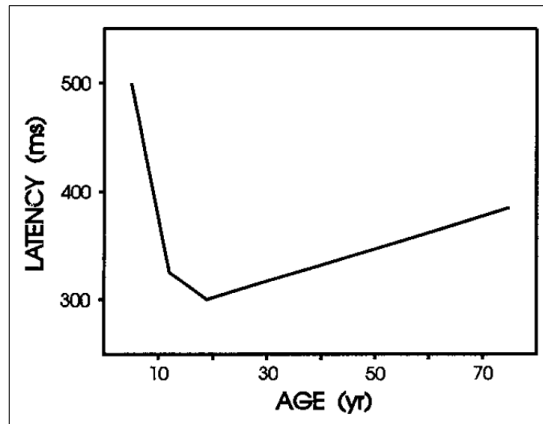
A célstimulus valószínűsége és a P300 amplitúdója közti kapcsolathoz érdemes még egy megjegyzést fűzni. Courchesne 1977-ben végzett kísérlete során betűdetekcióra kérte fel vizsgált alanyait. A betűdetekció során az alanyoknak azt kellett figyelnie, hogy mikor jelenik meg a „B” betű, ez szolgált célingerként. Háttéringernek azonban két külön esetet választott: csak „A” betűk jelentek meg vagy az ABC minden betűje megjelent, csak a „B” nem. Mindkét háttérstimulus alkalmazása ugyanahhoz a megfigyeléshez vezetett. A 0.15-ös valószínűséggel előforduló „B” karakter nagy, amíg a háttér karakterek kicsi amplitúdóval – vagy azzal egyáltalán nem - rendelkező P300-at eredményeztek. A megfigyelés alapján a P300 jel és a valószínűség kapcsolata inkább, a cél-háttér kategóriák közötti valószínűségen alapszik, mint az egyes ingerek egyénekenkénti valószínűségén. [14]

A P300 hullám rögzítéséhez a mérésvezetők az alanyoknak különböző típusú és nehézségű feladatokat határoznak meg. A feladatok nehézsége adódhat a feladat komplexitásából vagy a célinger háttértől való disztinválásából. Ahogyan az alany a komplexitásra és a megkülönböztetésre máshogy reagál, úgy a P300 jel is módosul. A feladat nehézségével a P300 hullám késleltetése nő, egyes esetekben amplitúdója csökken. Ezt az effektust azonban nem a feladat komplexitása, hanem az alany bizonytalanságának mértéke okozza. Vizsgált alanyoknak három különböző feladatot kellett végrehajtaniuk Kutas 1977-ben tartott kísérlete során. Az első feladat a „Nancy” név észrevétele volt a „David”-ek között, a második női nevek észrevétele férfi nevek között, a harmadik pedig „prod” szó szinonimáinak felismerése egyéb szavak között. A kísérlet eredményei szerint a P300 hullám amplitúdója nem változott jelentősen a

feladatok között, azonban az első feladattól a harmadikig a jel késleltetése igen. Kutas kísérletén kívül további mérések eredményei alapján azaz elmélet állt fel, hogyha különböző komplexitású feladatok azonos magabiztossággal végezhetőek el, akkor a P300 amplitúdójában nem keletkezik változás, tehát a hullám csúcserkének kitérése attól függ, hogy a vizsgált személy milyen magabiztosan képes a célstimulus kiválasztására a háttérstimulusok közül. Ez a tény a jelentősen hozzájárul a P300-hoz kötődő feladatok összeállításához. A célinger-háttér különválasztásának nehézségéből adódó probléma azonban nem azt vonja maga után, hogy minél egyszerűbb a célstimulus disztingválása, annál nagyobb amplitúdóval rendelkező P300 keletkezik. A vizsgált alanyok, ahhoz hogy P300 hullám keletkezzen nagy fokú odafigyelésre van szüksége, ugyanis a figyelem csökkenésével arányosan a keletkezett jel amplitúdója is csökken. Ebből következik, hogyha egy feladat nem elég megerőltető vagy nem kíván nagy fokú odafigyelést, akkor az alany könnyebben veszíti el fókuszát, fogékonyabbá válik a külső, zavaró ingerekre. A P300 karakterisztikája – amplitúdója és késleltetése – egy fajta módon azt is jelzi, hogy az általános ébredési szint, amely a feldolgozási rendszert modulálja, mennyi figyelmi erőforrást allokált a P300-at kiváltó inger által jelzett feladat teljesítésére. Gyakran, hogyha a feladat nem hordoz különösebb komplexitást magában, a célstimulusok számolását kérik az alanytól a mérésvezetők, így fenntartva a megfelelő állandó figyelmet. [13][14]

A figyelem azonban nemcsak a feladat nehézségétől vagy érdekességétől függ, hanem a vizsgált alanytól is. Az alany kora, mentális állapota, fizikai, pszichológiai egészsége, valamilyen szer által való befolyásoltsága jelentős hatással van az ingerelt P300 hullámra. A P300 késleltetése már a gyermekkortól változik és változása a felnőtt korban sem áll meg. Körülbelül 5-től 12 éves korig bezárólag nagyjából 25ms/év a késleltetés csökkenése, majd a folyamat 1-5ms/év-re lassul kb. 20 éves korig. A legkisebb időintervallum az inger megjelenése és a P300 csúcsa között valahol 15 és 20 éves kora között van az egyénnek. 20 éves kor után lassú folyamatos, megközelítőleg 1.3ms/éves – 31 ms-os standardizált regressziós hiba mellett - növekedés jellemzi a csúcs késleltetését az öregedéssel. Klinikai rendellenességben szenvedő személyek – pl. schizofrenia, súlyos depresszió vagy autizmus – P300 jeleiben amplitúdócsökkenés figyelhető meg, illetve egyes esetekben a jel térbeni disztribúciója is eltérő a hasonló vagy egyéb betegségekben nem szenvedő személyekétől. Habár általánosságban az alkohol és egyéb drogok negatív módon befolyásolják a P300 jellegét – kisebb

amplitúdó, nagyobb késleltetés -, érdekes tény, hogy bizonyos nehézségű feladatok és figyelési stratégiák mellett nem volt megfigyelhető csökkenés valamilyen hatás alatt álló alany kiváltott jelében. [14]



4. Abra: P300 késleltetésének alakulása az életkor függvényében [14]

3 Brain-Computer Interface rendszerek

Az orvostudomány és a számítástechnika fejlődése olyan agyjelek és agyhullámok megfigyelését tette lehetővé, amelyek rengeteg információt hordoznak az agy és azon keresztül az egész emberi test működésével kapcsolatban. Azonban ezek a jelek nem csak az emberi agy – akár elme – működésével kapcsolatos kutatások vagy betegségek diagnosztizálása során nyújtanak segítséget, hanem megfelelő feldolgozásukkal Brain-Computer Interface (BCI) rendszerek kidolgozása, fejlesztése és alkalmazása is lehetséges. A BCI rendszerek lehetőséget nyújtanak felhasználóik számára, hogy gondolattal irányítsanak valós vagy virtuális objektumokat, ezáltal az orvosi felhasználástól kezdve a szórakoztató iparig rengeteg területen megtalálhatók. Segítségükkel pl. súlyos idegrendszeri sérülést szenvedett vagy bezárt szindrómás emberek válnak képessé mozogni vagy kommunikálni a külvilággal. A BCI rendszerek bemeneteként az agyból valamilyen jelrögzítő vagy képalkotó technológiával – pl. EEG, EcoG, MEG, MRI, fMRI vagy CT – vett jelek szolgálnak, kimeneteként pedig az az esemény, amelyet felhasználója el szeretne érni. A bemenet és kimenet között komplex jelfeldolgozás történik. Több BCI típus különböztethető meg, amelyek más és más típusú bemenetet fogadnak, illetve más és más ideológia szerint működnek. A rendszerek használata nem egyszerű, a használó alanynak meg kell tanulnia irányítani, továbbá más az agy működéséből eredő következmények is megnehezítik a használatukat. Szerencsére a BCI rendszerek fejlesztése már több, mint 30 éve folyik – pl. Farwell és Donchin P300 ERP jelalapú betűző rendszere – és napjaink technológiájával már akár mind a négy végtagjuk mozgását elveszítő emberek is képesek visszanyerni mozgásukat pl. BCI vezérelt exoskeletonok segítségével. [1][3]

3.1 BCI rendszerek osztályozása és felépítése

Egy Brain-Computer Interface rendszer célja, hogy kapcsolatot teremtsen a használójának agya és a jeleket feldolgozó számítógép között. A kapcsolat által lehetőség nyílik az emberi test bizonyos részeinek – pl. neuromuszkuláris csatornák – megkerülésére. A teljesség igénye nélkül pár példa olyan betegségekre vagy sérülésekre, amelyek esetében a BCI rendszerek segítséget nyújtanak – akár folyamatos hétköznapi használatként, akár időszakos terápiás vagy rehabilitációs jelleggel - :

gerinc- és/vagy nyakcsigolya törés, amiotrófiás laterálszklerózis (ALS), sclerosis multiplex, agytörzsi stroke, izomsorvadás. [1]

3.1.1 Rendszerek osztályozása

A BCI-eket több módon is lehetséges osztályozni, ezek közül a továbbiakban két csoportosítási mód kerül említésre. Az első kategorizálási módban, aszerint különülnek el a rendszerek egymástól, hogy a jelrögzítési technológia, amin alapszanak invazív, félig invazív vagy non-invazív. Többnyire a non-invazív és a félig invazív rendszerek kerülnek alkalmazásra, mivel kevesebb kockázattal – pl. műtéti komplikáció – bírnak és több szabadságot biztosítanak a felhasználóik számára. Non-invazív esetben többnyire EEG vagy magnetoencefalográfia (MEG) – lehetséges CT, MRI vagy fMRI használata is – biztosítja az agyban ingerelt jelekről az információt. A jelrögzítés a fej felszínéről történik elektródákon keresztül. A MEG eljárás előnye az EEG-vel szemben, hogy nagyobb térbeli felbontást tud nyújtani az agyról, így pontosabb adatot ad a jelek lokalizációjáról. Azonban az EEG során a páciens mozgásában kevésbé limitált és az felhasznált technológia is olcsóbb. Elektrokortikográfia (ECoG) fél-invazív eljárások során kerül alkalmazásra. Az elektródák a koponyán keresztül fúrt lyukon keresztül kerülnek beültetésre, nem közvetlen az agyba. Az elektródák nem az egész fej felületét fedik le, hanem csak azokat a területeket, amelyeken a mérni kívánt jelek keletkeznek. Az ECoG eljáráshoz használt elektródák beültetése előtt MEG, MRI vagy CT segítségével meghatározásra kerülnek azok a pontos területek, ahol azok a funkcionális területek helyezkednek el, amelyek jelei szükségesek a rendszer helyes működéséhez. A második csoportosítási mód szerint független és nem-független BCI-k különülnek el. A független rendszerek új, az emberi szervezetben létezőktől független kommunikációs csatornát létesítenek az agy és a számítógép között. A rendszert használó személy gondolataival, cselekvések vagy mozdulatok elképzelésével, továbbá az agyában külső ingerre keletkezett jelek – kiváltott potenciálok, pl. P300, VEP, stb. - segítségével tud feladatokat teljesíteni. A függő BCI rendszerek használatához szükségesek egyes motoros funkciók, nem teremtenek független csatornák az ember agy és a számítógép között. [4][9]

3.1.2 Rendszerek felépítése

Minden BCI rendszer alapépítőkövei a következők: agyi aktivitást figyelő és rögzítő egység, jelfeldolgozó, dekóder és visszajelzést biztosító egység.

Az agyi aktivitást figyelő egység feladata, hogy mérje az agyhullámokat és jeleket, továbbá ezeket elektrophysiológiai jelekből, elektromossá transzformálja. Ilyen eszköz pl. MEG, EEG, ECoG vagy az fMRI. A jelfeldolgozó egység elvégzi a szükséges jelerősítést, szűrést és jelanalízist. Különböző algoritmusok segítségével választja ki a releváns jelkomponenseket, amelyet továbbít a dekódernek. A dekóder feladata, hogy a kapott adatokat parancsokká alakítsa át. A parancsok hordozzák a felhasználó által kívánt cselekményt, amelyről a visszajelzést biztosító egység ad információt az alanyoknak. [1] [17]

A BCI-vel való kommunikáció lehet felhasználó-vezérelt vagy ingerlés-vezérelt. Az első esetben a felhasználó a gondolataival és képzelő erejével generál olyan agyi aktivitást, amellyel képes a BCI-t irányítani. A második esetben a BCI stimulusok sorozatával váltja ki a felhasználó agyában a feladat elvégzéséhez szükséges jelet. [17]

3.2 Rendszerek alkalmazhatósága

Ahhoz, hogy egy BCI hasznos segítségével legyen a felhasználónak, gyors kommunikációt és a kívánt aktivitásról pontos predikciót kell biztosítani. A két elvárás szimultán teljesítése koránt sem egyszerű feladat és további nehézség, hogy az alkalmazott jelrögzítési technológia befolyásolja az előző két feltétel határait.

A non-invazív MEG, MRI és CT segítségével ugyan pontos kommunikációt tud biztosítani, de sebességében távol maradnak az EEG-től. Egyelőre a félig invazív ECoG jelenti a középutat az elvárások teljesítése és a műtéti, illetve későbbi egészségügyi komplikációk elkerülése között. A vezeték nélküli kommunikációra alkalmas ECoG sebességét maga a kommunikáció sebessége a feldolgozó egységgel limitálja, példa erre a [3]-ban használt epidurális ECoG.

A BCI bemenetén kapott jelek alapján meghatározott cselekmény helyessége a felhasználón is múlik. A teljes rendszer sikeres működése nemcsak a BCI-on múlik, hanem annak használóján. A rendszer irányításának elsajátítása az esetek csak egy bizonyos részében nélkülözhető – [4]-ben publikált egyik eset során a vizsgált alany, egy nap alatt, csupán a mozgás szándékával képes volt a megfelelő irányba mozdítani

a képernyőn látható kurzort - , többségben - akár csak gyermekkorban a járást vagy a beszédet – a rendszer megfelelő használatát meg kell tanulni. A tanulási folyamathoz, továbbá a rendszer sikeres jövőbeli használatához elengedhetetlen, hogy a BCI valamilyen visszajelzéssel rendelkezzen. Ez a visszajelzés lehet pl. a felhasználó által választott betű megjelenése a képernyőn vagy az exoskeleton segítségével megmozdított végtag. (Az egészséges ember számára is létezik visszajelzés: hogyha az ember meg szeretne fogni egy poharat, akkor érzékeli, hogy megemeli karját és megérinti, majd megmarkolja a poharat.)

Az agyban keletkező jelek és hullámok karakterisztikái és lokalizációi mentális, fizikai és egészségügyi állapot szerint, de egyénenként is változnak. A pillanatnyi állapoton túl az idő múlása és az öregedés is kihatással van jelekre – a változás nem csak években, hanem hetekben is már mérhető - . Ahhoz, hogy egy BCI folyamatosan helyesen funkcionáljon, bizonyos időközönként frissíteni kell a paramétereit. A frissítés a BCI használatának megszakításával jár, ezért a rendszer fejlesztésénél jelentős szempont az, hogy minél ritkábban legyen szükség a beállítások újrakalibrálására. [1] Alapvetően kétféle hozzáállás létezik a probléma megoldására: az első esetben [3]-hoz hasonlóan a rendszerbe adaptív dekóder – pl. Adaptive Brain Signal Decoder (ABSD) keretrendszer - , amíg a másodikban hasonlóságok felismerésén alapuló dekóder kerül integrálásra a rendszerbe. Az adaptív dekóder új adatok gyűjtésével és új minták tanulásával alkalmazkodik a megváltozott állapothoz, míg a hasonlóságokon alapuló verzió az új és a régi állapot közötti hasonlóságokat keresi és ismeri fel. [5]

Alkalmazhatóság szempontjából biztató tény, hogy ugyanazon feladat megoldása lehetséges az agy eltérő területeiről származó jelek segítségével is. [4]-ben bemutatott négy eset azt hivatott alátámasztani, hogy akár olyan agyi területekről is el lehet végezni egy motoros BCI irányítását, amelyek korábban nem számítottak ideális területnek. Fél-invazív ECoG alapú BCI alkalmazásával mind a négy esetben ugyanazon feladat teljesítése, a képernyőn megjelenő kurzor adott célpontba való eljuttatása volt a cél. Mindegyik esethez külön páciens tartozott, továbbá mind a négy alany agyának eltérő területeire kerültek beültetésre az ECoG elektródák. A feladat során a kurzor irányítását a vizsgált pácienseknek különböző testrészeik – pl. nyelv, jobb és bal kar, lábfej - mozgatásának, illetve ismerős hangok – pl. mobiltelefon csörgés vagy családtag hangja - elképzelésével kellett elvégezniük. Mindegyik esetben, mind a négy páciens sikeresen tudta a feladatot abszolválni, csupán 2-7 napnyi tanulás után. Ugyan a kutatásban kevés

alany vett részt, továbbá az egyénnel kapcsolatban több egészségügyi faktor nem került kivizsgálásra, az eredmény fontos jelentőséggel bír. A kutatás konklúziója, hogy az agy több területének jelei is alkalmasak arra, hogy megfelelő bemenetként szolgáljanak motorikus BCI-eknek. Alkalmazhatóság szempontjából ez a tény akkor válik előnyössé, amikor egy sérült agyi terület funkcionalitását kell megvalósítani egy egészséges terület segítségével.

3.3 Létező rendszerek bemutatása

A számítástechnika és az orvostudomány fejlődése, továbbá a technológia utáni érdeklődés csak az elmúlt 20 évben vált olyan mértékűvé, hogy több kutatási projekt is indulhasson és futhasson a témában. Ennek ellenére kutatások BCI-ek fejlesztésével már több, mint 30 éve is foglalkoztak. 1988-ban publikálta Farwell és Donchin a P300 betűző fejlesztéséről és teszteléséről szóló cikküket. Az általuk megalkotott P300 ERP alapú BCI a mai napig gyakran alkalmazott rendszer, amely több további hasonló eszköz fejlesztését alapozta meg. 2019 októberében pedig egy új BCI került bemutatásra, amely segítségével egy mind a négy végtagjára lebénult, 28 éves férfi mozgáskorlátozottsága csökkent drasztikusan. A következő két fejezetben az előbb is említett régebbi, azonban meghatározó szereppel rendelkező P300 betűző és az új, napjaink technológiai fejlettségét jól tükröző BCI rendszerek kerülnek bemutatásra.

3.3.1 P300 speller

A betűző megalkotásának és tesztelésének fő célja a P300 event-related potential (ERP) jellel való kommunikáció lehetőségeinek és határainak a felderítése volt. Két fő szempont szerint került kiértékelésre a kifejlesztett rendszer: a kommunikáció sebessége és a betűzés pontossága. A BCI fejlesztéskor a célközönséget legfőképp azok az emberek alkották, akik szinte teljesen elvesztették motorikus kontrolljukat – pl. kéz- vagy lábmozgatás, beszédképesség -, akár betegség, akár sérülés miatt.

A készülék egy 6x6 mátrix, amelyet az angol ABC betűi, pár speciális karakter és hat darab egyszavas parancs – ezek között a törlés, space, kilépés, szöveg kimondatása – alkot. Az agyban ingerelt jelek rögzítését elektroencefalográfiával (EEG) végezték, amelynek elektródáit a Pz – parietális lebeny középvonali lokalizációja – környékére helyezték a 10-20-as standardizált rendszernek megfelelően. A Pz választásának szempontja, hogy amíg a P300 korai komponense a P3a inkább a frontális lebenynél

(Fz) mérhető, addig a P3b – amely a P300 amplitúdó csúcsának tekintendő – legjelentősebben a parietális területek környékén mérhető. A jeleket rögzítő elektródák anyaga ezüst-ezüst-klorid, amely pontos felvételt biztosít lassú potenciál változások esetén is, továbbá impedanciájuk nem haladta meg az ökölszabályként alkalmazott 5kOhm-os határt. A jelek erősítés után először egy 35 Hz-en limitált alul-áteresztő, majd egy 0.02 Hz-en határolt felül-áteresztő szűrőn keresztül lettek kondicionálva. Végző lépésként 50 Hz-en kerültek digitalizálásra. A melléktermékek kiszűrhetősége érdekében az alany jobb szeme alá és fölé helyezett – elektro-oculogram (EOG) által felhasznált – elektródák kerültek.

MESSAGE					
BRAIN					
Choose one letter or command					
A	G	M	S	Y	*
B	H	N	T	Z	*
C	I	O	U	*	TALK
D	J	P	V	FLN	SPAC
E	K	Q	W	*	BKSP
F	L	R	X	SPL	QUIT

5. Abra: Farwell és Donchin P300 alapú betűzője [2]

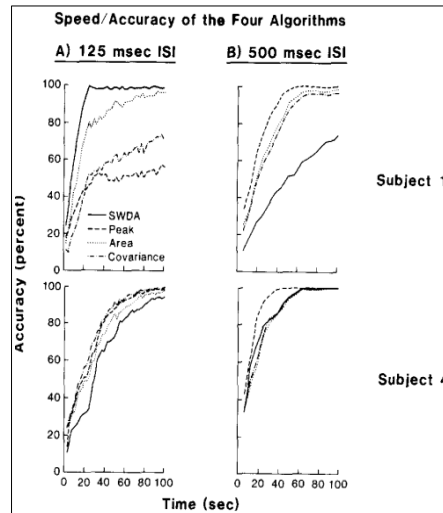
A kísérletre alanyként négy egészséges állapotú – 20 és 36 év közötti - önkéntes jelentkezett és a mérés a következőkben összefoglaltak szerint történt. Az alanyoknak először meg kellett ismerkedniük a mérés folyamatával és az eszköz használatával, ezután pedig el kellett végezniük a meghatározott feladatot. A feladat sikeres teljesítéséhez az jelentkezőknek a „BRAIN” szót kellett lebetűzniük és a géppel kimondatniuk a „TALK” utasítás – amely használata a betűk kiválasztásával megegyező módú – segítségével. A betűzés folyamata során a mátrix oszlopai – majd ezután sorai – véletlenszerű sorrendben villantak fel. P300 jel kiváltásához egy ritka célingerre és háttéringerekre van szükség. A betűző estében a célingert annak a sornak – vagy oszlopnak – a felvillanása jelentette, amely tartalmazta a kívánt karaktert – vagy parancsot –, valószínűsége 1 a 6-hoz. Háttérstimulusokként azok a sorok – vagy oszlopok – szolgáltak, amelyek nem tartalmazták célkaraktert. Az egy hat darab villanásból álló sorozat alkot egy próbát. Azonban az ingerelt P300 jel amplitúdója gyakran nehezen felismerhető, ezért több a kiváltó inger többszöri megismétlésére, azaz

több próbára van szükség, amelyeknek az eredményeit átlagolni kell, hogy csökkenjen a signal-to-noise-ratio (SNR). Az átlagolási folyamat, illetve az előre meghatározott pontosság eléréséhez szükséges próbák száma dominánsan korlátozzák a betűző használatának sebességét. A próbák ismétlésének számának meghatározása során további tényezőként szolgált az EOG által érzékelt jel. A páciens pislogása vagy arcmozdulata eredményeként az adott próba semmisnek minősült.

Megfelelő mennyiségű adat felvételéhez minden alanyak tízszer – tíz feladat blokkban megoldási - kellett elvégeznie a betűzési feladatot. A tíz felvételi blokk kétszer öt darabra lett szétválasztva. Két páciensnek az első öt blokkban 125 ms-os, a második ötben pedig 500 ms-os stimulus közti időközzel (ISI) villantak fel a sorok – vagy oszlopok -. A másik két páciensnek ugyanúgy kellett elvégeznie a feladatot, de felcserélt időközökkel. (Az ISI, angolul inter-stimulus interval a két villanás, azaz a két inger inicializálása között eltelt időt adja meg. Az ingerek közti időköz hatással van a P300 amplitúdójára és késleltetésére, ezáltal a betű kiválasztásának pontosságára.) Egy feladatblokk 30 – egyes esetekben a pislogás vagy betűhibázás, törlés és újraválasztás miatt kicsivel több – próbából állt. A próbák további 600 ms-os időtartamú al-próbákra lettek felbontva. Mivel egy stimulus 100ms-nyi világításból és 25 ms-nyi szünetből – ekkor a mátrixon nem világít semmi –, ezért ez azt jelenti, hogy 5 darab stimulus van egy al-próbában. A jelfeldolgozás al-próbánként történt, ami azt jelenti, hogy feldolgozás közben minden próba között volt 125 ms-nyi – azaz egy ingernyi – átfedés, amely azonban nem rontotta a feldolgozást minőségét jelentős mértékben.

Betűzés során az alanyok az éppen következő betű villanásának megfelelő fokú figyelésével és a betű villanásának számolásával támogatták az agyukban kialakuló P300 jel ingerlését. Minden sor-/oszlopvillanás, amit észlelt a páciens, ERP-t ingerelt az agyában. Ahogyan feljebb is említésre került, eközben a feldolgozás al-próbánként történt. Minden al-próbához egy pont került hozzárendelve, amely azt mutatta meg, hogy milyen mértékben volt jelen P300 benne. A pont hozzárendelést négy különböző típusú algoritmussal is elvégezték, ez a négy fajta algoritmus a lépcsőzetes diszkrimináns analízis (SWDA), csúcsválasztó (peak picking), területszámítás, és kovariancia számítás volt. (Mind a négy típusú algoritmusnak rendelkezik előnnyel és hátránnyal is.) Minden betűhöz ezáltal egy szummázott pont tartozott, ami annak a sor és oszlop pontoknak az összegeként alakult ki, amelybe tartozott. A rendszer a pontozás alapján választotta ki azt a legmagasabb pontértékkel rendelkező mátrix sor- és

oszlopmetszetet - ezáltal a metszetben található betűt - , amelyre a felhasználó gondolt. Azonban ahhoz, hogy egy bizonyos százalékos pontossággal meg tudja határozni a rendszer a kívánt betűt, több al-próba pontozásának eredménye is szükséges.



6. Abra: Pontozó algoritmusok teljesítménye két alany és két ISI esetében [2] (módosított kép)

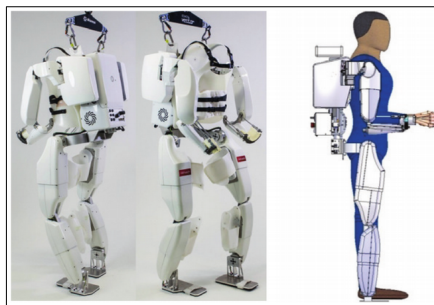
A kísérlet eredményei ígéretes jövőképet mutattak az EEG és P300 alapú BCI rendszerek alkalmazásáról, tovább fontos tanulságokat is magukban hordoztak. Az eredmények kiértékelése a felvett adatok alapján offline történt. Bootstrapping technikával becsülték meg a rendszer pontosságát, amelyhez véletlenszerűen válogatták össze az adathalmazokat. 2-40 próba utáni eredményekig gyűjtöttek össze 1000-1000 mintát. A kiértékelés a pontosság és a meghatározott számú próbák elvégzéséhez szükséges idő függvényeként történt. A legjobb eredmények a következők: 80%-os pontosság eléréséhez átlagosan 20.9 s-ra, míg 95%-os pontosság eléréséhez átlagosan 26 s-ra volt szüksége a felhasználónak és a rendszernek közösen. Az átlagos időtartamok magukba foglalják mind a al-próbák, mint a próbák között eltelt szünetek időtartamait is. A kiértékelés során használt különböző típusú algoritmusok eltérő eredményeket mutattak a négy páciens esetében. Mint ahogyan 2.3.2-ben is említésre került, minden ember eltérő, sajátos karakterisztikájú jeleket produkál. Az algoritmusok teljesítménye közti eltérés oka abban rejlik, hogy a páciens agyában kiváltott P300 általánosan kisebb vagy nagyobb amplitúdóval rendelkezik, illetve a jel késleltetésében milyen mértékű a variancia. Példaként: a csúcsválasztó az amplitúdó alapján végzi a jelek szeparálását, tehát rosszabbul teljesít általánosan laposabb P300-ak detektálásában, mint pl. az SWDA, azonban az SWDA a késleltetés varianciájára érzékeny, amelyre a csúcsválasztó nem, így ifőben variánsabb jelek esetén az SWDA a csúcsválasztóhoz

képeket marad alul. Végző eredményben a 125 ms-os ISI alkalmazása esetében az SWDA, 500 ms-os ISI alkalmazása mellett pedig a csúcsválasztó algoritmus bizonyult a leggyorsabban a legpontosabbnak négyből három páciensnél.

A Farwell és Donchin által publikált BCI tesztelésének folyamata, annak eredményei, továbbá a megalkotott eszköz is jelentős fontossággal bír a BCI rendszerek technológiája és az ERD jelek kutatásának területein. [2]

3.3.2 Exoskeleton irányítása vezeték nélküli ECoG alapú BCI rendszerrel

2019 októberében a Lancet Neurol publikálásában megjelent koncepció bizonyítás jól tükrözi napjaink state-of-the-art BCI technológiai fejlettségének állását és alkalmazhatóságát. A program célja, hogy napjaink képalkotó és jelrögzítő, továbbá végtagprotézis technológiáját felhasználva, az agyi jelekről eddig szerzett információk alapján olyan korszerű BCI rendszer valósuljon, amely hosszútávon képes megkönnyíteni mozgáskorlátozott emberek életét. A kutatás során egy 28 éves korú, mind a négy végtagjára lebénult férfi kapta vissza részben mozgásképességét. Az alkalmazott BCI rendszer részeit két, egyenként 32 elektródával ellátott fél-invazív módon beültetett vezeték nélküli elektro-kortikográfiai (ECoG) egység, a jelfeldolgozást és robottestet irányító számítógép és az exoskeleton alkotja.

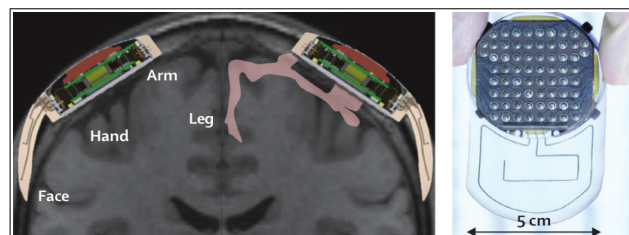


7. Abra: Exoskeleton és hátulján a jelfeldolgozásért és az irányításért felelős számítógép [3]

A kutatás vizsgált alanyaként egy férfi vett részt, aki egy C4-C5-ös gerincvelő sérülés miatt felső végtagjaira részleges – baloldali kézi könyökét és csuklóját részben képes volt használni a sérülés után - , alsó végtagjaira betíg teljesen lebénulást szenvedett. A program előtt egy joystickkal irányítható kerekesszéket használt kizárólag, mint kiegészítő eszköz. A programban való részvétel össze követelményének megfelelt, amelyek között olyan kizáró okok szerepeltek, mint pl. korábbi agyműtét, kontraindikáció a magneto-encefalográfiára (MEG), elektroencefalográfiára (EEG) vagy mágnesrezonancia-

képalkotásra (MRI), szerfüggőség megléte vagy depresszív állapot. Az alany önként jelentkezett a programra. A férfin kívül további egy alany nyújtotta be jelentkezését a programra, azonban sajnos az ECoG elektródák beültetése során fellépet komplikációk miatt kizárásra került a programból.

Az alany fejébe két darab, egyenként 32 darab platinum-iridium és 5 darab referencia elektródával ellátott, vezeték nélküli ECoG került beültetésre fél-invazív módon, azaz az elektródák nem hatoltak be az agyba. Az ECoG használatának előnye, hogy nagy térbeli felbontásra és ms-os nagyságrendbeli jelrögzítésre képes. A két vezeték nélküli ECoG egység kommunikációját és energia ellátását rendre egy szilikon fülbe csomagolt ultra-magas frekvenciás (402-405 MHz), továbbá egy magas-frekvenciás (13.56 MHz induktív) antenna biztosítja. Az ECoG egységek beültetése a szenzomotoros kéreg fölé került. Az elektródák pontos lokalizációja előzetesen feltérképezésre került. Ehhez az alanynak MEG és fMRI alatt mozgatnia valósan és virtuálisan felső és alsó végtagjait. Minden egyes mozgást MEG alatt 100-szor, fMRI alatt 80-szor meg kellett ismételnie. A MEG segítségével a szenzomotoros kéreg neuronjainak aktivációja által kiváltott és az fMRI-vel a vér-oxigén-szint függő (BOLD) megfigyelt jelek alapján készült képek összegzésre kerültek. Az így kapott összegzett kép alapján került meghatározásra annak a szenzomotoros kéregbeli területnek a pontos térbeli koordinátái, amely a legaktívabban részt vett a páciens által végzett feladatokban. A kijelölt pont feletti ECoG elektródái kizárólag a felső végtagok mozgatásért felelős területet fedik le.



8. Abra: A fél-invazív módon beültetett vezeték nélküli ECoG jelvevők [3] (módosított kép)

A páciens agyi jelei és az exoskeleton közötti kapcsolatot az Adaptive Brain Signal Decoder (ABSD) keretrendszer valósította meg, amelybe a kutatási protokoll, a szoftver és a dekóder algoritmus is integrálásra került. Az ABSD tanítása és frissítése valós időben feldolgozott adatokkal történt. Egy meghatározott feladat végrehajtása közben a az ECoG jelei 586 Hz-en kerültek mintavételezésre. Az ECoG jelei mellett, a kéz pozíciójának koordinátáit, a csukló szögét és a séta indításáért felelős állapotot is rögzítve lett. A tanulási fázis alatt a kézmozgatás feature-einek kiválasztása a

kézmozgatás koordinátáinak idő szerinti függvénye szerint történt. Ez szintén igaz a csukló szögének forgatására – szög az idő függvényében - és a sétálási mód változtatására is. Az epidurális ECoG jelek feature szelektására a komplex folytonos hullám transzformáció (CCWT) lett kiválasztva. Az adaptív dekóder lehetővé teszi a későbbi valós-idejű tanulást, ezáltal megnövelve a dekóder frissítések között eltelt időtartamot. A dekódert két algoritmus alkotja: egy rekurzív, exponenciálisan súlyozott, n-utas, részben legkisebb négyzetes regresszió és egy Markov váltó lineáris modell. A második algoritmus a dekódolás stabilitásának és a pihenési állapot támogatásának fejlesztéséért szükséges. A 8 szabadságfokban kontrollálni képes dekóder frissítési ciklusának időtartama elérheti akár a 7 hetet is.

A páciens tanulását különböző fizikai és virtuális segédeszközök támogatták. Virtuális környezetben az 1D-s mozdulatok elsajátításához Pong-szerű játék ütőjét kellett irányítani az alanynak. Ugyancsak virtuális környezetben, avatar segítségével fejlesztette a különböző 2D-s és 3D-s mozgásokhoz, illetve a járás be- és kikapcsolásához szükséges tudását. Valós környezetben a felszerelt exoskeleton karok irányítását LED-ek irányították: több különböző pozícióban elhelyezett LED körkörös felvillanását figyelve a csuklóját tanulta meg forgatni, továbbá a véletlenszerűen felvillanó LED megérintésével a fel-, alkar és az ujjak mozgását tudta fejleszteni. [3]

4 Single-Trial klasszifikáció Deep Learning segítségével

Relatíván nehéz lenne olyan területet említeni - legyen az pénzügyi, egészségügyi, autóipari stb. - , amelyen még tettek volna kísérletet, vagy alkalmaznák a Machine Learning (ML) technológiáját. A ML egy tág halmaza azon alkalmazásoknak – pl. Artificial Neural Network (ANN), Support Vector Machine (SVM) - , amelyek segítségével különböző adatok analízisét vagy eredmények becslését végzik. Múltja kb. az 1960-as évekig nyúlik vissza, de napjainkban érzékelhető térnyerése csak az elmúlt 15-20 évben történt. Ez annak is köszönhető, hogy szintén ugyanezen idő alatt nőtt ugrásszerűen a számítógépek számítási teljesítményének növekedése. A többnyire kétdimenziós konvolúciós neurális hálók (2D CNN) már bizonyították (főként képfeldolgozás terén), hogy kiváló teljesítményt képesek nyújtani nagy komplexitású feladatok megoldása során is. Az egydimenziós CNN-k használata még nem olyan mértékben elterjedt, mint a kétdimenziósaké, pedig bizonyos feladatok elvégzésére és bizonyos hardvereken előnyösebb lehet alkalmazásuk. Az 1D CNN-k további előnyei között szerepel architektúrájuk egyszerűsége, amely ezáltal kevesebb számítási kapacitást igényel, így lehetővé téve egyszerűbb készülékekbe – pl. mobiltelefon - integrálásukat is. Deep Learning (DL) architektúrák során a hálók rétegzésével válik a rendszer teljesítőképesebbé, még bonyolultabb feladatok megoldó képességévé. Az orvostudományban a ML és a DL modellek többnyire képfeldolgozási feladatokat látnak el diagnosztikai eljárások során. A Brain-Computer Interface rendszerek az agyi jelek rögzítésével és feldolgozásával segítik felhasználóik életminőségének javulását. A BCI-k kielégítő működéséhez a megfelelő jelek felismerése és feldolgozása kritikus feladat, amelyre a CNN-k jó megoldásként szolgálhatnak. A BCI rendszerek többnyire EEG vagy ECoG által rögzített jeleket használnak, amelyek feszültség-idő vagy feszültség-frekvencia térben analizálhatók ki.

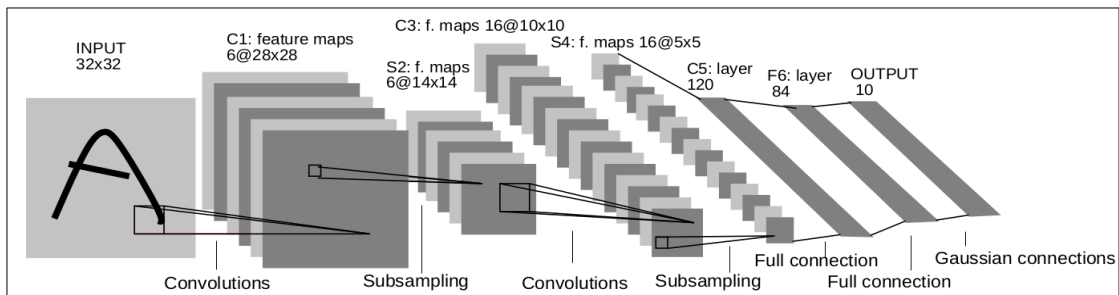
A következő fejezetekben először bemutatom a konvolúciós neurális hálók elméleti hátterének főbb részeit, majd a felhasznált adathalmaz felvételének körülményeit és felépítését. Ezután kísérletet teszek a nyers adatok feldolgozására kizárólag 1D CNN alkalmazásával és az eredmények alapján következtetést teszek további fejlesztési javaslatokra.

4.1 Konvolúciós neurális hálók

A legtöbb osztályozási feladat során a bemeneti adatot előre kell feldolgozni, szegmentálni és ún. feature-öket kell kisselektálni. A feature szelekció során az adatból azok a hasznos információk kerülnek kinyerésre, amely alapján képes a klasszifikáló algoritmus kategorizálni. A konvolúciós háló előnye, hogy a feature szelekciót és a klasszifikációt is elvégzi, továbbá nem szükséges a bemeneti adaton szegmentálást alkalmazni, így a modell lényegesen robusztusabbá és egyszerűbbé válik.

4.1.1 Rövid elmélet a konvolúciós neurális hálók működéséről

A neurális hálók alapegysége a neuron, amelyhez bemenetek, a bemenetein súlyok, kimenetén pedig egy ún. aktivációs függvény található. Egy réteget több neuron alkot, amelyekhez az előző és az adott réteg utáni réteg neuronjai csatlakoznak. Egy hálóban a fully-connected rétegek között mindegyik neuron csatlakozik mindegyikhez, azonban a konvolúciós hálók esetében ez nem igaz.



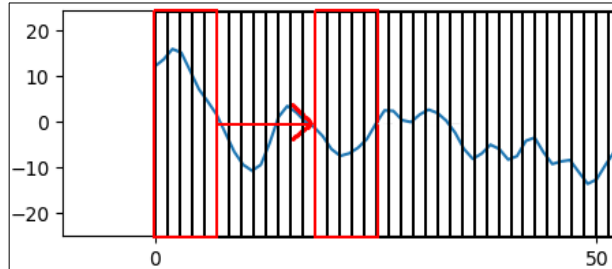
9. Abra: LeNet-5 architektúra, egy korai konvolúciós neurális háló modell példa [19]

Minden konvolúciós réteget feature map-ek alkotnak. A feature map-eket olyan pontok alkotják, amelyek az előző réteg feature map-jeinek és egy kernel – konvolúciós ablak - konvolúciója során állnak elő. Mindegyik feature map a bemenő adat egy feature-ét tartalmazza például élek egy képen vagy színek. A rétegek közötti kapcsolat a következő függvény írja le:

$$y_k^l = f(a) = f\left(b_k^l + \sum_{i=1}^{N_{l-1}} \text{conv1D}(w_{ik}^{l-1}, s_i^{l-1})\right) \quad (1)$$

, ahol y_k^l az l réteg kimenete, b_k^l az l réteg k -ik neuronjához tartozó bias érték, w_{ik}^{l-1} az $l-1$ réteg i -ik és az l réteg k -ik neuronok közötti kernel, s_i^{l-1} az $l-1$ réteg i -ik neuronjának kimenete, conv1D a konvolúciós művelet és $f()$ egy nem-lineáris függvény. [18] A háló mélyítésével a rétegek száma és azzal együtt a feature

map-ek száma is növelhető, amely során egyre bonyolultabb és összetettebb információ nyerhető ki a adatból. A map-ek száma és a kernel mérete jelentősen befolyásolja a háló helyes működését.



10. Abra: Konvolúció szemléltetése 1 dimenziós EEG adaton. A bemeneti adatvektor a standardizált feszültség értékeket tartalmazza, egy rácsegység 1 adatpontnak felel meg. A piros ablak a kernel, amely az adaton végighaladva végzi el a konvolúciós műveletet.

4.1.2 Aktivációs függvények

Az $f()$ egy ún. aktivációs függvény, amely azt a nem-linearitást valósítja meg a neurális hálóknban és ezáltal válnak a modellek képessé bonyolult adatok feldolgozására és komplex problémák megoldására. Az agyi neuronok működéséhez hasonló viselkedés a neurális hálóknban az aktivációs függvényekkel valósul meg. A agy neuronjainak aktivációja a sejt bemenetére érkező ingerek mértékétől függ; amennyiben az ingerlés mértéke elér egy küszöbértéket a neuron aktivál és kiad magából egy impulzust. A gyakorlatban aktivációs függvényként a leggyakrabban a Rectified Linear Unit (ReLU) vagy annak valamely módosított változata, a Softmax és a Sigmoid kerül alkalmazásra, de léteznek egyéb választások is. A ReLU többnyire jó kiinduló választás egy háló architektúrájának összeállításakor, a függvény a következő:

$$f(x) = \begin{cases} 0, & \text{ha } x \leq 0 \\ x, & \text{ha } x > 0 \end{cases} \quad (2)$$

, kimenetének tartománya $[0, \infty)$. Látható, hogy negatív bemeneti tartományra nem ad kimenetet a függvény. A LeakyReLU finomítja a ReLU-t és az $[0, \infty)$ bemeneteket $\alpha \cdot x$ kimeneti értékekké transzformálja át. α értéke általában 0.01 körül mozog. A neurális hálók kimenetei azok az értékek, amelyek azt adják meg, hogy a bemeneti adat mekkora valószínűséggel tartozik az előre meghatározott osztályokba. A valószínűségi értékek kifejezése 0 és 1 között történik, amelyek reprezentációjára többnyire Softmax-et használnak. A Softmax függvény, amely kimeneti értéke $(0, 1)$ tartományban helyezkedik el, a következő:

$$f_i(x) = \frac{e^{x_i}}{\sum_{j=1}^J e^{x_j}}, i=1,2,\dots,J \quad (3)$$

4.1.3 Gradiens-alapú tanulás

A hálók tanítása gyakran felügyelt módon, gradiens-alapú-backpropagation (BP) folyamaton keresztül történik. A háló tanulása során minden bemenethez prediktál egy kimeneti értéket. A prediktált és a valós kimeneti érték közötti különbséget a veszteségi függvény adja meg. Backpropagation során a cél azoknak a súlyparamétereknek a megtalálása, amellyel a háló vesztesége minimális. A veszteség minimalizálásának egy gyors és hatékony módja a veszteségi függvény súlyok szerinti gradiensének képzése. BP során a gradiensképzés a háló kimenetétől indulva halad a bemenete fele a következőképpen:

$$\begin{aligned} \frac{\partial E^p}{\partial W_n} &= \frac{\partial F_n}{\partial W} (W_n, X_{n-1}) \frac{\partial E^p}{\partial X_n} \\ \frac{\partial E^p}{\partial X_{n-1}} &= \frac{\partial F_n}{\partial X} (W_n, X_{n-1}) \frac{\partial E^p}{\partial X_n} \end{aligned} \quad (4)$$

, ahol az első egyenlet a $E^p(W)$ veszteségi függvény gradiensét adja meg, a második egyenlet pedig a parciális gradienst propagálja visszafele. [18][19] A gradiens minél pontosabb és gyorsabb számításához különböző optimalizáló algoritmusokat alkalmaznak. A leggyakoribb algoritmusok a stochastic gradient descent (SGD), RMSprop, ADAM és az Adagrad.

4.2 P300 detekció 1D CNN-el

P300 detekcióra a szakirodalomban számos olyan példa van, amely valamilyen ML modellt alkalmaz. Azonban a modellek többsége a legjobb tudomásom szerint SVM-re, Bayesian klasszifikátorra vagy 2D CNN-re épül, ezért a következőkben kísérletet teszek P300 jel klasszifikációjára 1D CNN alkalmazásával. A háló a bemeneti adatot nyers állapotában kapja meg, nem kerül elé feature detektor pl.: ICA vagy PCA. A modell kialakítása során kétfajta architektúra megvalósítására, összehasonlítására kerül sor és kiértékelésére kerül sor. A hyper-paraméterek finomhangolását random

search segítségével végeztem, majd az eredményeket a háló predikciójának pontossága alapján hasonlítottam össze.

4.2.1 Adathalmaz

A P300 jel klasszifikációhoz a BCI Competition 3 2-es szettje szolgál tanító és teszt adathalmazként. A két alanytól származó adathalmaz rögzítése BCI2000 általánosan használható BCI rendszerrel történt. Mindkét alany esetében 64 elektródával, azaz 64 csatornán rögzítették az agyi aktivitást. A Farwell és Donchin - féle betűzőhöz hasonló eszköz egy 6x6-os mátrix, amelynek soraiban és oszlopaiban az ABC betűi, továbbá 0 és 9 közötti számok szerepeltek, illetve egy alsóvonalas (_) karakter kapott helyet. A sorok/oszlopok véletlenszerű felvillantatása 5.7 Hz-es gyakorisággal történt. Az sorok/oszlopok világítása 100 ms-ig tartott, majd a következő ingerig eltelt szünet 75 ms-ig (ekkor a mátrixon nem világított egyik sor/oszlop sem). Az összesen 12 sor és oszlopok felvillantása egy karakter felvételi cikluson belül 15-ször került ismétlésre, így egy karakterpróba alatt összesen 180 stimulust érzékeltek az alanyok. A jelek 0.1 és 60 Hz között sávszűrésre , 240 Hz-en digitalizálásra kerültek.

Tanító adathoz 85 karakter felvételére került sor, amíg a teszt adat 100 karaktert tartalmaz. Az adatok *.mat formátumban kerültek elmentésre, előre felcímkézve. A struktúra tartalmazza a karaktereket, amelyre az alanynak koncentrálnia kellett, annak a sornak/oszlopok a sorszámát, amelyik felvillant, a villanáshoz tartozó címkét és azt, hogy éppen világít vagy sötét a panel. [15]

A jel zajossága miatt az adatot mielőtt a háló bemenetére került, Butterworth sávszűrővel szűrtem meg 0.3 és 30Hz között. A szűrés után standardizáltam is az adatot, mivel a neurális hálók teljesítményét jelentősen javíthatja a bemeneti adatok egységnyivé és null középpontosítottá tétele.

A következőkben pár olyan számítást írok le, amely az architektúra kialakításánál fontos szerepet játszott. Az adat 240 Hz-en került digitalizálásra, azaz az adatpontok között 4,17 ms telik el. Az ISI idő 175 ms, amely 42, az ITI 6*175 ms, ami 252 adatpontot jelent. A P300 jel többnyire 200-250 ms késleltetéssel jelenik meg és kb. 500 ms-ig tart, azonban előfordulhat, hogy egy-egy jelnek nagyobb a késleltetése. A P300 jel időtartama kb. 250-300 ms, amelyet 60-72 adatpont fed le.

4.2.2 Architektúra

Két fő architektúra típus kialakításával próbálkoztam meg. Az egyik egy általánosabb egybemenetű, amíg a másik több bemenettel rendelkezik. Mindegyik modell legalább x konvolúciós, 1 maxpool, újra x konvolúciós, 1 global-averagepool, 1 dropout és 1 fully-connected réteget tartalmaz. A rétegek és hyperparaméterek meghatározására random search algoritmust alkalmaztam.

A konvolúciós rétegek változtatásánál a várt viselkedése a modellnek, hogy minél mélyebb, azaz minél több réteg alkotja, annál pontosabbá válik, mivel annál több paraméterrel rendelkezik. Ennek azonban lehet hátulütője: a több paraméter több ideig tartó tanítást is jelent, továbbá egy bizonyos mérték után a háló túltanul (szakirodalomban idegen nyelven overfit). A túltanulás problémája azt jelenti, hogy a modell a túlságosan is a tanító adathoz alkalmazkodik és képtelenné válik új, addig még nem látott adat helyes feldolgozására. A túltanulás megakadályozható a háló sekélyebbé tételével, illetve pooling és dropout rétegek alkalmazásával is.

A kernel méretek megállapításánál a hipotézisem a következő volt: a kernelek segítségével a konvolúciós háló lokális mintákat keres. Azonban az agyi jel szűrés után is relatíve zajos, változatos marad, ezért a túl kicsi mérettel rendelkező kernelekkel nem lesz képes a modell felismerni a P300 karakterisztikáját. Másik oldalról a túl nagy kernelek más jelkomponenseket is számításba vesznek, amelyek a P300-at elfedik. A kernel méretek ezért 3 adatpontos mérettől 40-ig kerültek megadásra, azaz a P300 jel lefedettsége így 5%-tól a 67%-ig változott a kernel által.

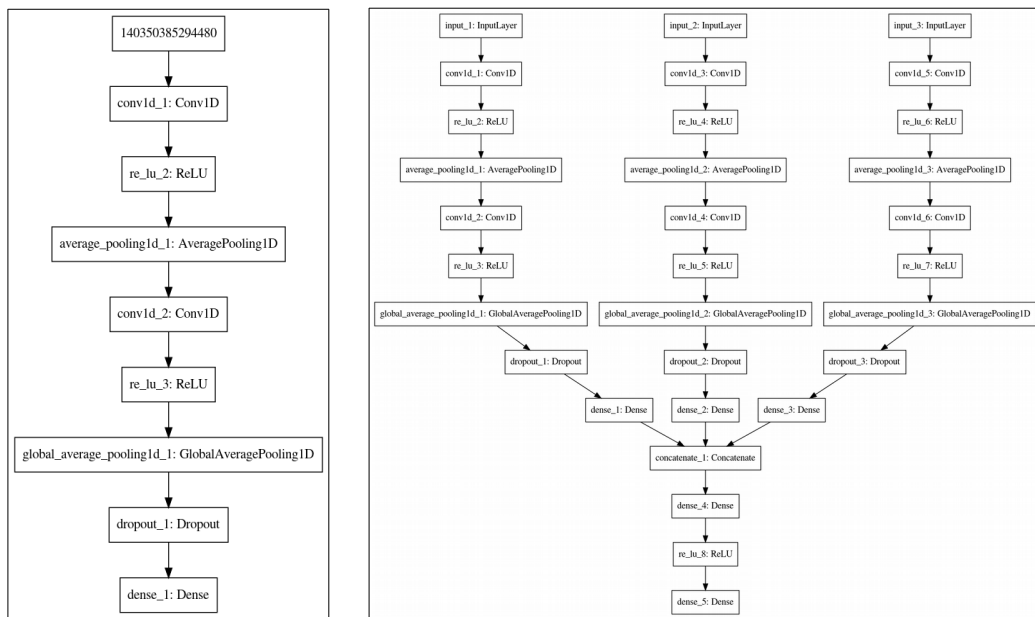
Pooling rétegek közül a maxpool és az averagepool rétegek kerültek összehasonlításra. Mindkét pooling célja, hogy minél toleránsabbá tegye a modellt az idővarianciával szemben. Túl nagy pooling méretek azonban, azt érhetik el, hogy a modell nem fogja figyelembe venni a jel időbeli karakterisztikáját, ezért nem képes klasszifikálni a jelet.

A dropout réteg alkalmazásának a célja, hogy a hálót egyrészt megakadályozza a túltanulástól és attól, hogy bizony súlyok nem kerülnek megfelelően kiszámításra. A tanulás folyamata alatt a dropout a paraméterként megadott százalékát kiejti a súlyoknak, ezzel ráerőltetve a modellre a nem kiejtett súlyok figyelembe vételét.

Aktivációs függvényeknek a ReLU-ra, a LeakyReLU-ra és az elu-ra esett választásom, mivel általában a ReLU-t a szakirodalom egy jó kiindulási választásnak számít. Több aktivációs függvény kipróbálása azért célszerű, mert előfordulhat, hogy valamelyik

függvény jobban passzol az adathoz. Az aktivációs függvényekhez hasonlóan kerültek kiválasztásra az optimalizáló algoritmusok (SGD, RMSprop, Adam) és a veszteségi függvények is (categorical crossentropy, mean squared error, mean squared logarithmic error, logcosh), amelyek teljesítményéről szintén a random search segítségével kapunk képet. (A veszteségi függvényeknél a logcosh-ra korábbi tapasztalat alapján került sor. Radar jelek feldolgozása során - amelyek szintén 1D-s adatok – az egyik legjobban teljesítő veszteségi függvényként szerepelt.)

A két fő architektúrából az egyik egy egybemenetű modell. Az első architektúra számára Cx210-es mátrixok voltak a bemeneti adatok, ahol C a csatornák, azaz az elektródák számát jelöli. A második architektúra egy pipeline, vagy másképp shared-vision modell. A modell C darab bemenetére, C darab 1x210-es vektor kerül, majd a csatornákhöz tartozó külön hálók kimenetei egy közös fully-connected rétegben futnak össze.



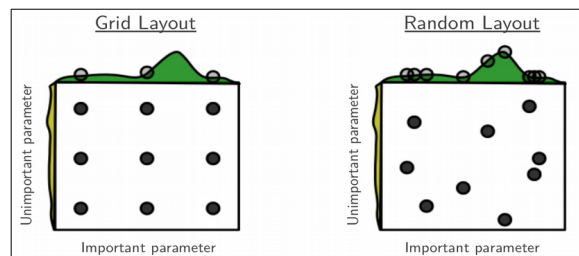
11. Abra: Egy- és több-bemenetű architektúra

A két architektúra közti fő különbség az, hogy az egybemenetű Cx210-es bemenettel táplált háló a csatornák között is keres összefüggést, illetve mintákat, amíg a több bemenettel rendelkező modell a fully-connected rétegig nem. Az elektródák az agy frontális lebenyi (Fz) területtől egészen a nyakszirti területig (Oz) lefedik az agyat. A P300 jel a 2.3.2-ben leírtak szerint kis késleltetéssel jelenik meg a különböző elektródák területén. A különböző elektródákon eltérő időpontokban megjelenő P300 miatt a több-

bemenetű architektúrának előnye lehet az egybemenetűvel szemben. Az előny abból származhat, hogy a háló elektródaspecifikusan tanul a mintákra, majd a fully-connected réteg területén dönt a csatornák hasznosságáról és az eltérő al-háló predikciói alapján határozza meg a végleges eredményt.

4.2.3 Hyperparameter tuning

Egy neurális háló megfelelő működéséhez rengeteg paraméter finomhangolása szükséges. Pár példa a paraméterekre: használt aktivációs és veszteségi függvény, optimalizáló algoritmus, rétegek száma, pooling méret, dropout mérték, learning-rate (a minimális hiba – veszteség - keresése közbeni lépések nagysága), az optimalizáló algoritmus paraméterei (pl. momentum érték), kernel méret és az egy konvolúciós rétegen belüli feature map-ek száma. Ezek a paraméterek mind hyperparaméterek. Látható, hogy már a példák listája is nagy mennyiségű összeállítási lehetőséget jelent. Az a folyamat, amely során az optimális hyperparaméter összeállításának megtalálása a cél, a hyperparameter tuning. A tuning leggyakrabban előforduló két módja a grid search és a random search. A grid search során minden paraméter egy előre meghatározott érték, amelyeknek a variációján halad végig az algoritmus a tuning során.

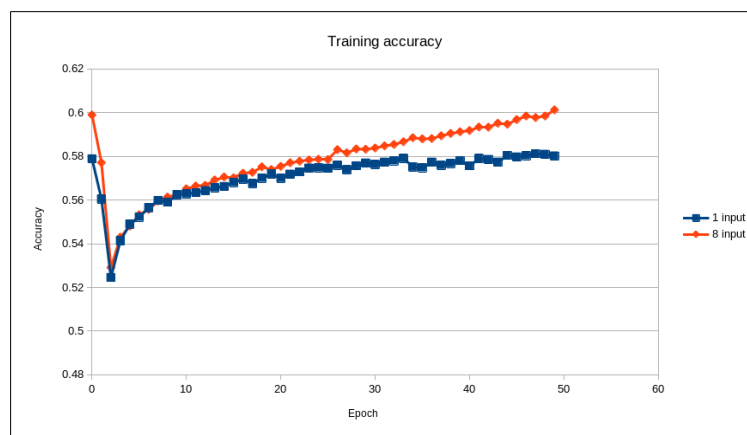


12. Abra: Grid search és Random search [10]

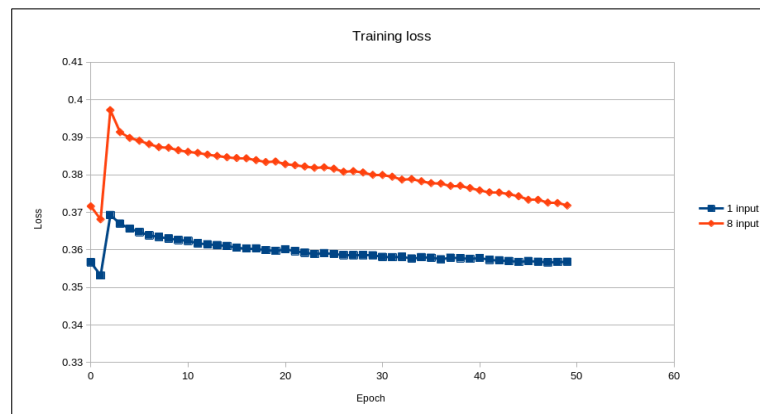
Ezzel szemben a random search esetében, nem konkrét értékek, hanem intervallumok (konfigurációs terek) kerülnek megadásra, amelyeken belül az algoritmus véletlenszerűen választja ki a paramétereket. A grid search előnyei közé tartozik, hogy a tanulási folyamatba jobban bele lehet látni, illetve kisebb méretű adathalmazokon jobban is tud teljesíteni, mint a random search. Azonban, ahogyan nő a keresendő paraméterek száma és a használt bemeneti adatok halmazának mérete, a random search válik hatékonyabbá. További előnye a random searchnek, hogyha a folyamat valamilyen hiba folytán megszakad, bármikor előről kezdhető, minimális veszteséggel. [10]

4.2.4 Eredmények

A két architektúra tanítása ugyanazon a módon történt. Mindkét architektúra modelljei 50 epoch-ig lettek tanítva, véletlenszerűen generált paraméterekkel. Összesen 30-30 különböző hyper-paraméter összeállítással rendelkező modell került tanításra és tesztelésre. Az eredményeket a tanítás során egy callback függvény mentette ki *.csv formátumú fájlalba. Két ábrán mutatom be a modellek pontosságainak és veszteségeinek alakulásait. Az ábrák ordinátatengelyén a pontosság, illetve a veszteség látható, az abszcisszán pedig az befejezett epoch-ok száma. A függvényeken a két architektúra 30-30 modelljének adatai láthatók, lefutott epoch-ok szerint átlagolva.



13. Abra: Átlagos pontosság alakulása a lefutott epoch-ok függvényében



14. Abra: Átlagos veszteség alakulása a lefutott epoch-ok függvényében

Az 1 bemenetű modellek pontossága és vesztesége is hamar szaturálni kezd. 40 epoch felett az látható, hogy a modell kb. 58%-os pontosságnál már nem tanul tovább. Ezzel szemben a pipeline architektúra pontossági és veszteségi függvényének meredekségében az látszik, hogy a háló 50 epoch-nál még nagy ütemben tanult. A különbséget, hogy a pipeline architektúra még tanult 50 epoch-nál is, a háló mélysége,

ezzel együtt pedig a paraméterek száma okozhatja. Kb. 1-2 nagyságrenddel több paramétert tartalmaznak a több-bemenetű modellek, mint az egybemenetűek. Azonban a pontosság folyamatos növekedése utalhat arra, hogy a pipeline típusú modellek jobban képeket tanulni, annak köszönhetően, hogy az elektródák elszeparáltan kerültek feldolgozásra.

A 4.2.2-ben említett kernel mérethez kapcsolódó észrevételek. A pontosság szempontjából legjobban teljesítő 30 modell 23%-nál szerepelt a 40-es, míg csak 8%-nál 3-as kernel méret. A nagyobb kernel méret tehát jobban passzolhat ehhez az adattípushoz, mint a nagyon kicsik. Érdeemes megjegyezni, hogy a két kernel méret között a többi mérethez (6, 12, 20, 24) tartozó előfordulások száma nem lineárisan nő. A 6-os és 24 ablakméret ritkábban fordult elő, mint a 12 és 20-as. Továbbá fontos azt is megjegyezni, hogy az előfordulások számát jelentősen befolyásolhatja a random search algoritmus is.

[11]-ben PCA alkalmazásával növelték egy egyszerű nem-lineáris regressziós modell pontosságát. Amíg a legjobb eredménnyel teljesítő PCA+NLR (non-linear regression) pontossága 63.05%, addig pipeline architektúráé 61.7%. [12]-ben nem ERP, hanem motor imagery jelek klasszifikációját végezték. A modell egy 1D CNN + SAE (stacked auto encoder) párosából áll össze. A BCI competition II 3-as adathalmazán 90%-os pontosságot tudott elérni a modell, amely kitűzött célként fog szolgálni a jövőbeli elérendő eredmények tekintetében.

Érdeemes megjegyezni, hogy több epoch-ig tartó tanítás lehetségesen már a jelenlegi paraméterekkel is növelné a modell pontosságát. Habár a jelenlegi architektúra pontossága még csak 61.7%, a jövőben ez az eredmény még tovább növelhető.

5 Konklúzió és jövőbeli tervek

A pipeline architektúra hordoz magában lehetőséget EEG jelek 1D CNN-el való feldolgozásához. A feladat során nem tettem kísérletet a bemeneti adat különösebb feldolgozására. Az adat előfeldolgozásával, például Fast Fourier transform, Short time Fourier transform (mint [12]-ben) alkalmazásával előfordulhat, hogy háló olyan feature-ök szelektálását tudná elvégezni, amely alapján jelentősen nőne a teljesítménye. A háló architektúrájába mélyebben belenyúlva, az optimalizáló algoritmusok és az aktivációs függvények paramétereinek állításával szintén pontosabbá lehetne tenni a predikciót. További ötlet a transfer learning alkalmazása, amely során a modell egy adott alany jeleire kerül betanításra, majd a betanított modell bizonyos számú rétege (esetleg egésze) egy másik alany agyjeleivel is tanításra kerül.

Irodalomjegyzék

- [1] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, Theresa M. Vaughan: *Brain–computer interfaces for communication and control*, Invited review – Elsevier, *Clinical Neurophysiology* 113, 2002, pp. 767–791
- [2] L.A. Farwell and E. Donchin: *Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials* - Elsevier Scientific Publishers Ireland, Ltd., *Electroencephalography and clinical Neurophysiology* 70, 1988, pp. 510-523
- [3] Alim Louis Benabid, Thomas Costecalde, Andrey Eliseyev, Guillaume Charvet, Alexandre Verney, Serpil Karakas et al.: *An exoskeleton controlled by an epidural wireless brain–machine interface in a tetraplegic patient: a proof-of-concept demonstration* - *Lancet Neurol* 2019
- [4] Elizabeth A. Felton, M.S., J. Adam Wilson, M.S., Justin C. Williams, PhD., P. Charles Garell, M.D.: *Electrocorticographically controlled brain–computer interfaces using motor and sensory imagery in patients with temporary subdural electrode implants*, Report of four cases - *J Neurosurg* 106, 2007, pp. 495–500
- [5] David Sussillo, Sergey D. Stavisky, Jonathan C. Kao, Stephen I Ryu, Krishna V. Shenoy: *Making brain-machine interfaces robust to future neural variability* - *Nature Communications* 7, 2016
- [6] M. Teplan: *Fundamentals of EEG measurement*, Measurement Science Review - Institute of Measurement Science, Slovak Academy of Sciences , Vol. 2, Section 2, 2002
- [7] Parthana Sarma, Prakash Tripathi , Manash Pratim Sarma, Kandarpa Kumar Sarma: *Pre-processing and Feature Extraction Techniques for EEG-BCI Applications*, A Review of Recent Research – Sarma, ADBU-Journal of Engineering Technology, ISSN: 2348-7305, Vol. 5, 2016, pp. 0051604(8pp)
- [8] Hubert Cecotti and Axel Gräser: *Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces* - *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 33, No. 3, 2011, pp. 433-445
- [9] Jessy Parokaran Varghese: *Analysis of EEG Signals For EEG-based Brain-Computer Interface* - School of Innovation, Design and Technology Mälardalen University, 2009
- [10] James Bergstra, Yoshua Bengio: *Random Search for Hyper-Parameter Optimization* - *Journal of Machine Learning Research*, 13, 2012, pp. 281-305

- [11] Nand Sharma: *Single-trial P300 Classification using PCA with LDA, QDA and Neural Network* - Department of Mathematics, Colorado State University, 2017
- [12] Yousef Rezaei Tabar, Ugur Halici: *A novel deep learning approach for classification of EEG motor imagery signals* - Journal of Neural Engineering, 14, 2016
- [13] John Polich: *Updating P300: An integrative theory of P3a and P3b*, Invited Review – Elsevier, Clinical Neurophysiology 118, 2007, pp. 2128–2148
- [14] Terence W. Picton: *The P300 Wave of the Human Even-Related Potential* – Journal of Clinical Neurophysiology 9(4), 1992, pp. 456-479
- [15] Wadsworth BCI Dataset (P300 Evoked Potentials): *Documentation* - BCI Competition III Challenge, 2004
- [16] Jeffrey S. Kreutzer, John DeLuca, Bruce Caplan: *Encyclopedia of Clinical Neuropsychology* – Springer, ISBN 978-0-387-79947-6
- [17] Touradj Ebrahimi: *Recent advances in Brain-Computer Interfaces* - IEEE 2007 MMSP Workshop, 2007
- [18] Serkan Kiranyaz, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, Daniel J. Inman: *1D Convolutional Neural Networks and Applications*, A survey - arXiv, Signal Processing, 2019
- [19] Yann LeCun, Patrick Haffner, Léon Bottou, Yoshua Bengio: *Object Recognition with Gradient-Based Learning* - Shape, Contour and Grouping in Computer Vision, 1999, pp. 319-345

Ábrák

- [20] DIY tDCS: *Where Do The Electrodes Go?* - <https://www.diytdcs.com/tag/1020-positioning/> , 2012
- [21] R.F. Simons, W.M. Perlstein: *A Tale of Two Reflexes: An ERP Analysis of Prepulse Inhibition and Orienting* – <http://rsimons.psych.udel.edu/simons.htm>, Figure 5, 2019