



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

NÁSTROJ PRO AUTOMATICKÉ HODNOCENÍ KVALITY EEG DAT

TOOLBOX FOR AUTOMATIC EEG DATA QUALITY ASSESSMENT

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Jan Meloun

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Martin Lamoš, Ph.D.

BRNO 2023

Diplomová práce

magisterský navazující studijní program **Bioinženýrství**

Ústav biomedicínského inženýrství

Student: Bc. Jan Meloun

ID: 186676

Ročník: 2

Akademický rok: 2022/23

NÁZEV TÉMATU:

Nástroj pro automatické hodnocení kvality EEG dat

POKYNY PRO VYPRACOVÁNÍ:

1) Proveďte literární rešerši zpracování dat skalpové elektroencefalografie (EEG), zaměřte se na hodnocení kvality naměřených záznamů. 2) Na základě literatury definujte minimálně 5 metrik hodnocení kvality dat. 3) Navrhněte koncept nástroje, který by umožňoval automaticky ohodnotit předložený EEG záznam pomocí definovaných metrik kvality. Nástroj by měl umožňovat i rozšiřitelnost o další metriky hodnocení kvality. 4) Implementujte takto navržený nástroj v programovém prostředí MATLAB, výstupem nástroje by měl být přehledný report o kvalitě předloženého záznamu. 5) Otestujte použitelnost nástroje (spolehlivost jednotlivých metrik, skórovacího algoritmu, výsledného reportu) na reálných EEG datech z Laboratoře multimodálního a funkčního zobrazování CEITEC MU. 6) Diskutujte dosažené výstupy jednotlivých bodů zadání.

DOPORUČENÁ LITERATURA:

- [1] MICHEL, Christoph M., et al. (ed.). Electrical neuroimaging. Cambridge University Press, 2009.
- [2] FICKLING, Shaun D., et al. Good data? The EEG quality index for automated assessment of signal quality. In: 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE, 2019. p. 0219-0229.

Termín zadání: 6.2.2023

Termín odevzdání: 22.5.2023

Vedoucí práce: Ing. Martin Lamoš, Ph.D.

doc. Ing. Radim Kolář, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato práce se zabývá návrhem nástroje pro automatické hodnocení kvality EEG dat. V teoretické části práce se nachází popis vzniku a šíření akčního potenciálu nervovou soustavou. Dále také teoretickým popisem EEG záznamu a jeho artefakty. Následuje popis používaných metod pro detekci artefaktů. V praktické části práce se nachází popis návrhu nástroje pro automatické hodnocení kvality EEG včetně diskuze výsledků na poskytnutých datech.

KLÍČOVÁ SLOVA

Elektroencefalografie, EEG artefakty, předzpracování dat, metriky kvality, automatické skórování, MATLAB, Fieldtrip toolbox

ABSTRACT

This thesis deals with the design of a tool for the automatic evaluation of EEG data quality. The theoretical part of the thesis contains a description of the formation and propagation of the action potential through the nervous system. Furthermore, a theoretical description of the EEG recording and its artifacts. The following is a description of the methods used to detect artifacts. In the practical part of the thesis, there is a description of the design of the tool for automatic EEG quality assessment, including a discussion of the results based on the provided data.

KEYWORDS

Electroencephalography, EEG artifacts, data preprocessing, quality metrics, automatic scoring, MATLAB, Fieldtrip toolbox

MELOUN, Jan. *Nástroj pro automatické hodnocení kvality EEG dat*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství, 2023, 52 s. Diplomová práce. Vedoucí práce: Ing. Martin Lamoš, Ph.D.

Prohlášení autora o původnosti díla

Jméno a příjmení autora: Bc. Jan Meloun
VUT ID autora: 186676
Typ práce: Diplomová práce
Akademický rok: 2022/23
Téma závěrečné práce: Nástroj pro automatické hodnocení kvality EEG dat

Prohlašuji, že svou závěrečnou práci jsem vypracoval samostatně pod vedením vedoucí/ho závěrečné práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

podpis autora*

* Autor podepisuje pouze v tištěné verzi.

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Martinu Lamošovi, Ph.D. za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Obsah

Úvod	11
1 Skalповý záznam elektroencefalografie	12
1.1 Vznik a šíření potenciálu	12
1.1.1 Morfologie neuronu	12
1.1.2 Rytmy a vlny v elektroencefalografických datech	14
1.2 Přístrojové řešení	16
1.2.1 Elektrody	16
1.2.2 Montáže	17
1.2.3 Další komponenty přístroje	19
2 Zpracování elektroencefalografických dat	21
2.1 Artefakty a rušení	21
2.1.1 Artefakty a rušení biologického původu	22
2.1.2 Artefakty a rušení nebiologického původu	22
2.2 Předzpracování dat	22
2.2.1 Filtrace	23
2.2.2 Analýza nezávislých komponent	24
2.2.3 Interpolace	24
3 Kvalita elektroencefalografických dat	25
3.1 Analýza elektroencefalografického signálu	25
3.1.1 Analýza signálu v časové oblasti	25
3.1.2 Analýza signálu ve frekvenční oblasti	25
3.1.3 Analýza signálu v časově-frekvenční oblasti	26
3.1.4 Analýza signálu v prostorové oblasti	26
3.2 Extrakce příznaků	27
3.2.1 Příznaky složitosti	27
3.2.2 Příznaky kontinuity	27
3.2.3 Příznaky konektivity	28
3.3 Metody detekce odlehlých hodnot	29
3.3.1 Statistické metody	29
3.3.2 Metody založené na učení	30
4 Návrh nástroje a jeho programové řešení	31
4.1 Načtení elektroencefalografických dat	32
4.1.1 Použité datové formáty	32
4.2 Přítomnost značek a elektrokardiografického signálu	33

4.3	Předzpracování dat	33
4.4	Výpočet příznaků	34
4.4.1	Statičnost signálu	34
4.4.2	Koeficient špičatosti	34
4.4.3	Síťové rušení	35
4.4.4	Maximální gradient	36
4.4.5	Amplituda kvadratického průměru	36
4.4.6	Síla v užitečném pásmu	36
4.4.7	Míra průchodu nulou	36
4.4.8	Rozšířitelnost o výpočet dalších příznaků	37
4.5	Vyhodnocení příznaků	37
4.6	Detekce špatných úseků signálů a kanálů	37
4.7	Report o kvalitě dat	38
5	Vyhodnocení detekce artefaktů	39
5.1	Detekce artefaktových úseků signálu	40
5.2	Detekce artefaktových kanálů	43
	Závěr	44
	Literatura	45
	Seznam zkratek	49
	Seznam příloh	51
	A Obsah elektronické přílohy	52

Seznam obrázků

1.1	Morfologie neuronu	13
1.2	Základní EEG vlny	15
1.3	Rozložení elektrod systému 10-20	17
1.4	Rozložení elektrod systému využívajícího 256 elektrod	18
1.5	Schéma zapojení bipolární montáže	18
1.6	Schéma zapojení referenční montáže	19
1.7	Schéma zapojení průměrovací montáže	19
2.1	Obecné blokové schéma zpracování EEG signálu	21
2.2	Obecné blokové schéma předzpracování EEG signálu	23
3.1	Harmonický signál a jeho amplitudové spektrum	26
4.1	Blokové schéma navrženého nástroje	31
4.2	Blokové schéma extrakce příznaků a statistických veličin	35

Seznam tabulek

3.1	Tabulka se seznamem běžně používaných příznaků složitosti	27
3.2	Tabulka se seznamem běžně používaných příznaků kontinuity	28
3.3	Tabulka se seznamem běžně používaných příznaků konektivity	29
4.1	Tabulka podporovaných datových formátů podle výrobců	32
5.1	Tabulka s veškerými poskytnutými soubory a jejich vlastnostmi	39
5.2	Tabulka unikátních značek, jejich četností a kanálu EKG	40
5.3	Tabulka s hodnotami z matice záměn a výsledků metrik úspěšnosti	42
5.4	Tabulka anotovaných a detekovaných kanálů a hodnot poměru	43

Úvod

Skalповá elektroencefalografie je neinvazivní neurozobrazovací technika, která zaznamenává elektrickou aktivitu z povrchu mozku. Ukázalo se, že je cenným nástrojem v různých oblastech neurověd, včetně kognitivní psychologie a klinické diagnostiky. Kvalita záznamů EEG však hraje zásadní roli při zajištění spolehlivé a přesné analýzy a interpretace zaznamenaných dat. Proto je nezbytné vyvinout účinné metody pro hodnocení kvality naměřených EEG signálů. Tato práce si klade za cíl řešit problém hodnocení kvality dat EEG provedením komplexního přehledu literatury o zpracování dat EEG na skalpu, se zvláštním zaměřením na hodnocení kvality naměřených záznamů.

Na základě poznatků z rešerše literatury bude tato práce definovat alespoň pět metrik hodnocení kvality dat. Tyto metriky budou sloužit jako kvantitativní měřítka pro hodnocení spolehlivosti a věrnosti EEG záznamů, zachycující různé aspekty, jako je poměr signálu k šumu, impedance elektrody, spektrální charakteristiky a další relevantní parametry. Výběr těchto metrik se bude řídit jejich zavedeným významem v oboru a jejich potenciálem poskytnout komplexní pohled na kvalitu EEG dat.

Dále tato práce navrhne koncept nástroje, který umožní automatické vyhodnocení zaslaných EEG záznamů pomocí definovaných ukazatelů kvality. Nástroj bude navržen tak, aby byl flexibilní a rozšiřitelný, což umožní snadnou integraci dalších metrik hodnocení kvality. Implementace nástroje bude provedena v programovacím prostředí MATLAB s využitím volně dostupného Fieldtrip toolboxu.

K ověření účinnosti a použitelnosti vyvinutého nástroje budou využita skutečná EEG data z laboratoře CEITEC MU Multimodálního a funkčního zobrazovací. Testování odhalí účinnost detekce artefaktových kánlů či úseků signálu. Výsledky těchto testů poskytnou cenné poznatky o výkonu nástroje a jeho potenciálních nedostacích.

Na závěr si tato práce klade za cíl přispět do oblasti zpracování EEG dat skalpu poskytnutím komplexního přehledu metod hodnocení kvality, definováním nových metrik kvality dat a navržením a implementací nástroje pro automatické vyhodnocování EEG záznamů. Výsledky výzkumu mají potenciál zvýšit spolehlivost a validitu analýzy dat EEG a usnadnit pokrok v různých oblastech neurovědního výzkumu a aplikací.

1 Skalповý záznam elektroencefalografie

Elektroencefalografie (EEG) je elektrofyziologická metoda záznamu elektrické aktivity mozku. Z průběhu naměřeného EEG signálu lze usuzovat funkční stav jednotlivých oddílů mozku, proto tato metoda patří mezi základní diagnostické metody v neurologii a psychiatrii. [29], [31]

Z hlediska záznamu dat se zpravidla jedná o neinvazivní metodu, kdy nepolarizovatelné elektrody jsou přikládány na pokožku hlavy. Méně často je metoda invazivní, kdy jsou použity podkožní vpichové, nasofaryngeální či intercerebrální elektrody. Z pohledu místa snímání signálu rozlišujeme skalповý nebo intrakraniální záznam EEG. U prvně zmíněného záznamu jsou malé elektrody umístěny na pokožce hlavy. Zde je nutný dobrý mechanický a elektrický kontakt. Pro intrakraniální záznam EEG jsou speciální elektrody implantované do mozku. Zvláštním případem měření EEG je metoda elektrokortikogram (ECoG), při němž je měření prováděno přímo z kortikálního povrchu pomocí subdurálních elektrod. V této práci je využit skalповý záznam EEG. [22], [31]

1.1 Vznik a šíření potenciálu

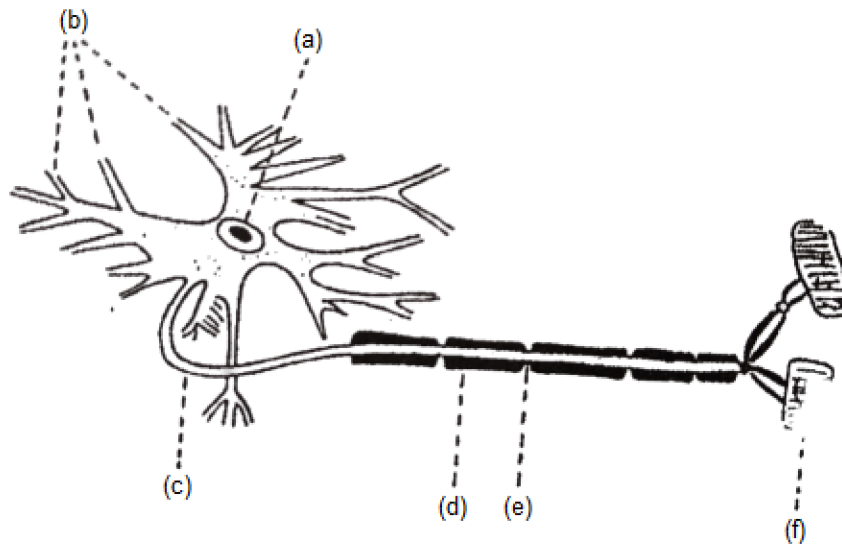
Pro pochopení EEG dat je nutné vysvětlit vznik a šíření akčního potenciálu (AP) centrální nervovou soustavou člověka (CNS). CNS je tvořena mozkem (*cerebrum*) a míchou (*medulla spinalis*), kde mozek se dále dělí na tři hlavní části: koncový mozek (*telencephalon*), mozeček (*cerebellum*) a mozkový kmen (*truncus cerebri*). Nejvýznamnějším oddílem CNS je mozková kůra (*cortex cerebri*) tvořící plášť koncového mozku rozčleněný na závitý výrazným způsobem zvětšující funkční plochu. Zde rozlišujeme levou a pravou hemisféru oddělenou centrální rýhou. Základní dělicí jednotkou je lalok (*lobus*), dále se skládající ze závitů (*gyrus*) od sebe oddělenými pomocí zářezů (*sulcus*). [7], [29], [31]

Nervový systém se skládá z neuronů a neuroglií - podpůrných buněk, které mají velký rozsah funkcí. Například tvoří podporu neuronální sítě, zajišťují výživu neuronů, provádějí fagocytózu a fungují jako izolanty tvorbou myelinu. Stavbě a funkci neuronu jakožto zdroji EEG signálu je věnovaná následující kapitola. [7], [29]

1.1.1 Morfologie neuronu

Neuron znázorněný na obrázku 1.1 tvořící základní stavební jednotkou CNS se skládá z buněčného těla, dendritů a axonu. Uvnitř buněčného těla (*soma*) se nachází jediné jádro a probíhá zde většina buněčného metabolismu, obzvláště toho spjatého se

syntézou proteinů. Druhou význačnou částí neuronu jsou krátké a velmi bohatě větvené výběžky - dendrity, ze kterých vystupují dendritické trny účastnící se synapsí. Jako synapse je označované spojení mezi dvěma neurony nebo neuronem a smyslovou buňkou za účelem předávání vzruchů. Dendrity jsou spojeny buď s axony nebo jinými dendrity okolních neuronů. Na membránách dendritů vzniká podráždění - odstupňovaná místní elektrická odpověď. Šíření probíhá přes membránu buněčného těla až k iniciálnímu segmentu. Válcovitý výběžek neuronu nazývaný axon (*neurit*) je zpravidla obalen myelinovými pochvami, mezi kterými se nachází Ranvierovy zářezy. Látka myelin je produkována Schwannovými buňkami, což jsou jedny z dříve zmíněných neuroglií. Pochvy mají specifickou strukturu a stavbu významnou pro přenos informací, čímž se rozumí přenos nervových vzruchů z těla neuronu. Taktéž důležitý pro přenos informací je axonový hrbol, protože u myelinizovaných axonů přechází v iniciální segment. Zde je místní elektrická odpověď při dosažení prahových hodnot převáděna na AP. [7], [23], [31]



Obr. 1.1: Morfologie neuronu: (a) - jádro, (b) - dendrity, (c) - axon, (d) - myelinová pochva, (e) - Ranvierův zářez, (f) - synapse. [34]

V CNS je velké množství neuronů propojeno do složité prostorové sítě, kde AP probíhá po nervovém vlákne a šíří se od těla jednoho neuronu k jeho periférii (centrifugálně). Rychlost šíření je závislá na vlastnostech nervového vlákna. Šíření místních elektrických proudů dráždící sousední úsek axonu je podstatou vedení vzruchů. Šíření vzruchu je vždy jednosměrné a bez dekrementu. Díky přítomnosti myelinu se rychlost šíření zvyšuje díky saltatornímu vedení, kdy AP neprobíhá po celém vlákne neuronu, ale přeskakuje mezi Ranvierovými zářezy. Takto vzruch doputuje skrze axon až ke spojení konce axonu s dalším neuronem - synapsi. Základním typem synapse je synapse chemická, v níž z presynaptické membrány jsou uvolněny mediátory do synap-

tické štěrby. Množství uvolněného mediátoru je přímo úměrné vzruchové aktivitě. Z prostoru štěrby jsou mediátory navázány na postsynaptickou membránu u níž mohou způsobit hyperpolarizaci nebo depolarizaci. Hyperpolarizace označuje změnu negativního membránového napětí na více negativní. Zatímco depolarizace označuje změnu membránového napětí na méně negativní, v některých případech až na pozitivní. Běžně se hodnota membránového napětí pohybuje okolo -70 mV . Excitační postsynaptický potenciál (EPSP) označuje depolarizaci postsynaptické membrány vlivem navázání mediátorů. Dostatečné množství stimulů se projeví vybavením akčního potenciálu na axonovém hrbolku. Inhibiční postsynaptický potenciál (IPSP) je zapříčiněn inhibicí z důvodu hyperpolarizace postsynaptické membrány. [23]

1.1.2 Rytmus a vlny v elektroencefalografických datech

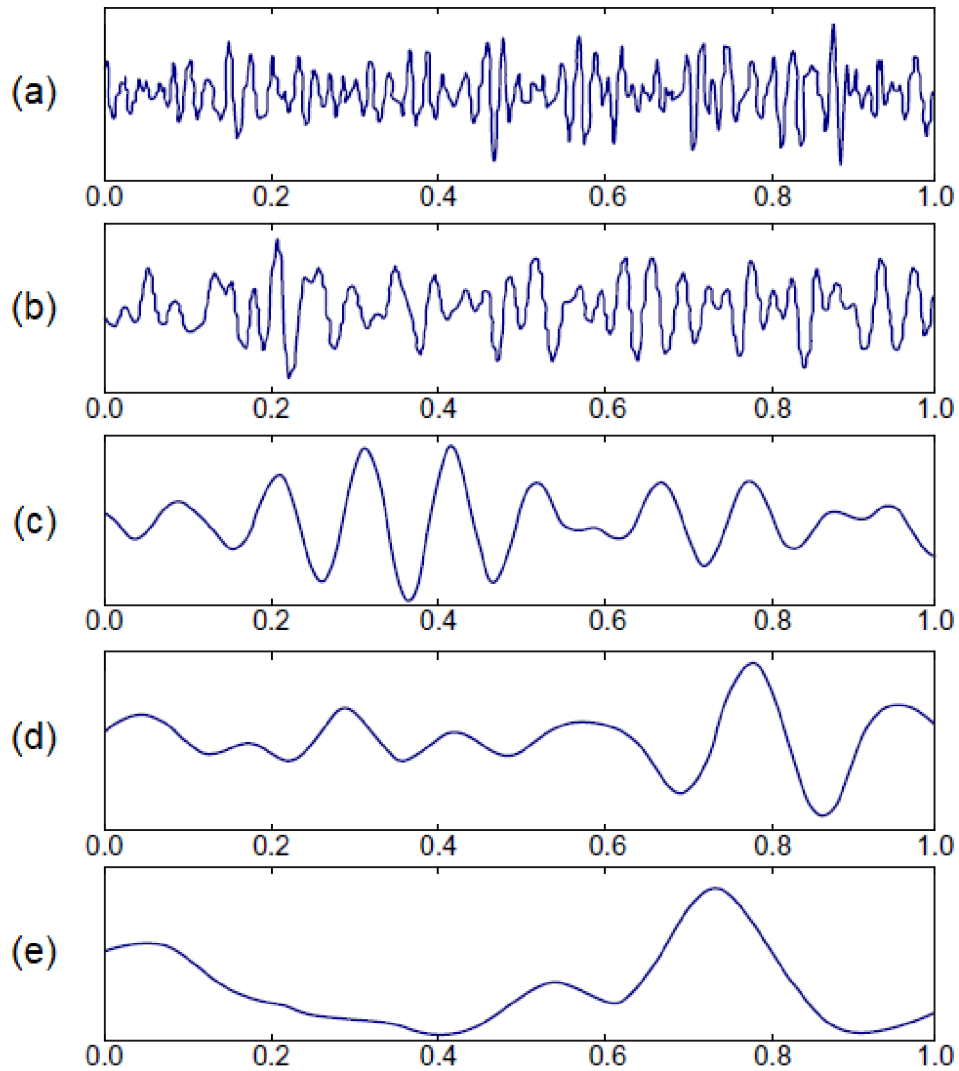
Jako rytmus je označována elektrická aktivita mozkové kůry pro její repetitivní oscilující chování. Pro popis rytmů se využívá jejich relativní amplituda a rozsah frekvencí. Amplituda je dána mírou synchronizace, s níž neurony kůry interagují. Vysoká amplituda je výsledkem sumace signálů ze součinných neuronů v závislosti na čase. Pokud dojde k opakování výše zmíněných synchronních excitací, vzniká tak rytmický EEG signál obsahující rytmy s vysokou amplitudou. Zatímco asynchronní excitace produkuje nepravidelně vypadající EEG s nízkou amplitudou signálu. [31]

Pro skalpový záznam EEG dosahuje amplituda až do $100\ \mu\text{V}$ a frekvence se pohybuje od $0,5$ do $40\ \text{Hz}$. Podle behaviorálního stavu se v tomto frekvenčním rozsahu nachází pět rozlišitelných specifických frekvenčních pásem znázorněných na obrázku 1.2. Interpretace se odvíjí od věku, stavu vědomí, medikace a podmínek provádění měření. [21], [31]

Nejnižší frekvenci dosahuje δ rytmus. Ten nikdy nepřesáhne hodnotu $4\ \text{Hz}$ a má poměrně vysokou amplitudou klesající při depresivních stavech. Pro dospělý subjekt je jeho přítomnost typická a důležitá při hlubokém spánku, zatímco u bdění je patologická a svědčí o poškození mozku. U subjektů v kojeneckém věku je v bdělém stavu běžná. [19], [21]

V pásmu od 4 do $7\ \text{Hz}$ se nachází θ rytmus s amplitudou do $15\ \mu\text{V}$ zvyšující se při usínání. U dospělých subjektů se vyskytuje při bdělém stavu a otevřením očí dochází k útlumu rytmu. Někdy může chybět úplně, někdy může být superponován na α rytmus. Pro podezření z patologického procesu na mozku by asymetrie mezi hemisférami neměla přesáhnout $50\ \%$. [19], [21]

α rytmus je vymezen rozsahem frekvencí mezi 8 a $13\ \text{Hz}$ a amplitudou od 20 do $80\ \mu\text{V}$. U zdravých dospělých subjektů se projevuje v klidu, bdění a zavřených očích. α rytmus vymizí při otevření očí a po jejich uzavření a zklidnění subjektu se opět objeví. K vymizení může taktéž dojít z důvodu ospalosti nebo spánku vyšetřovaného



Obr. 1.2: Základní EEG vlny: (a) - γ , (b) - β , (c) - α , (d) - θ a (e) - δ . [2]

subjektu. Obecně platí, že frekvence by měla být konstantní v rámci měření a též by neměla být rozdílná mezi hemisférami. Avšak s rostoucím věkem může klesnout. Taktéž lze usuzovat, že α rytmus svým způsobem reflektuje perfuzi mozku, kdy s klesající frekvencí α rytmu klesá i prokrvení mozku. [19], [21], [31]

Pro β rytmus je typický rozsah frekvencí 18 až 25 Hz a amplituda jež není vyšší než 20 μV . Tento rytmus je měřitelný v celé snímané oblasti hlavy, avšak je nejvíce zřetelný nad frontálními laloky v přední části skalpu a temeni hlavy. β rytmus se běžně vyskytuje při normální až zvýšené pozornosti v bdělém stavu. Významná je symetrie mezi hemisférami, kdy nesymetrie značí patologii. S rostoucím věkem roste i zastoupení β rytmu a obecně je vyšší u žen než u mužů. Krátkodobý vzrůst β rytmu je pozorovatelný při usínání. [19], [31]

V rámci EEG má γ rytmus frekvence vyšší než 30 Hz , kdy v literatuře se horní

hranice liší, často se uvádí rozsah až do 100 Hz. Amplituda se pohybuje od 3 do 5 μV . γ rytmus vychází z mozkové kůry a hrbolu mezimozkového (*thalamus*) a je odrazem aktivního zpracování informace kůrou za účelem kognitivní nebo motorické funkce. [19], [29], [31]

Krom výše zmíněných pravidelných rytmů se mohou náhodně vyskytovat strmé vlny nebo specifické hroty s nepravidelností a nepředvídatelností. Výskyt těchto útvarů je typický pro subjekty se špatnou funkcí neuronů, často trpícími epilepsií nebo jinými mozkovými chorobami. [19]

1.2 Přístrojové řešení

Elektrická aktivita mozkové kůry je zaznamenávána jako EEG signál, kdy její hodnota je v řádu mikrovoltů (μV) a proto je nutné ji zesílit pro zobrazení. Většina signálu pochází z postsynaptických proudů pyramidových neuronů v šedé hmotě, které se propagují na kůži hlavy. Aktivace velkého množství neuronů může generovat dostatečně velký potenciál zaznamenanatelný s pomocí povrchových elektrod na hlavě. [19]

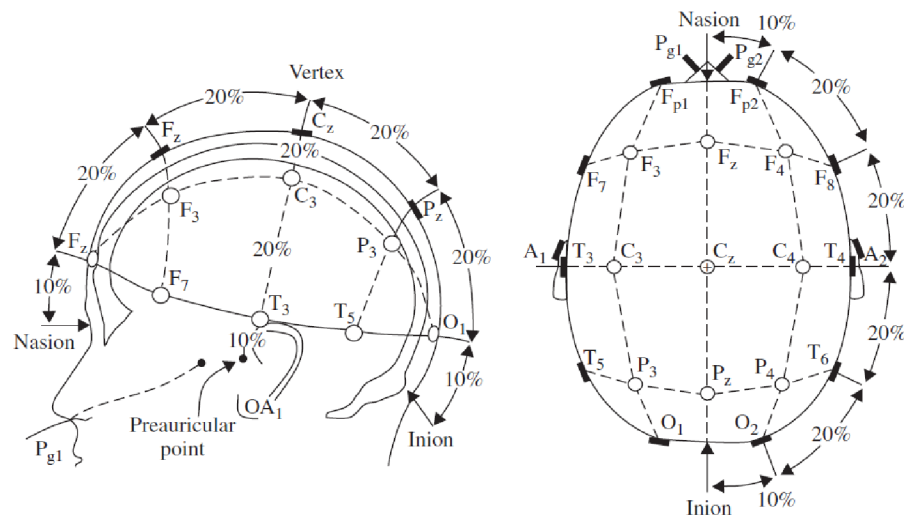
Lidská hlava se skládá z různých vrstev tkání, které různou mírou tlumí signál postsynaptických proudů detekovaný na povrchu hlavy. Mezi tři druhy tkáně, které se nejvíce podílí na útlumu patří pokožka hlavy, lebka a mozek. Lebka tlumí výrazně signál oproti měkké tkáni, chová se jako filtr typu dolní propust. Další dvě zmíněné vrstvy jsou zdrojem šumu v signálu - mozek generuje interní šum, zatímco pokožka hlavy vytváří systémový šum nebo vnější šum. [29]

1.2.1 Elektrody

Pro snímání sumy všech elektrochemických dějů odehrávajících se v mozku z povrchu hlavy jsou použity nejčastěji stříbrné argentchloridové elektrody (Ag/AgCl), ale jako vodivá část elektrody může být použit i jiný kov jako například zlato či platina. Pro zajištění vyšší vodivosti je využíván vodivý gel, který se aplikuje mezi elektrodu a kůži. Podle konkrétní aplikace je zvoleno množství elektrod v sadě pro měření. [14], [29]

Nejčastěji využívaný systém 10-20 pracuje s 19 elektrodami a až dvěma referenčními elektrodami. Jako místa možného umístění referenčních elektrod může být temeno hlavy, spojené levé a pravé ucho, spojené bradavkovité výběžky lebky (*processus mastoideus*), nebo špička nosu a další. Rozložení elektrod je znázorněno na obrázku 1.3. V tomto systému je změřena vzdálenost mezi kořenem nosu (*nasion*) a koncovým hrbolkem v týlu (*inion*) v sagitální rovině. Na obou koncích jsou označeny body ve vzdálenosti rovnající se 10 % celkové naměřené vzdálenosti. Mezi těmito

body je vzdálenost rozdělena na čtvrtiny, tudíž vzdálenost mezi těmito body se rovná 20 % celkové naměřené vzdálenosti. Ve frontální rovině je naopak změřena vzdálenost od levého chrupavčitého výstupku před zvukovodem (*tragus*) k pravému. Ale následný postup je stejný - odměření 10 % vzdálenosti na obou koncích a následně rozdělení zbytku na čtyři díly s 20 % původní délky. Rozměřování v horizontální rovině je taktéž založené na podobném principu, avšak zde se měří vzdálenosti od bodu, kde bylo v sagitální rovině změřeno 10 % od kořene nosu, až do bodu ve vzdálenosti 10% od koncového hrbolku v týlu. [14], [18], [29]

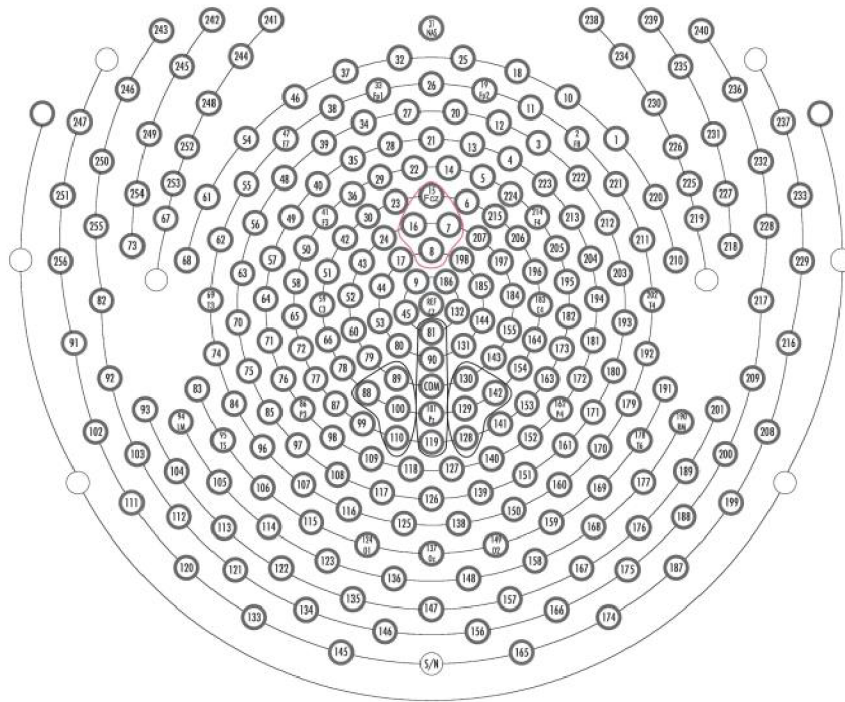


Obr. 1.3: Rozložení elektrod systému 10-20. [29]

Pro účely výzkumu lze využít i více elektrod, například 128 a více. V této práci je využito záznamu využívajícího 256 elektrod, jak je znázorněno na obrázku 1.4. Obecně však platí, že využití systému s více elektrodami nemusí poskytnout lepší prostorové rozlišení z důvodu například odporu lebky. [14], [18]

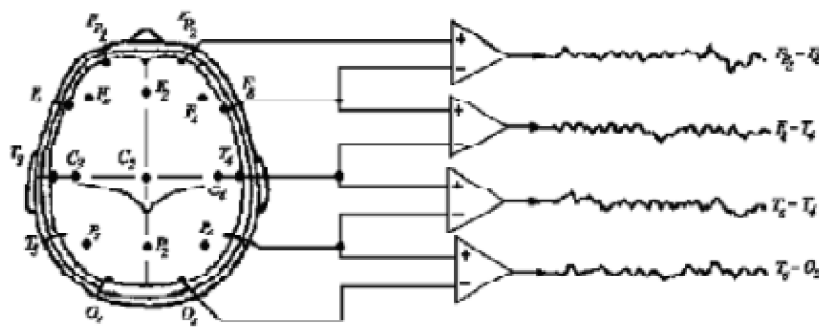
1.2.2 Montáže

Montáž označuje předpis pro vytvoření EEG signálu pomocí kombinace zvolených svodů. Používají se jak standardní montáže, tak i vlastní dle aplikace a účelu. Mezi používané druhy montáží patří bipolární, referenční, průměrovací a například také Laplaciánská. Vývod každé elektrody je připojen k diferenčnímu zesilovači. Tato součástka vychází z operačního zesilovače a výstupní napětí je úměrné rozdílu napětí na vstupech. Tudíž stejné složky v obou signálech budou na výstupu silně potlačeny. Vždy je na jeden vstup zesilovače přiveden vývod určité elektrody a druhý vstup v závislosti na zvoleném druhu montáže. [19], [29], [31]



Obr. 1.4: Rozložení elektrod systému využívajícího 256 elektrod. [18]

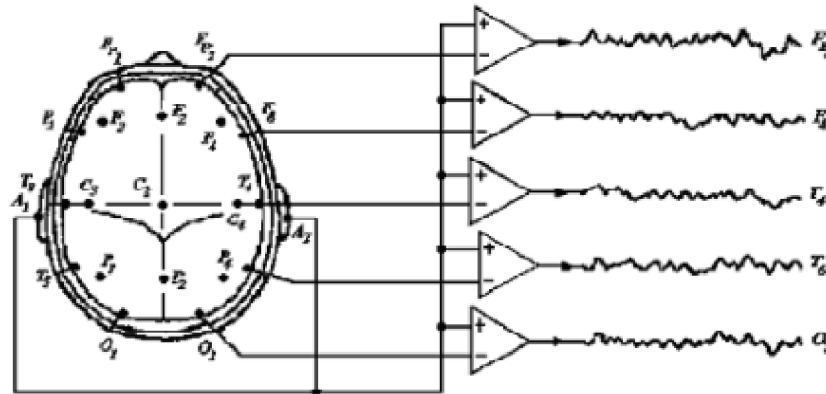
Bipolární montáž využívá jako vstupy do zesilovače sousední elektrody zapojené do temporální řady. Tato montáž je vhodná pro lokalizaci místa vzruchu i včetně koncových elektrod. Výhodou této montáže oproti referenční je menší úroveň artefaktů, avšak největší nevýhodou je nízký výstupní signál pro sousední elektrody s podobnou aktivitou. Schématické znázornění bipolární montáže je znázorněno na obrázku 1.7. [19], [29], [31]



Obr. 1.5: Schéma zapojení bipolární montáže. [16]

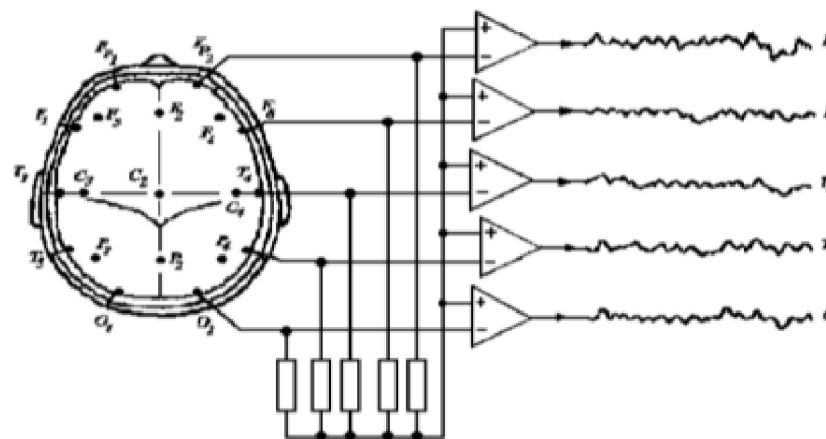
Referenční montáž označovaná jako unipolární porovnává daný svod s referenčním. Jako referenční svod mohou být vybrány různé elektrody podle vybraného vyšetření nebo patologie. U této snímákové montáže je místo vzruchu nalezeno jako

maximální výchylka. Vzhledem k podstatě snímání nedochází při této montáži ke zkreslení tvaru či frekvence při stavu, kdy oba vstupy jsou na stejném potenciálu. Ovšem při této montáži může nastat problém, pokud referenční elektroda detekuje také mozkovou aktivitu. Řešení tohoto problému se odráží v dalších dvou používaných montážích. [19]



Obr. 1.6: Schéma zapojení referenční montáže. [16]

Průměrovací montáž řeší výše zmíněný problém s referenční elektrodou tak, že jako druhý vstup do zesilovače je použit průměr přes všechny elektrody, krom elektrod blízkých očím. Další možností je Laplaciánská montáž, kdy se jako reference použije průměr ze čtyřech okolních elektrod. [19]



Obr. 1.7: Schéma zapojení průměrovací montáže. [16]

1.2.3 Další komponenty přístroje

Mezi další důležité komponenty přístroje patří analogově digitální převodník (A/D převodník) a zesilovač. [14]

A/D převodník transformuje vstupní analogový signál na signál digitální, dále zpracovatelný počítačem. Proces digitalizace začíná vzorkováním s pevně daným časovým intervalem - vzorkovací frekvencí (f_{vz}), kdy každý takto vzniklý vzorek je kvantován na digitální reprezentaci A/D převodníkem a uložen do paměti počítače. Kvantování je závislé na rozlišení převodníku. Běžně jsou převodníky minimálně osmibitové, častěji však 12-bitové. Volba vzorkovací frekvence se řídí podle Nyquistova teorému vyjádřeného vztahem:

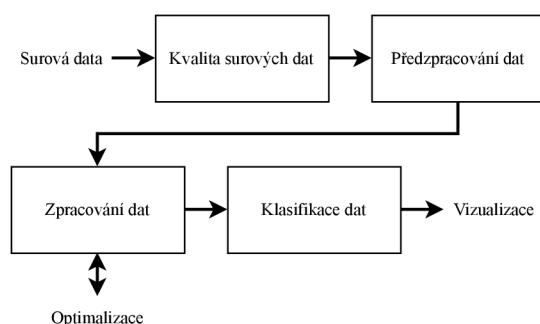
$$f_{vz} = 2 \cdot f_{max}, \quad (1.1)$$

kde f_{vz} je vzorkovací frekvence v $[Hz]$ a f_{max} je maximální frekvence vyskytující se v analogovém signálu v $[Hz]$. Platí, že efektivní užitečné pásmo frekvencí bývá obvykle nižší než je právě polovina vzorkovací frekvence. Při nedodržení tohoto teorému dochází k aliasingu. V současné době při snímání biosignálů, kam EEG patří je využíváno převzorkování. Při němž je využita vzorkovací frekvence výrazně vyšší než je nutné. Benefitem je že není nutné použít tak ostrý filtr z hlediska strmosti jeho frekvenční charakteristiky a použití vyšší frekvence neklade zvýšené finanční požadavky na A/D převodník. Problém aliasingu je dnes vyřešen převážně instalací antialiasingového filtru do převodníku. [14]

Další důležitou součástí přístroje je zesilovač výstupního signálu. Na něj jsou kladeny požadavky: vysoký vstupní odpor, malý výstupní odpor, malý vlastní šum, vysoká hodnota činitele potlačení soufázového zesílení (CMRR), ochrana pacienta před unikajícím proudem z obvodu a ochrana před nebezpečným vstupním napětím - defibrilační pulz. [14]

2 Zpracování elektroencefalografických dat

V současné době je EEG signál nejčastěji zpracováván v digitální formě. Obecně je jeho zpracování v počítači provedeno pomocí několika na sebe navazujících kroků znázorněných na obrázku 2.2. Při předzpracování jsou v signálu nalezeny úseky signálu či celé kanály s artefakty. Naleznout je lze manuální kontrolou nebo využitím nástroje, který provede tuto průzkumnou analýzu automaticky a poskytne report s informacemi o zarušených úsecích signálu, či celých kanálech. O náhledech na tuto problematiku a výpis používaných metod je napsán v další kapitole. Cílem této práce je takový nástroj navrhnout. Odstranění artefaktů lze dosáhnout prostou filtrací nebo metodou analýzy nezávislých komponent (ICA). Při běžném předzpracování poté následuje prostorová interpolace pro doplnění chybějících hodnot. Taktéž při předzpracování dat je vhodné provést kontrolu hodnot, zda se v signálu vyskytují chybějící hodnoty a extrémní hodnoty ležící mimo rozsah zbytku dat. V některých případech je vhodné provést standardizaci či normalizaci dat. Při samotném zpracování předzpracovaných dat jsou z nich extrahovány příznaky popisující signál, kdy tento proces může podléhat iterativní optimalizaci. Ze všech možných příznaků jsou selektovány jen ty nejpodstatnější pomocí nichž lze provést klasifikační analýzu nebo vizualizaci. [12], [20], [36]



Obr. 2.1: Obecné blokové schéma zpracování EEG signálu.

2.1 Artefakty a rušení

Pro dobrou eliminaci artefaktů a rušení jen nutné správně definovat jejich zdroje. Ty lze v surovém EEG signálu rozdělit do dvou velkých skupin podle jejich původu. První skupina má biologický původ - jiná než neuronální aktivita měřeného subjektu se projeví do měřeného signálu. Druhou skupinu tvoří artefakty a rušení vzniklé interferencí okolní elektroniky s měřicí během měření signálu. [19], [29], [31]

2.1.1 Artefakty a rušení biologického původu

Pohyby očí produkují elektrickou aktivitu zaměnitelnou s pomalou EEG aktivitou - θ a δ vlny. Dalším častým zdrojem artefaktu je mrkání, projevující se jako prudké změny v signálu. Ve spektru signálu se toto projevuje jako vysokofrekvenční složky. Oba dva tyto zdroje rušení lze však snadno odstranit pomocí elektrookulografického signálu. Elektrookulografie (EOG) je metoda sledující pohyby oka a směru pohledu při pozorování určitého objektu. Je zkoumán vztah mezi potenciály rohovky a sítnice pomocí elektrod umístěných blízko oka. Naměřené napětí odpovídá úhlu pohledu a jeho velikost primárně závisí na blízkosti elektrody k oku a na směru, ve kterém se oko pohybuje. Existuje korelace mezi EEG a EOG signály. [19], [29]

Dalším velkým zdrojem biologického rušení EEG signálu je činnost svalů. Jakákoliv kontrakce v blízkosti EEG elektrody se projeví rušením. Avšak jako u pohybů oka i zde je možné rušení nulovat pomocí záznamu elektrické aktivity svalů. Elektromyografie (EMG) je metoda měřící právě tuto aktivitu, její signál může být tvořen nízkými amplitudami až po šum připomínající signál. [29], [31]

Posledním velkým zdrojem rušení mohou být artefakty spjaté se srdeční činností. V případě, že se elektroda EEG nachází blízko tepny, jsou díky její pulzaci detekovatelné pulsové artefakty. Taktéž elektrická aktivita srdeční, ač na skalpu dosahuje relativně nízkých hodnot oproti EEG, může být špatně odhadnuta jako patologická aktivita. [19]

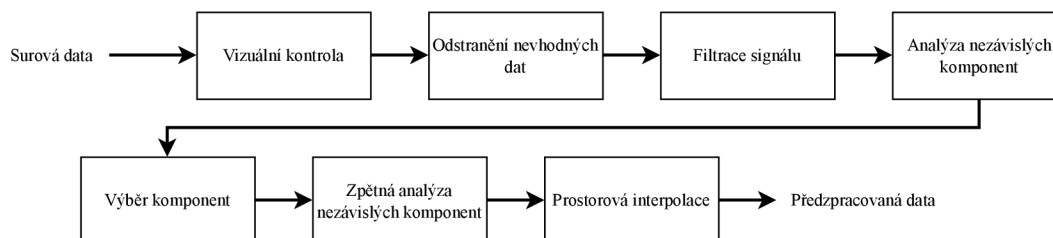
2.1.2 Artefakty a rušení nebiologického původu

Nejčastějším zdrojem rušení nebiologického původu je síťové rušení. Do spektra EEG signálu se projevuje harmonické rušení napájením ze sítě o frekvenci 50 Hz. Toto rušení lze jednoduše odstranit filtrací a to úzkopásmovou zádrží. Dalším typickým artefaktem je elektrodový popping, kdy díky poškozeným elektrodám nebo špatnému technickému stavu zařízení dojde ke změně impedance rozhraní elektroda/kůže. Toto se v signálu projevuje kolísáním signálu, především jeho izolínie s postupným návratem do původní hodnoty. Kolísání izolínie je odstranitelné průměrováním celého signálu. Problematickou může též být interference mezi okolními elektrickými zařízeními a EEG přístrojem. Zde se využívá prevence, kdy je snaha zajistit vhodné podmínky pro průběh měření. Je vhodné zmínit, že do této kategorie taktéž spadají artefakty spjaté s nedokonalostmi zesilovačů a převodem na digitální signál. [19]

2.2 Předzpracování dat

Blok předzpracování dat lze rozdělit na dílčí kroky znázorněné na obrázku 2.2. Prvním krokem je vizuální kontrola dat. Je vhodné v signálu označit úseky nevhodné

pro další analýzu z důvodu vysokého podílu šumu. Taktéž je možné úplně vynechat z následné analýzy celé signály určitých elektrod, z důvodu zatížení artefakty od svalové aktivity. Poté jsou metodou ICA vybrány vhodné komponenty a je provedena prostorová interpolace. [12], [36]



Obr. 2.2: Obecné blokové schéma předzpracování EEG signálu.

2.2.1 Filtrace

Obecně rozlišujeme filtry na lineární a nelineární podle toho, zda u nich platí princip superpozice. Dále lineární filtry dělíme na dva druhy - filtry s konečnou impulzní charakteristikou (FIR) a filtry s nekonečnou impulzní charakteristikou (IIR). Další možností dělení filtrů je podle propustného pásma na dolní a horní propust, pásmovou propust a zádrž. [12], [36]

FIR filtr má často nerekurzivní realizaci, jeho návrh je více intuitivní a je absolutně stabilní. Fázová charakteristika systému je vždy lineární bez ohledu na symetrii impulsní charakteristiky. Nevýhodou FIR filtrů je nutnost velmi dlouhé impulsní charakteristiky pro dosažení velké strmosti mezi propustným a nepropustným pásmem, což se negativně projevuje na výpočetní náročnosti a celkové zpoždění systému. [12], [36]

IIR filtry jsou realizovány rekurzivně a je proto nutné ověřit jejich stabilitu. Taktéž vždy má filtr nelineární fázovou charakteristiku. Oproti FIR filtrům dosahuje IIR filtr s nižším řádem mnohem strmějších přechodů a kratších výpočetních dob. [12], [36]

Pro filtraci EEG signálu je taktéž důležitá amplitudová frekvenční charakteristika. Je snaha o co největší hladkost křivky v prostupném pásmu, zvlnění zde odpovídá zesílení nebo zeslabení určité frekvence a co největší strmost při přechodu do nepropustného pásma, ve kterém by měl být maximální útlum. Typickými filtry splňujícími tyto kritéria je IIR Butterworthův filtr, nebo IIR Chebyshevův filtr. Obecně tedy lze říct, že se využívají oba druhy filtrů, IIR filtry pro rychlost výpočtu a FIR pro lineární fázovou frekvenční charakteristiku. [12], [36]

2.2.2 Analýza nezávislých komponent

Výše zmíněná filtrace je vhodná pro odstranění rušení, které je ve frekvenční oblasti odlišitelné od EEG signálu. Pro případy, kdy tento stav nenastává lze využít metodu nezávislých komponent. Matematicky tuto metodu lze vyjádřit pomocí vztahu:

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S}, \quad (2.1)$$

kde \mathbf{S} je zdrojový signál o délce m , \mathbf{X} je snímaný signál a \mathbf{A} je směšovací matice o rozměrech $m \times m$. Cílem metody ICA je nalezení matice \mathbf{W} , jenž je inverzní k směšovací matici a výsledkem jejího součinu se snímaným signálem bude odhad původního zdrojového signálu. [20]

Algoritmy pro výpočet ICA se dělí do dvou skupin podle postupu řešení. První skupina je založená na minimalizaci vzájemné informace, dobře se přizpůsobuje variacím v datech, avšak je závislá na správném určení rozložení dat. Naopak druhá metoda je založena na maximalizaci nenormality dat. Výběr metody se odvíjí na konkrétní aplikaci. [20]

2.2.3 Interpolace

Při odhalení elektrody poskytující chybový signál je tento signál odstraněn a nahrazen interpolací signálů z okolních elektrod. Metod interpolace je mnoho, pro interpolaci EEG dat se často využívá 3D sférického splinu prvního řádu. Zde pro každý chybějící vzorek chybového signálu je vypočítána hodnota z okolních elektrod pomocí lineární interpolace. Zde bude vypočítaná hodnota ležet v rozsahu naměřených hodnot, což by při použití splinů s polynomem vyššího řádu nemuselo nastat. Pro pozdější analýzu EEG signálu je vhodné v co nejmenší míře používat interpolaci, neboť jejím výsledkem je umělý signál. Taktéž je problematická interpolace okrajových elektrod, kde není dostatek elektrod v okolí. [12]

3 Kvalita elektroencefalografických dat

Pro zjištění, zda naměřená data jsou vhodná pro další analýzu je nutné zhodnotit, zda jsou dostatečně kvalitní. Tedy je nutné odhalit jestli jsou v signálu jednotlivých kanálů přítomny artefakty a rušení. Přítomnost lze odhalit analýzou signálu v určité doméně, odkud jsou získány příznaky popisující kvalitu. [8], [28], [39]

3.1 Analýza elektroencefalografického signálu

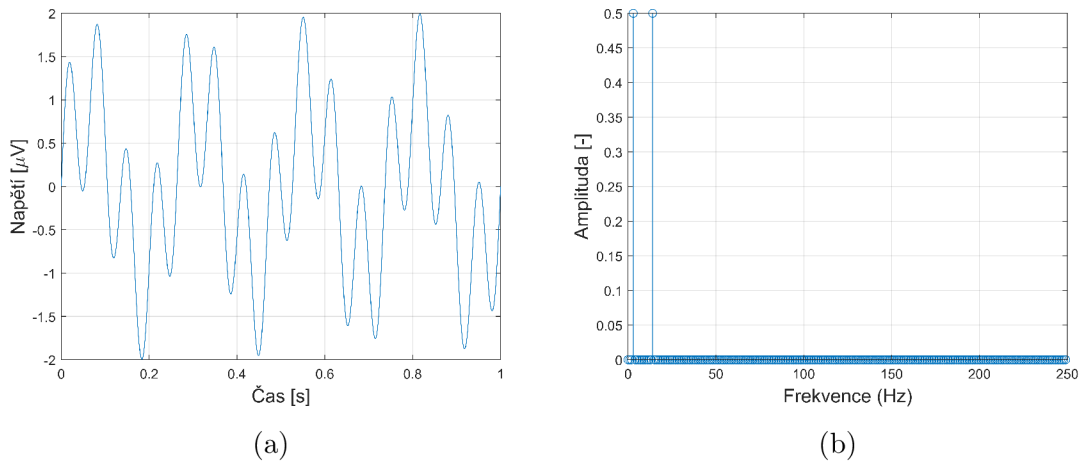
Analýza EEG signálů zahrnuje extrakci smysluplných informací ze zaznamenaných dat, aby bylo možné identifikovat vzorce nebo abnormality. Použité metody se zaměřují na analýzu v časové, frekvenční, časově-frekvenční a prostorové oblasti. Každá doména nabízí jedinečné perspektivy a metody pro odhalování cenných poznatků z EEG dat. [8], [28], [39]

3.1.1 Analýza signálu v časové oblasti

Analýza EEG signálů v časové oblasti hraje klíčovou roli v pochopení časové dynamiky a charakteristik mozkové aktivity. Tento přístup zahrnuje zkoumání nezpracovaných EEG dat v jejich časové doméně, bez jakýchkoli frekvenčních nebo prostorových transformací. K získání náhledu na amplitudu, variabilitu a časové vzorce signálu se používají různé funkce a techniky. Amplitudová analýza poskytuje informace o síle a intenzitě nervové aktivity, zatímco míry jako průměr, rozptyl, šikmost a špičatost kvantifikují distribuci a statistické vlastnosti signálu. Frekvence průchodů nulou odhaluje rychlost oscilací. Kromě toho Hjorthovy parametry, včetně aktivity, mobility a složitosti, zachycují různé aspekty energie signálu, změny tvaru vlny a nepravidelnosti. [12], [28]

3.1.2 Analýza signálu ve frekvenční oblasti

Analýza ve frekvenční oblasti je dalším z možných přístupů v analýze EEG signálu využívající spektrální vlastnosti dat. Použitím Fourierovy transformace je signál EEG transformován z časové oblasti do frekvenční oblasti, jak je znázorněno na obrázku 3.1. To umožňuje zkoumat rozložení a sílu různých frekvenčních složek, jako jsou vlny α , β , θ a γ . Analýza ve frekvenční oblasti poskytuje cenné informace o dominantních frekvenčních pásmech a jejich relativních amplitudách a odhaluje důležité poznatky o oscilační aktivitě mozku. [12], [28]



Obr. 3.1: Harmonický signál a jeho amplitudové spektrum, kdy na obrázku (a) je časový průběh harmonické funkce tvořené dvěma frekvencemi a na obrázku (b) je její amplitudové spektrum. [12]

3.1.3 Analýza signálu v časově-frekvenční oblasti

Analýza signálu v časově frekvenční oblasti kombinuje jak časové, tak spektrální vlastnosti dat. Tento přístup umožňuje zkoumat, jak se frekvenční obsah EEG signálu mění v průběhu času, čímž je zachycena dynamická povaha mozkové aktivity. Mezi používané techniky patří krátkodobá Fourierova transformace a vlnková transformace. Aplikací těchto transformací se mohou identifikovat a zkoumat přechodné události, spektrální změny související s událostmi a frekvenční dynamiku, které nemusí být zřejmé ani v časové nebo frekvenční doméně samotné. Časově-frekvenční reprezentace poskytuje detailní pohled na časově proměnnou distribuci výkonu nebo amplitudy v různých frekvenčních pásmech, odhaluje časový vývoj neurálních oscilací a jejich vztah ke kognitivním procesům. [12], [28], [32]

3.1.4 Analýza signálu v prostorové oblasti

Analýza signálu v prostorové oblasti se zaměřuje na prostorovou distribuci a topografické charakteristiky mozkové aktivity. Tato analýza zahrnuje zkoumání elektrických potenciálů zaznamenaných elektrodami umístěnými na pokožce hlavy, aby se získal pohled na nervové procesy probíhající v různých oblastech mozku. Techniky, jako jsou skalповé mapy a topografické grafy, se běžně používají k vizualizaci amplitudy nebo distribuce výkonu napříč umístěními elektrodami, odhalující prostorové vzorce mozkové aktivity. Analýza prostorové domény umožňuje posouzení funkční konektivity mezi oblastmi mozku, identifikaci mozkových sítí a lokalizaci abnormalit nebo fokálních aktivací. [28], [32]

3.2 Extrakce příznaků

Extrakce příznaků umožňuje extrakci relevantních informací ze surových nezpracovaných EEG dat. Signály EEG jsou složité a obsahují obrovské množství informací, takže je nutné extrahovat specifické rysy, které zachycují základní charakteristiky mozkové aktivity nebo artefaktů a rušení. Tyto metody zahrnují transformaci surových dat do souboru smysluplných a reprezentativních příznaků, které mohou zahrnovat například statistické míry, spektrální koeficienty, deskriptory časové domény nebo míry konektivity. Cílem extrakce příznaků v analýze kvality signálu EEG je odhalení přítomnosti artefaktů a rušení. [8], [39]

3.2.1 Příznaky složitosti

Příznaky složitosti poskytují cenné poznatky o složité a nelineární dynamice mozkové aktivity. Tyto funkce zachycují složitost, nepravidelnost a informační obsah EEG signálu a nabízejí hlubší pochopení základních nervových procesů. Pro kvantifikaci složitosti EEG signálu se běžně používají míry jako fraktální dimenze, entropie a multiškálová entropie. Fraktální dimenze charakterizuje soběpodobnost a škálovací vlastnosti signálu, zatímco entropie měří náhodnost a informační obsah. Víceškálová entropie zkoumá složitost napříč různými časovými měřítky a poskytuje pohled na časovou organizaci mozkové aktivity. [28]

Tab. 3.1: Tabulka se seznamem běžně používaných příznaků složitosti.

Příznaky složitosti	Stručný popis
Shannonova entropie	Aditivní míra stochasticity signálu [12]
Tsalisova entropie	Neaditivní míra stochasticity signálu [9]
Hjortova mobilita	Střední frekvence signálu [25]
Hjortova komplexita	Rychlost změny Hjortovy mobility [25]
Množství informace	Entropie vlnkově rozloženého signálu [30]
Koeficienty cepstra	Rychlost změny výkonu spektrálního pásma signálu [27]
Lyapunovův exponent	Separace mezi signály s podobnou trajektorií [37]
Falešný nejbližší soused	Spojitosť a plynulosť signálu [10]

3.2.2 Příznaky kontinuity

Příznaky kontinuity nabízí náhled na časovou dynamiku a plynulosť mozkové aktivity. Tyto funkce zachycují kontinuitu nebo pravidelnost signálu EEG a nabízejí

cenné informace o stabilitě a konzistenci nervových procesů. K extrakci rysů kontinuity ze signálů EEG se běžně používají autokorelace, derivace v časové doméně a vlnková koherence. Autokorelace kvantifikuje korelaci mezi signálem a jeho zpožděnými verzemi, odhaluje časové závislosti a opakující se vzorce. Derivace v časové oblasti hodnotí rychlost změny nebo sklon signálu, což naznačuje hladkost nebo náhlé přechody. Vlnová koherence zkoumá koherenci nebo fázovou konzistenci mezi různými frekvenčními složkami a poskytuje pohled na synchronizaci a konektivitu neurálních oscilací. [12], [28], [32]

Tab. 3.2: Tabulka se seznamem běžně používaných příznaků kontinuity.

Příznaky kontinuity	Stručný popis
Medián frekvence	Medián spektrálního výkonu [20]
Výkon δ pásma	Spektrální výkon v pásmu 0 až 3 Hz
Výkon θ pásma	Spektrální výkon v pásmu 4 až 7 Hz
Výkon α pásma	Spektrální výkon v pásmu 8 až 15 Hz
Výkon β pásma	Spektrální výkon v pásmu 16 až 31 Hz
Výkon γ pásma	Spektrální výkon v pásmu nad 32 Hz
Směrodatná odchylka	Průměrný rozdíl mezi hodnotou signálu a jeho průměrnou hodnotou [20]
Poměr α/δ	Poměr výkonové spektrální hustoty v pásmech α a δ [35]
Zpomalení difúze	Ukazatel maximální spektrální hustoty výkonu [11]
δ impuls po hrotu	Nárůst δ po špičce ve srovnání s δ před špičkou [11]
Počet impulsů	Počet amplitudových impulsů
Délka impulsu	Statistické vlastnosti impulsu
Síla pásma impulsu	Spektrální síla impulsu
Počet potlačení	Segmenty se souvislým potlačením amplitudy
Délka potlačení	Statistické vlastnosti potlačení

3.2.3 Příznaky konektivity

Pohled na funkční konektivitu a komunikaci mezi různými oblastmi mozku umožňují příznaky konektivity, jenž popisují interakce, synchronizaci a koherenci neuro-nové aktivity. K extrakci příznaků se používá například koherence, hodnota fázového uzamčení (PLV) a metriky založené na teorii grafů. Koherence kvantifikuje lineární korelaci a synchronizaci mezi různými frekvenčními složkami, což ukazuje na funkční konektivitu mezi oblastmi mozku. PLV měří fázovou konzistenci nebo fázový zámeček

mezi nervovými oscilacemi, což odráží stupeň synchronizace a komunikace. Pro využití metrik založených na teorii grafů je nutné mít matici konektivity, na kterou se aplikují síťové metriky jako centralita, modularita a další. [28], [32]

Tab. 3.3: Tabulka se seznamem běžně používaných příznaků konektivity.

Příznaky konektivity	Stručný popis
Koherence	Koherence ve výkonu 0 až 4 Hz mezi signály [35]
Vzájemná informace	Míra závislosti [28]
Grangerova kauzalita	Míra kauzality [5]
Index fázového zpoždění	Asociace mezi okamžitou fází signálů [33]
Velikost vzájemné korelace	Maximální korelace mezi signály [12]
Zpoždění vzájemné korelace	Časové zpoždění, které maximalizuje korelaci mezi signály [12]

3.3 Metody detekce odlehlých hodnot

Metody detekce odlehlých hodnot při extrakci příznaků EEG mají za cíl identifikovat datové body nebo úseky, které se významně odchyľují od očekávaného nebo normálního chování. Tyto metody využívají různé techniky, jako jsou statistické přístupy, algoritmy strojového učení nebo doménově specifická heuristika, k identifikaci odlehlých hodnot a jejich odlišení od základního EEG signálu. [28], [39]

3.3.1 Statistické metody

Detekce odlehlých hodnot na základě histogramu je používaným přístupem, jehož cílem je identifikovat a označit datové body, které se výrazně odchyľují od očekávaného rozložení. Tato metoda zahrnuje sestavení histogramu, který představuje frekvenční rozložení hodnot dat, a poté použití statistických technik k identifikaci odlehlých hodnot. Analýzou tvaru, šíření a hustoty histogramu lze odlehlé hodnoty identifikovat jako datové body, které spadají mimo určitý rozsah nebo vykazují výrazně odlišné rozložení. Detekce odlehlých hodnot na základě histogramu poskytuje jednoduchý, ale účinný způsob, jak detekovat a odstranit artefaktové datové body. [13], [28]

Lokální odlehlý faktor (LOF) je robustní metoda detekce odlehlých hodnot, fungující na principu posouzení stupně odlehlosti datového bodu na základě jeho místní hustoty ve srovnání s jeho sousedními datovými body. Zvážením místní hustoty datových bodů může LOF efektivně identifikovat odlehlé hodnoty, které se odchyľují

od okolního kontextu. LOF počítá skóre pro každý datový bod, přičemž vyšší skóre značí vyšší pravděpodobnost, že se jedná o odlehlou hodnotu. V kontextu extrakce rysů EEG lze LOF použít k identifikaci abnormálních segmentů EEG nebo rysů, které vykazují neobvyklé vzory, jako jsou artefakty nebo vzácné fyziologické události. [6], [28]

Angle-Based Outlier Detector (ABOD) je výkonná metoda detekce odlehlých hodnot, která měří úhly mezi datovými body ve vícerozměrném prostoru. ABOD vypočítává relativní úhly vytvořené datovým bodem vzhledem k jeho sousedním datovým bodům, což mu umožňuje identifikovat odlehlé hodnoty na základě jejich odchylky od očekávaných úhlových vzorů. Při zohlednění úhlů spíše než vzdáleností je ABOD méně citlivý na vliv irelevantních nebo hlučných dimenzí v prostoru prvků. V kontextu extrakce rysů EEG lze ABOD využít k detekci anomálních vzorců EEG nebo rysů, které vykazují významné odchylky od sousedních bodů. [17], [28]

Jednotřídová metoda podpůrných vektorů (OCSVM) je robustní metoda, kdy řízený algoritmus učení se učí charakteristiky normálních datových bodů a poté identifikuje odlehlé hodnoty jako datové body, které se výrazně odchylují od naučeného normálního vzoru. Modelováním normálního chování funkcí EEG může detektor OCSVM účinně rozlišovat mezi normálními a abnormálními vzory. [1], [28]

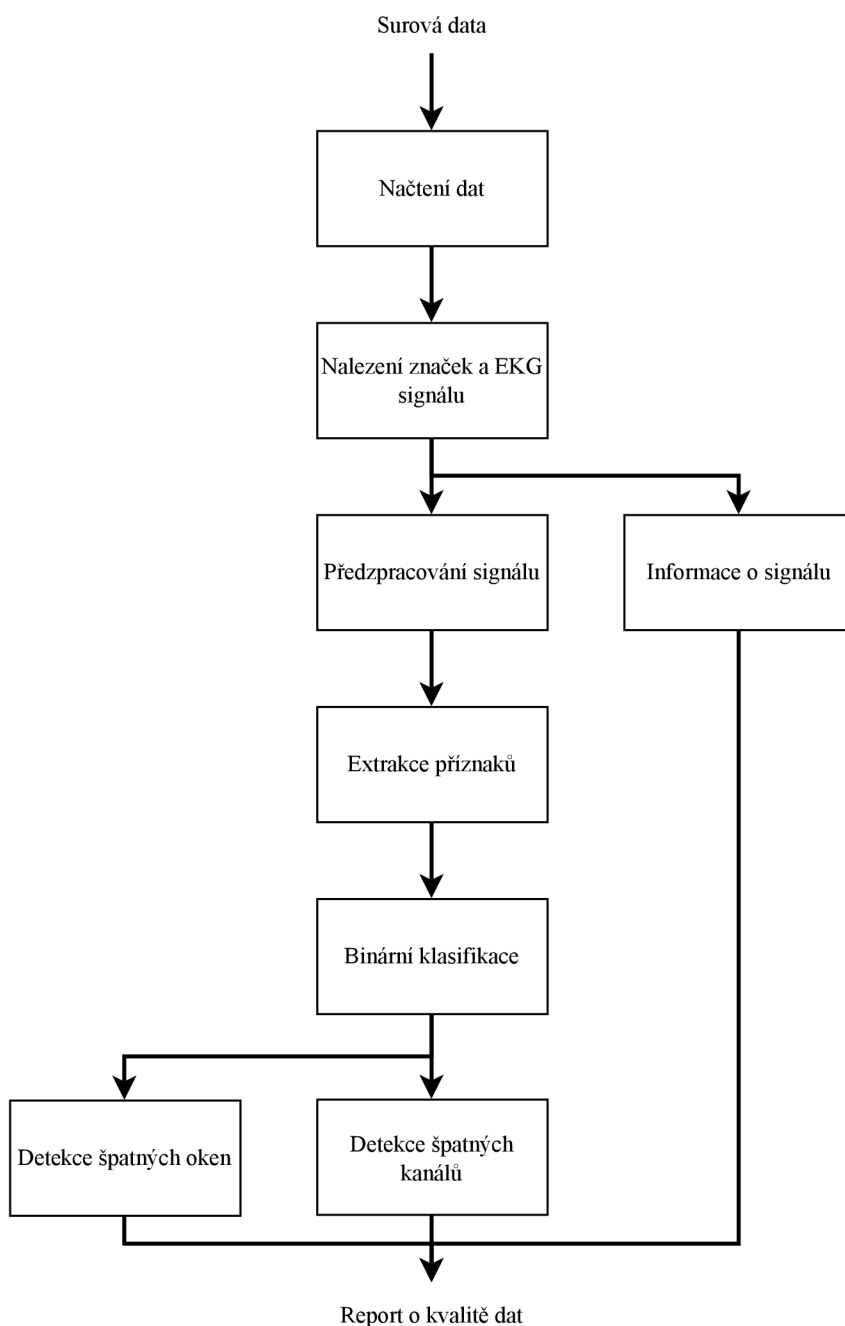
Lokálně selektivní kombinace v paralelních odlehlých souborech (LSCPOE) je pokročilá metoda detekce odlehlých hodnot zahrnujících několik algoritmů v EEG datech. Selektivním kombinováním výsledků těchto algoritmů na základě místních charakteristik a vzorů dosahuje LSCP robustní a přesné detekce odlehlých hodnot. Tento přístup je zvláště užitečný při práci s různými odlehlými vzory a složitými distribucemi dat. [28], [38]

3.3.2 Metody založené na učení

Metody založené na učení jsou efektivními přístupy k detekci odlehlých hodnot. Tyto metody využívají sílu algoritmů strojového učení k automatickému rozpoznání vzorů a charakteristik normálních EEG signálů. Tréninkem na označených datech nebo použitím technik učení bez učitele mohou tyto metody zachytit přirozenou složitost a variabilitu EEG dat. Přístupy učení s učitelem využívají anotovaná tréninková data k trénování klasifikátorů, které dokážou rozlišit mezi normálními a abnormálními EEG vzory. Na druhé straně metody učení bez dozoru mají za cíl odhalit odlehlé hodnoty modelováním normální distribuce rysů EEG a identifikací případů, které se významně odchylují od této naučené reprezentace. Metody založené na učení nabízejí výhody adaptability, škálovatelnosti a schopnosti zvládnout různé typy odlehlých hodnot, včetně artefaktů, anomálií nebo vzácných událostí. [3], [28]

4 Návrh nástroje a jeho programové řešení

Smysl navrženého nástroje je před předzpracování surových EEG dat provést průzkumnou analýzu a nalézt úseky signálu či celé kanály s artefakty a rušením. Nástroj automaticky nalezne a doporučí, o která místa se jedná. Výstupem nástroje je report o kvalitě dat. Nástroj je navržen tak, aby bylo možné ho rozšířit o další metriky kvality. Blokové schéma nástroje je znázorněno na obrázku 4.1.



Obr. 4.1: Blokové schéma navrženého nástroje.

4.1 Načtení elektroencefalografických dat

Prvním krokem nástroje je načtení naměřených dat z datové struktury měřicího přístroje a převést je do formátu lépe počítačově zpracovatelného. Toho lze dosáhnout s pomocí volně dostupné knihovny FileIO jež je součástí Fieldtrip toolboxu. EEG data mohou nabývat různých datových formátů podle výrobce zařízení. Seznam podporovaných datových formátů a výrobců je uveden v tabulce 4.1.

Tab. 4.1: Tabulka podporovaných datových formátů podle výrobců. [26]

Výrobce EEG přístroje	Datové formáty
ANT Neuro	.avr, .cnt, .trg
Besa	.avr, .swf
Biosemi BDF	.bdf
Brain Products	.eeg, .seg, .dat, .vhdr, .vmrk
Cambridge Electronic Design	.smr
EEGLAB	.set, .fdt
Magstim Electrical Geodesics, Inc.	.egis, .ave, .gave, .ses, .raw, .mff
NeuroScan	.eeg, .cnt, .avg
Nexstim	.nxe
TMSi	.Poly5
Generické standardní formáty	.edf, .gdf

Pro načtení dat je použit FieldTrip toolbox, což je sada softwarových nástrojů pro interaktivní programové prostředí MATLAB, jež je vhodná pro analýzu dat z EEG, magnetoencefalografie (MEG) a intrakraniální elektroencefalografie (iEEG). Nabízí předzpracování a pokročilé analytické metody, jako je časově-frekvenční analýza, rekonstrukce zdroje pomocí dipólů a neparametrické statistické testování. Podporuje datové formáty všech hlavních systémů MEG a nejpobulárnějších systémů EEG a iEEG. [26]

4.1.1 Použité datové formáty

Poskytnutá EEG data z laboratoře multimodálního a funkčního zobrazování CEI-TEC MU jsou pouze ve dvou datových formátech - Meta File Format (MFF) a Brain Vision File Format (BVFF) z důvodu použití EEG systémů od výrobců Magstim Electrical Geodesics, Inc. (Magstim EGI) a Brain Products (BP).

MFF je široce používaný formát souboru v analýze signálu EEG od výrobce Magstim EGI. Je speciálně navržen pro ukládání a organizování EEG dat spolu se

souvisejícími metadaty. MFF poskytuje komplexní strukturu pro ukládání různých komponent EEG experimentů, včetně nezpracovaných EEG záznamů, značek událostí, pozic elektrod, informací o účastnících a experimentálních protokolů. Tento formát umožňuje snadný přístup, správu a sdílení dat EEG napříč různými analytickými platformami a softwarovými nástroji. Formát MFF podporuje interoperabilitu a standardizaci v analýze signálu EEG, usnadňuje reprodukovatelnost výzkumných zjištění a spolupráci mezi výzkumníky v oblasti neurověd. Veškeré informace je možné získat načtením souboru s koncovkou `.mff`. [26]

BVFF je široce používaný formát souborů při analýze signálu EEG od výrobce BP. Je speciálně navržen pro ukládání a výměnu EEG dat spolu s přidruženými metadaty. BVFF poskytuje standardizovanou strukturu pro ukládání nezpracovaných záznamů EEG, konfigurací elektrod, značek událostí a dalších relevantních informací. Tento formát podporuje kompatibilitu mezi různými systémy EEG a softwarovými nástroji, což usnadňuje sdílení dat a spolupráci mezi výzkumníky. Oproti předchozímu formátu je nutné pro správnou analýzu načíst tři samostatné soubory. Meta data jsou uložena v souboru s koncovkou `.vhdr`, soubor textových značek s koncovkou `.vmrk` obsahující informace o událostech v datech a hodnoty napětí EEG uložené jako binární soubor s koncovkou `.eeg`. Výhoda BVFF oproti MFF je, že se jedná o schválený formát standardu Brain Imaging Data Structure (BIDS). Ten vznikl jako jednoduchý a snadno použitelný způsob organizace komplikovaných dat neurologických experimentů. [26]

4.2 Přítomnost značek a elektrokardiografického signálu

Pro určení kvality EEG záznamu je taktéž nutné stanovit přítomnost značek v záznamu. Značky v datech EEG slouží k označení a pojmenování konkrétních událostí nebo časových bodů zájmu během záznamu EEG. Při snímání EEG dat je často snímán i elektrokardiologický signál (EKG). Po odhalení přítomnosti EKG signálu je nalezený kanál odstraněn z dat pro další analýzu a do reportu o kvalitě je uložena logická hodnota o přítomnosti EKG.

4.3 Předzpracování dat

Po načtení dat je nutné provést jejich předzpracování, avšak oproti informacím v kapitole 2.2 jsou tyto operace provedeny pouze jako pomocný nástroj s co nejmenším vlivem na informaci v surových datech. Předzpracování se skládá z filtrace a transformace dat.

Pro odstranění lineárního posunu v surových EEG datech je provedena filtrace filtrem typu horní propust s mezní frekvencí 1 Hz . Specificky byl použit Butterworthův filtr druhého řádu. Kvůli požadavkům funkce v MATLAB toolboxu je nutné normalizovat mezní frekvenci pomocí vztahu:

$$f'_m = \frac{f_m}{\frac{1}{2}f_{vz}}, \quad (4.1)$$

kde f'_m je normalizovaná mezní frekvence filtru, f_m je mezní frekvence filtru [Hz] a f_{vz} je vzorkovací frekvence [Hz].

Po provedení filtrace je provedena transformace dat - zkrácení celkové délky na délku N' , protože metody v následné analýze extrahují příznak pro časové okno o délce jedné vteřiny.

$$N' = N - (N \bmod f_{vz}), \quad (4.2)$$

kde N' je délka zkrácených dat, N je délka původních dat a $N \bmod f_{vz}$ je výsledek operace modulo - celočíselný zbytek po dělení původní délky dat vzorkovací frekvencí.

4.4 Výpočet příznaků

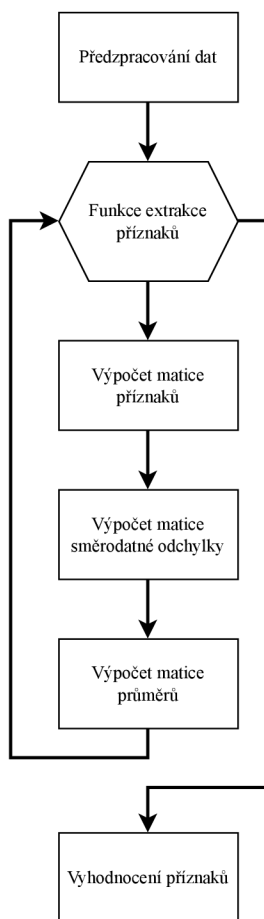
Filtrovaná a zkrácená data z předchozí funkce jsou transformována aby byla lépe využita již existujícími MATLAB funkcemi a tím byl zkrácen výpočetní čas. Na tyto data jsou aplikovány veškeré funkce extrakce příznaků podle schématu znázorněného na obrázku 4.2. Popis významu dílčích funkcí je uveden níže. Taktéž pro každý extrahovaný příznak je spočítána směrodatná odchylka a průměrná hodnota pro daný kanál. Tyto hodnoty budou sloužit ke klasifikaci výsledků.

4.4.1 Statičnost signálu

Detekce statického signálu vychází z předpokladu, že signály s malou směrodatnou odchylkou (SD) nebo absolutní střední odchylkou (MAD) budou pravděpodobně konstantní/extrémně slabé. První metoda tedy detekuje konstantní signály pomocí výpočtu SD a MAD a je porovnána s velmi malou hodnotou napětí ($10^{-9}\mu\text{V}$). Kromě toho jsou také detekována okna obsahující jakékoliv obecně neplatné hodnoty. Všechna tato detekovaná okna jsou považována za špatná, z důvodu že se v nich nenachází EEG signál. [39]

4.4.2 Koeficient špičatosti

Tento koeficient je standardní statistická míra závažnosti konců distribuce vzorků, tedy udává, s jakou pravděpodobností vzorek obsahuje odlehlou hodnotu. Čím vyšší



Obr. 4.2: Blokové schéma extrakce příznaků a statistických veličin.

je pravděpodobnost odlehlých hodnot ve vzorku, tím větší je hodnota koeficientu. Podobně, čím rovnoměrnější je rozdělení, tím nižší je koeficient špičatosti vzhledem k normálnímu rozdělení. Pomocí následujícího vztahu je možné vypočítat koeficient špičatosti pro signál x délky N

$$K = \frac{\frac{1}{N} \sum_{n=1}^N (x(n) - \bar{x})^2}{\left[\frac{1}{N} \sum_{n=1}^N (x(n) - \bar{x})^2 \right]^2}, \quad (4.3)$$

kde $x(n)$ je hodnota vzorku signálu na aktuální pozici, \bar{x} je průměrná hodnota vzorků v okně.

4.4.3 Síťové rušení

Popis vzniku a vlivu artefaktu síťového rušení je uveden v kapitole 2.1.2. Pro změření síly síťového rušení v EEG signálu, je vypočítána průměrná hodnota jednostranného amplitudové spektrum signálu v rozsahu $50 \pm 1Hz$.

4.4.4 Maximální gradient

Maximální gradient signálu EEG je největší rozdíl mezi všemi sousedními vzorky v okně. Toto je běžně používaná metoda detekce artefaktů v analýzách EEG, kde se krok $10\mu V/ms$ často používá jako práh pro indikaci jakéhokoli vysokoamplitudového či vysokofrekvenčního artefaktu, který vytváří rychlé změny snímaného napětí. Maximální gradient lze vypočítat pomocí vztahu

$$MG = \max[x(n) - x(n - 1)], \quad (4.4)$$

kde MG je hodnota maximálního gradientu, $x(n)$ je hodnota vzorku signálu na aktuální pozici a $x(n - 1)$ je hodnota vzorku signálu na předešlé pozici.

4.4.5 Amplituda kvadratického průměru

Odmocnina střední kvadratické amplitudy (RMSA) signálu EEG je obecným měřítkem velikosti signálu v celém okně, bez ohledu na frekvenci. Pro signál X délky N lze efektivní hodnotu RMSA vypočítat pomocí vztahu

$$RMSA = \sqrt{\frac{1}{N} \sum_{n=1}^N |x_n|^2}. \quad (4.5)$$

Platí, že okna obsahující artefakty s vysokou amplitudou by měla mít větší hodnotu RMSA než okna s amplitudami v normálním rozsahu EEG.

4.4.6 Síla v užitečném pásmu

V kapitole 1.1.2 je uvedeno, že mezi frekvencemi 1 až 50 Hz se nachází veškeré důležité rytmy. Pro zjištění, zda se v signálu nachází tyto rytmy je vypočítáno amplitudové jednostranné spektrum celého signálu a zprůměrováno přes frekvenční rozsah oblasti zájmu. Porovnáním těchto hodnot s hodnotami síťové rušení je určeno, zda se v signálu nachází interference mezi měřícím zařízením a okolními přístroji napájenými ze sítě.

4.4.7 Míra průchodu nulou

Míra průchodu nulou (ZCR) určuje míru s jakou se signál mění z pozitivních na negativní hodnoty. Je to ukazatel frekvence, při které je většina energie soustředěna ve spektru signálu. ZCR by se mělo zvyšovat u vysokofrekvenčních artefaktů a klesat u nízkofrekvenčních artefaktů, které způsobují posun EEG stopy od nulové čáry. Pomocí vztahu lze vypočítat

$$ZCR = \frac{1}{N} \sum_{n=1}^N |\operatorname{sgn}[x(n)] - \operatorname{sgn}[x(n - 1)]| \quad (4.6)$$

4.4.8 Rozšiřitelnost o výpočet dalších příznaků

Při dodržení stejného formátu vstupů je možné rozšířit toolbox přidáním dalších souboru s kódem.

4.5 Vyhodnocení příznaků

Matice hodnot příznaků X o počtu kanálů M a počtu oken N je poté binárně klasifikována na matici logických hodnot Y pomocí vztahu

$$y_{i,k} = \begin{cases} 1 & \mu - \sigma < x_{i,k} < \mu + \sigma \\ 0 & \text{jinak} \end{cases}, \quad (4.7)$$

kde $i = 1, 2, \dots, M$ a $k = 1, 2, \dots, N$.

4.6 Detekce špatných úseků signálů a kanálů

Matice logických hodnot dále vstupuje do dvou samostatných cest na jejichž výstupu jsou detekovaná okna a kanál, která jsou poškozená artefakty a rušením. Pro detekci oken je prvním krokem sumace oken klasifikovaných jako dobrá přes všechny okna a daný příznak. Hodnoty sumace jsou poté poděleny počtem oken aby byli získávány poměry. Tedy nízkých hodnot dosahují takové okna, kde je nízký počet dobrých oken. Jelikož pro každou metriku se pohybují poměry v různém rozmezí je provedena normalizace pomocí maximální a minimální hodnoty podle vztahu

$$y_{i,k} = \frac{x_{i,k} - \min(x_k)}{\max(x_k) - \min(x_k)} \quad (4.8)$$

Jako špatné okno je vybráno takové, jehož normalizovaný poměr je nižší než stanovený práh. Výzkumem byla hodnota prahu T stanovena pomocí vztahu

$$T = \begin{cases} 0,7 & \bar{x}_i > 0,7 \\ \bar{x}_i & \text{jinak} \end{cases}, \quad (4.9)$$

kde \bar{x}_i je průměrná hodnota normalizovaného poměru pro dané okno.

Detekce špatných kanálů probíhá stejně jen matice logických hodnot je sečtena přes všechny kanály a daný příznak, poté následuje dělení vzniklých sum počtem kanálů a normalizace poměrů. Jako špatný kanál je vybrán ten s nejnižší hodnotou normalizovaného poměru.

4.7 Report o kvalitě dat

Výstupem tohoto programu je report ve formátu `.pdf` vytvořeného pomocí toolboxu MATLAB Report Generator. V reportu se nachází jméno souboru, počet kanálů, vzorkovací frekvence, přítomnost značek, tabulka s unikátními značkami a jejich četností, přítomnost síťového rušení a EKG signálu v datech, tabulka detekovaných špatných oken a detekovaný kanál s jeho hodnotou standardizovaného poměru.

V první části reportu se nachází základní informace o souboru jakým je jeho jméno, vzorkovací frekvence a počet kanálů. V další části jsou informace o přítomnosti značek v signálu. Pokud se v signálu vyskytují, tak jsou zjištěny unikátní jména značek a jejich četnosti. Taktéž je uvedeno, zda se v souboru nachází EKG signál a síťové rušení. Následuje dvousloupcová tabulka s indexy vzorků začátku a konce detekovaného úseku s artefakty a rušením. V poslední části je uvedeno číslo kanálu s nejnižším normalizovaným poměrem - kanálem s artefakty a rušením.

5 Vyhodnocení detekce artefaktů

Nástroj byl implementován v prostředí MATLAB R2022b na přenosném počítači s procesorem Intel Core i7-4710HQ taktovaném na 2,50 GHz s 16 GB paměti RAM. Nástroj je implementován pomocí samostatně spustitelného skriptu `main.m` využívající toolbox s funkcemi.

Na poskytnutá EEG data z laboratoře multimodálního a funkčního zobrazování CEITEC MU byla aplikován základní průzkum a zjištěné vlastnosti dat jsou uvedeny v tabulce 5.1. Mezi hledané atributy patří počet kanálů v záznamu, vzorkovací frekvence f_{vz} , počet vzorků signálu a přítomnost značek a EKG signálu.

Tab. 5.1: Tabulka s veškerými poskytnutými soubory a jejich vlastnostmi.

ID souboru	Počet kanálů	f_{vz}	Počet vzorků	Značky	Přítomnost EKG
1	258	1000	211490	ANO	NE
2	258	1000	213745	ANO	ANO
3	258	1000	213595	ANO	ANO
4	258	1000	212691	ANO	ANO
5	258	1000	212652	ANO	ANO
6	258	1000	212159	ANO	ANO
7	258	1000	212479	ANO	ANO
8	258	1000	210927	ANO	ANO
9	258	1000	215159	ANO	ANO
10	258	1000	211591	ANO	ANO
11	258	1000	210961	ANO	ANO
12	258	1000	210401	ANO	ANO
13	258	1000	210975	ANO	ANO
14	258	1000	214494	ANO	ANO
15	258	1000	213861	ANO	ANO
16	258	1000	212603	ANO	ANO
17	258	1000	213856	ANO	ANO

Z tabulky vyplývá, že veškeré soubory mají stejný počet kanálů, vzorkovací frekvenci, obsahují značky a mají přibližně stejnou dobu trvání. Jediný první soubor se liší tím, že neobsahuje EKG signál v datech. První soubor se také proti zbytku liší tím, že se jedná o formát BVFF oproti ostatním ve formátu MFF.

Další analýzou dat byly získány informace v tabulce 5.2. Cílem této analýzy bylo nalezení unikátních jmen značek a jejich četností spolu s kanálem EKG. Například

zkratka značky DI16 označuje digitální vstupní kanál označený jako kanál 16. Znamená, že v daném časovém bodě během záznamu EEG došlo ke specifické události nebo spuštění. [26]

Tab. 5.2: Tabulka unikátních značek, jejich četností a kanálu EKG.

ID souboru	Jména značek	Četnosti značek	EKG kanál
1	DI16, DIN3	1, 102	-
2	DI16, DIN3	51, 50	258
3	DI16, DIN3	49, 50	258
4	DI16, DIN3	10, 50	258
5	DI16, DIN3	51, 50	258
6	DI16, DIN3	50, 50	258
7	DI16, DIN3	50, 50	258
8	DI16, DIN1, DIN3	46, 1, 50	258
9	DI16, DIN1, DIN3	50, 1, 50	258
10	DI16, DIN3	50, 50	258
11	DI16, DIN3	40, 50	258
12	DI16, DIN3	48, 50	258
13	DI16, DIN3	50, 50	258
14	DI16, DIN3	47, 50	258
15	DI16, DIN3	50, 50	258
16	DI16, DIN1, DIN3	50, 1, 50	258
17	DI16, DIN3	50, 50	258

5.1 Detekce artefaktových úseků signálu

První krokem otestování navrženého nástroje bylo ověření detekce špatných úseků signálu. Pro každý datový soubor byly poskytnuté soubory `.mrk` s vyznačenými počátky a konci artefaktů expertem. Platí, že úseky s artefakty se budou nacházet přes všechny kanály. Úpravou vznikla matice o velikosti $M \times N$, kde M je počet kanálů a N je počet vzorků signálu. Pro zachování validity jsou anotovaná data délkově zkrácena podle vztahu 4.2 na stejnou délku jako výstupní data z nástroje. Úseky s artefakty jsou ohodnoceny nulou, zbytek s hodnotou 1. Takto vzniklý binární obraz je porovnán s daty nástroje.

Výstupem nástroje jsou hodnoty odpovídající indexu okna detekovaného jako špatné. Pomocí transformace je každý index okna změněn na úsek signálu. Opět

vzniká černobílý obraz, který je porovnán s obrazem za účelem získání vizuálního výsledku. Po vizuální kontrole pro každou dvojici originální soubor - anotovaná data byla spočítána matice záměn, jejíž výsledky jsou uvedeny v tabulce 5.3. [20]

Hodnoty z matice záměn jsou následující: skutečně pozitivní (TP) - anotovaný artefaktový vzorek signálu detekovaný jako artefakt, falešně negativní (FN) - anotovaný artefaktový vzorek signálu nebyl detekovaný jako artefakt, falešně pozitivní (FP) - anotovaný vzorek signálu bez rušení detekovaný jako artefaktový a skutečně negativní (TN) - anotovaný vzorek signálu bez rušení nedetekovaný jako artefakt. Veškeré hodnoty jsou poděleny délkou signálu a jsou uvedeny v tabulce 5.3. Z původních hodnot jsou spočítány základní metriky úspěšnosti, mezi které patří senzitivita (TPR), specifická (TNR) a F1 skóre. [20]

TPR určuje pravděpodobnost pozitivního výsledku testu. Měří podíl detekovaných artefaktových vzorků vůči všem anotovaným vzorkům s artefaktem. Výpočet je možné provést pomocí vztahu

$$TPR = \frac{TP}{TP + FN}. \quad (5.1)$$

Zatímco TNR vyjadřuje pravděpodobnost negativního výsledku testu. V případě této práce jsou porovnány množství detekovaných vzorků bez rušení vůči množství anotovaných vzorků bez rušení. Pomocí vztahu 5.2 je možné vypočítat hodnotu TNR. [20]

$$TNR = \frac{TN}{TN + FP} \quad (5.2)$$

Jako měřítko přesnosti detekce bylo použito F1 skóre. To je dáno vztahem mezi dvěma veličinami: přesnost a úplnost. Vysoká přesnost detektoru pro případ této práce znamená, že nenalezne veškeré artefaktové vzorky signálu, ale bude minimalizovaný počet vzorků bez rušení detekovaných jako artefaktové. Zatímco vysoká úplnost nalezne veškeré artefaktové vzorky signálu i za cenu vyšší detekce vzorků bez rušení detekovaných jako artefaktové. Ze vzorků pro výpočet přesnosti a úplnosti je definovaný vztah pro výpočet F1 skóre. [20]

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (5.3)$$

Z výsledků v tabulce 5.3 jsou poté získány průměrné hodnoty metriky úspěšnosti $\overline{TPR} = 0,7603$, $\overline{TNR} = 0,7887$ a $\overline{F1} = 0,3050$. Průměrná TPR znamená, že detektor správně identifikoval přibližně 76 % pozitivních vzorků (artefaktové vzorky) z celkového počtu pozitivních vzorků v souboru dat. Obecně platí, že vyšší hodnoty citlivosti naznačují lepší výkon při správné detekci pozitivní třídy. Průměrná TNR naznačuje, že detektor správně identifikoval přibližně 79 % negativních vzorků (vzorků signálu bez rušení) z celkového počtu negativních vzorků v souboru dat.

Tab. 5.3: Tabulka s hodnotami z matice záměn a výsledků metrik úspěšnosti.

ID souboru	TP	FN	FP	TN	TPR	TNR	F1 skóre
1	0,0363	0,0237	0,0206	0,9195	0,6051	0,9781	0,6209
2	0,1356	0,0593	0,0334	0,7999	0,6959	0,9599	0,7454
3	0,0223	0,0000	0,1561	0,8216	1,0000	0,8404	0,2224
4	0,0184	0,0000	0,4627	0,5189	1,0000	0,5286	0,0738
5	0,1367	0,0040	0,1699	0,6894	0,9714	0,8023	0,6112
6	0,0168	0,0000	0,1247	0,8585	1,0000	0,8732	0,2124
7	0,0961	0,1668	0,0313	0,7050	0,3656	0,9576	0,4925
8	0,0037	0,0052	0,4582	0,5329	0,4177	0,5377	0,0159
9	0,0274	0,0014	0,0703	0,9009	0,9498	0,9276	0,4328
10	0,0212	0,0185	0,0736	0,8868	0,5344	0,9234	0,3151
11	0,0009	0,0000	0,3944	0,6048	1,0000	0,6053	0,0044
12	0,0009	0,0000	0,1562	0,8429	1,0000	0,8436	0,0116
13	0,0037	0,0000	0,4391	0,5571	1,0000	0,5592	0,0166
14	0,0000	0,0000	0,4626	0,5374	0,0000	0,5374	0,0000
15	0,0798	0,0028	0,1690	0,7483	0,9656	0,8158	0,4817
16	0,0707	0,0774	0,1369	0,7151	0,4774	0,8393	0,3975
17	0,0660	0,0041	0,1124	0,8175	0,9415	0,8791	0,5312

Vyšší hodnoty specificity indikují lepší výkon při správné klasifikaci negativní třídy. Průměrné F1 skóre představuje harmonický průměr přesnosti a úplnosti, taktéž naznačuje relativně nízkou celkovou výkonnost detektoru z hlediska vyvážení přesnosti a úplnosti. [20]

Celkově se na základě daných metrik zdá, že detektor má vyšší specificitu než citlivost, což ukazuje na lepší výkon při správné klasifikaci negativní třídy ve srovnání s pozitivní třídou. Navíc relativně nízké skóre F1 naznačuje, že může existovat prostor pro zlepšení, což lze provést přidáním extrakce dalšího příznaku. Možným důvodem nízké hodnoty F1 skóre může být fakt, že při práci se signál v kódu nástroje je rozdělen na vzájemně nepřekrývající se okna. Tudíž pokud počátek nebo konec artefaktu se nachází mezi vícero okny, krajní okna nebudou detekována. Dalším možným důvodem vysoký šum a rychlé změny hodnot napětí na začátku a konci měření. [20]

5.2 Detekce artefaktových kanálů

Druhým krokem otestování navrženého nástroje bylo ověření detekce artefaktových kanálů. K tomu byl poskytnut soubor s anotovaným artefaktovým kanálem. Do tabulky 5.4 jsou uvedeny anotované artefaktové kanály, jejich hodnota normalizovaného poměru, detekovaný kanál a jeho hodnota. Jak bylo zmíněno v kapitole 4.6, tak kanál je vybrán podle nejnižší hodnoty normalizovaného poměru.

Tab. 5.4: Tabulka anotovaných a detekovaných kanálů a hodnot normalizovaného poměru.

ID souboru	Anotovaný kanál	Hodnota poměru	Detekovaný kanál	Hodnota poměru
1	136	0,4119	136	0,4119
2	1	0,6957	25	0,3836
3	-	-	174	0,2503
4	-	-	70	0,2523
5	-	-	33	0,3056
6	-	-	131	0,4180
7	-	-	95	0,4100
8	-	-	162	0,2703
9	165	0,7521	85	2673
10	-	-	152	0,2400
11	48, 49, 55	0,5791, 0,3205, 0,2505	55	0,2505
12	-	-	95	0,2698
13	47, 48	0,3564, 0,3865	161	0,2502
14	-	-	222	0,3145
15	-	-	21	0,3202
16	256	0,7169	16	0,3018
17	209	0,7027	90	0,2359

Z výsledků v tabulce je patrné, že detekce artefaktových kanálů je nedostatečná, ač detekce úseků signálu s rušením není. Z tabulky vyplývá, že zvolené metriky nedokáží významně rozdělit mezi artefaktovým kanálem a tím bez rušení. Prostor pro zlepšení se nabízí extrakcí dalšího příznaku, který by lépe dokázal rozlišit mezi třídami.

Závěr

Cílem této práce bylo provést rešerši základních poznatků o elektroencefalografickém signálu. Důležitými body byl vznik a šíření signálu a jeho propagace v signálu. V další části práce byl popsán základní postup při předzpracování dat a jeho modifikace použitím tohoto navrženého nástroje. Následoval popis zdrojů artefaktů a rušení těchto dat. Hlavním cílem teoretické části práce byla rozsáhlá rešerše používaných metod a postupů pro odhalení artefaktů v signálu.

Pomocí těchto poznatků je v začátku teoretické části definovaný nástroj pro automatické hodnocení kvality elektroencefalografických dat s důrazem na pozdější rozšířitelnost. Smyslem nástroje je nalézt a doporučit úseky signály či celé kanály s podezřením na obsah artefaktů a rušení.

Ověření je provedeno na anotovaných datech z laboratoře CEITEC MU. Z výsledků vyplývá, že navržený nástroj dokáže detekovat artefaktové úseky, avšak kanály už ne tolik úspěšně. Vzhledem ke konstrukci nástroje je možné dosáhnout lepších výsledků přidáním extrakce dalších příznaků, nebo modifikací detekčního algoritmu.

Literatura

- [1] ABE, Shigeo. *Support Vector Machines for Pattern Classification*. 2nd ed. 2010. London: Springer London, 2010. Advances in Computer Vision and Pattern Recognition. ISBN 978-1-84996-098-4.
- [2] ABHANG, Priyanka A., Bharti W. GAWALI a Suresh C. MEHROTRA. *Introduction to EEG- and Speech-Based Emotion Recognition*. London: Elsevier Health Sciences, 2016. ISBN 978-0-12-804490-2.
- [3] AGGARWAL, Charu C. *Data Mining: The Textbook*. Cham: Springer International Publishing, 2015. ISBN 978-3-319-14142-8.
- [4] ALLEN, Ronald L. a Duncan W. MILLS. *Signal analysis: time, frequency, scale, and structure*. Piscataway: IEEE, 2004. ISBN 04-712-3441-9.
- [5] BLINOWSKA, Katarzyna J., Rafał KUŚ a Maciej KAMIŃSKI. Granger causality and information flow in multivariate processes. *Physical Review E*. 2004, **70**(5). ISSN 1539-3755. Dostupné z: doi:10.1103/PhysRevE.70.050902
- [6] BREUNIG, Markus M., Hans-Peter KRIEGEL, Raymond T. NG a Jörg SANDER. LOF. *ACM SIGMOD Record*. 2000, **29**(2), 93-104. ISSN 0163-5808. Dostupné z: doi:10.1145/335191.335388
- [7] ČECH, Svatopluk a Drahomír HORKÝ. *Přehled obecné histologie*. 2., přeprac. vyd. Brno: Masarykova univerzita, 2011. ISBN 978-80-210-5543-8.
- [8] FICKLING, Shaun D., Careesa C. LIU, Ryan C. N. D'ARCY, Sujoy GHOSH HAJRA a Xiaowei SONG. Good data? The EEG Quality Index for Automated Assessment of Signal Quality. *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2019, 2019, 0219-0229. ISBN 978-1-7281-2530-5. Dostupné z: doi:10.1109/IEMCON.2019.8936246
- [9] GEOCADIN, Romergryko G., Jitendran MUTHUSWAMY, David L. SHERMAN, Nitish V. THAKOR a Daniel F. HANLEY. Early electrophysiological and histologic changes after global cerebral ischemia in rats. *Movement Disorders*. 2000, **15**(S1), 14-21. ISSN 08853185. Dostupné z: doi:10.1002/mds.870150704
- [10] HEGGER, Rainer a Holger KANTZ. Improved false nearest neighbor method to detect determinism in time series data. *Physical Review E*. 1999, **60**(4), 4970-4973. ISSN 1063-651X. Dostupné z: doi:10.1103/PhysRevE.60.4970

- [11] HIRSCH, Lawrence J. a Richard P. BRENNER. *Atlas of EEG in Critical Care*. Chichester: John Wiley & Sons, 2010. ISBN 978-0-470-98786-5.
- [12] JAN, Jiří. *Medical image processing, reconstruction and restoration: concepts and methods*. 2. vydání. Boca Raton: Taylor & Francis, 2006. ISBN 0-8247-5849-8.
- [13] KALAYCI, Ilker a Tuncay ERCAN. Anomaly Detection in Wireless Sensor Networks Data by Using Histogram Based Outlier Score Method. *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2018, 2018, 1-6. ISBN 978-1-5386-4184-2. Dostupné z: doi:10.1109/ISMSIT.2018.8567262
- [14] KANIUSAS, Eugenijus. *Biomedical signals and sensors III: linking electric bio-signals and biomedical sensors*. Cham: Springer, [2019]. Biological and medical physics, biomedical engineering. ISBN 978-3-319-74916-7.
- [15] KAPLAN, Peter W. a Thien NGUYEN. *Clinical Electrophysiology: A Handbook for Neurologists*. Chichester: Wiley-Blackwell, 2011. ISBN 978-1-405-18529-5.
- [16] KRAJČA, Vladimír a Jitka MOHYLOVÁ. *Číslíkové zpracování neurofyziologických signálů*. V Praze: České vysoké učení technické, 2011. ISBN 978-80-01-04721-7.
- [17] KRIEGEL, Hans-Peter, Matthias SCHUBERT a Arthur ZIMEK. Angle-based outlier detection in high-dimensional data. *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA: ACM, 2008, 2008-08-24, 444-452. ISBN 9781605581934. Dostupné z: doi:10.1145/1401890.1401946
- [18] LUU, Phan, Zhongqing JIANG, Catherine POULSEN, Chelsea MATTSON, Anne SMITH a Don M. TUCKER. Learning and the Development of Contexts for Action. *Frontiers in Human Neuroscience*. 2011, **5**. ISSN 1662-5161. Dostupné z: doi:10.3389/fnhum.2011.00159
- [19] MARCUSE, Lara V., Madeline C. FIELDS a Jiyeoun YOO. *Rowan's Primer of EEG*. 2nd. Edinburgh: Elsevier Health Sciences, 2016. ISBN 978-0-323-35387-8.
- [20] MELOUN, Milan a Jiří MILITKÝ. *Kompendium statistického zpracování dat: metody a řešené úlohy včetně CD*. Praha: Academia, 2002. ISBN 80-200-1008-4.
- [21] MICHEL, Christoph M., Thomas KOENIG, Daniel BRANDEIS, Lorena R.R. GIANOTTI a Jiří WACKERMANN. *Electrical Neuroimaging*. New York: Cambridge University Press, 2009. ISBN 978-0-521-87979-8.

- [22] MORNSTEIN, Vojtěch. *Lékařská fyzika a biofyzika*. Brno: Masarykova univerzita, 2020. ISBN 978-80-210-8984-6.
- [23] MOUREK, Jindřich. *Fyziologie: učebnice pro studenty zdravotnických oborů*. 2., dopl. vyd. Praha: Grada, 2012. Sestra (Grada). ISBN 978-80-247-3918-2.
- [24] NICHOLLS, John G. *Od neuronu k mozku*. Praha: Academia, 2013. ISBN 978-80-200-2155-7.
- [25] OH, Seung-Hyeon, Yu-Ri LEE a Hyoung-Nam KIM. A Novel EEG Feature Extraction Method Using Hjorth Parameter. *International Journal of Electronics and Electrical Engineering*. 2014, 106-110. ISSN 2301380X. Dostupné z: doi:10.12720/ijeee.2.2.106-110
- [26] OOSTENVELD, Robert, Pascal FRIES, Eric MARIS a Jan-Mathijs SCHOFFELEN. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*. 2011, **2011**, 1-9. ISSN 1687-5265. Dostupné z: doi:10.1155/2011/156869
- [27] OPPENHEIM, A.V. a R.W. SCHAFER. Dsp history - From frequency to quefrequency: a history of the cepstrum. *IEEE Signal Processing Magazine*. 2004, **21**(5), 95-106. ISSN 1053-5888. Dostupné z: doi:10.1109/MSP.2004.1328092
- [28] SABA-SADIYA, Sari, Eric CHANTLAND, Tuka ALHANAI, Taosheng LIU a Mohammad M. GHASSEMI. Unsupervised EEG Artifact Detection and Correction. *Frontiers in Digital Health*. 2021, **2**. ISSN 2673-253X. Dostupné z: doi:10.3389/fdgth.2020.608920
- [29] SANEI, Saeid a Jonathon A. CHAMBERS. *EEG Signal Processing*. Chippenhams: Wiley-Interscience, 2008. ISBN 978-0-470-02581-9.
- [30] SHIN, Hyun-Chool, Xiaofeng JIA, Robert NICKL, Romergryko G. GEOCADIN a Nitish V. THAKOR. A Subband-Based Information Measure of EEG During Brain Injury and Recovery After Cardiac Arrest. *IEEE Transactions on Biomedical Engineering*. 2008, **55**(8), 1985-1990. ISSN 0018-9294. Dostupné z: doi:10.1109/TBME.2008.921093
- [31] SIULY, Siuly, Yan LI a Yanchun ZHANG. *EEG Signal Analysis and Classification: Techniques and Applications*. Cham: Springer International Publishing, 2017. ISBN 978-3-319-47652-0.
- [32] SÖRNMO, Leif a Pablo LAGUNA. *EEG Signal Processing. Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier, 2005, 2005, 55-179. ISBN 9780124375529. Dostupné z: doi:10.1016/B978-012437552-9/50003-9

- [33] STAM, Cornelis J., Guido NOLTE a Andreas DAFFERTSHOFER. Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping*. 2007, **28**(11), 1178-1193. ISSN 10659471. Dostupné z: doi:10.1002/hbm.20346
- [34] SVATOŠ, Josef. *Biologické signály I: geneze, zpracování a analýza*. Vyd. 2. Praha: České vysoké učení technické, 1998. ISBN 80-010-1822-9.
- [35] TJEPKEMA-CLOOSTERMANS, Marleen C, Fokke B VAN MEULEN, Gjerrit MEINSMA a Michel JAM VAN PUTTEN. A Cerebral Recovery Index (CRI) for early prognosis in patients after cardiac arrest. *Critical Care*. 2013, **17**(5). ISSN 1364-8535. Dostupné z: doi:10.1186/cc13078
- [36] TŮMA, Jiří. *Zpracování signálů získaných z mechanických systémů užitím FFT*. Praha: Sdělovací technika, 1997. ISBN 80-901-9361-7.
- [37] WOLF, Alan, Jack B. SWIFT, Harry L. SWINNEY a John A. VASTANO. Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*. 1985, **16**(3), 285-317. ISSN 01672789. Dostupné z: doi:10.1016/0167-2789(85)90011-9
- [38] ZHAO, Yue, Zain NASRULLAH, Maciej K. HRYNIEWICKI a Zheng LI. LSCP: Locally Selective Combination in Parallel Outlier Ensembles. *Proceedings of the 2019 SIAM International Conference on Data Mining*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2019, 2019-05-06, 585-593. ISBN 978-1-61197-567-3. Dostupné z: doi:10.1137/1.9781611975673.66
- [39] ZHAO, Lingling, Yufan ZHANG, Xue YU, et al. Quantitative signal quality assessment for large-scale continuous scalp EEG from a big data perspective. *Physiological Measurement*. 2022. ISSN 0967-3334. Dostupné z: doi:10.1088/1361-6579/ac890d

Seznam zkratek

ABOD	Konečná impulzní odezva (z angl. Angle-Based Outlier Detector)
A/D	Analogově-digitální
AP	Akční potenciál
CMRR	Činitel potlačení soufázového zesílení (z angl. Common-Mode Rejection Ratio)
CNS	Centrální nervová soustava
ECoG	Elektrokortikografie
EEG	Elektroencefalografie
EMG	Elektromyografie
EOG	Elektrookulografie
EPSP	Excitační postsynaptický potenciál
FIR	Konečná impulzní odezva (z angl. Finite Impulse Response)
ICA	Analýza nezávislých komponent (z angl. Independent Component Analysis)
iEEG	Intrakraniální elektroencefalografie
FN	falešně negativní (z angl. False Negative)
FP	Falešně pozitivní (z angl. False Positive)
FIR	Konečná impulzní odezva (z angl. Finite Impulse Response)
IPSP	Inhibiční postsynaptický potenciál
LOF	Lokální odlehlý faktor (z angl. Local Outlier Factor)
LSCPOE	Lokálně selektivní kombinace v paralelních odlehlých souborech (z angl. Locally Selective Combination in Parallel Outlier Ensembles)
MAD	Konečná impulzní odezva (z angl. Finite Impulse Response)
MG	Konečná impulzní odezva (z angl. Finite Impulse Response)
OCSVM	Jednotřídová metoda podpůrných vektorů (z angl. One Class Support Vector Machine)

PLV	Hodnota fázového uzamčení (z angl. Phase Lock Value)
RMSA	Konečná impulzní odezva (z angl. Finite Impulse Response)
SD	Směrodatná odchylka(z angl. Standard Deviation)
RMSA	Konečná impulzní odezva (z angl. Finite Impulse Response)
TN	Konečná impulzní odezva (z angl. Finite Impulse Response)
TNR	Konečná impulzní odezva (z angl. Finite Impulse Response)
TP	Konečná impulzní odezva (z angl. Finite Impulse Response)
TPR	Konečná impulzní odezva (z angl. Finite Impulse Response)
ZCR	Míra průchodu nulou(z angl. Zero-Crossing Rate)

Seznam příloh

A Obsah elektronické přílohy

52

A Obsah elektronické přílohy

/	kořenový adresář přiloženého archivu
+func	složka s funkcemi extrakce příznaků
_ A_StationarySignal.m	složka s funkcemi extrakce příznaků
_ Kurtosis.m	složka s funkcemi extrakce příznaků
_ LineNoise.m	složka s funkcemi extrakce příznaků
_ MaximumGradient.m	složka s funkcemi extrakce příznaků
_ RMSAmplitude.m	složka s funkcemi extrakce příznaků
_ WavePower.m	složka s funkcemi extrakce příznaků
_ ZeroCrossingRate.m	složka s funkcemi extrakce příznaků
ChannelComputing.m	funkce detekce artefaktových kanálů
EventDetection.m	funkce detekce značek v signálu
FeatureExtraction.m	funkce extrakce příznaků
FiltECG.m	funkce detekce a odstranění EKG signálu
LoadData.m	funkce načtení dat a toolboxu
main.m	hlavní spouštěcí skript
Preprocessing.m	funkce předzpracování dat
PrimaryParameters.m	funkce vyhodnocení příznaků
SignalQuality.pdf	výsledný .pdf report kvality EEG dat
WindowComputing.m	funkce detekce artefaktových úseků signálu