



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

FAKULTA BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

Katedra biomedicínské techniky

**Návrh metodiky automatizované klasifikace EEG signálu
s využitím adaptivní segmentace**

**Methodology proposal for EEG automated signal classification with
adaptive segmentation**

Diplomová práce

Studijní program: Biomedicínské inženýrství

Vedoucí práce: Ing. Jakub Šebek

Bc. Jaromír Chaloupek

Kladno 2023

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Chaloupek** Jméno: **Jaromír** Osobní číslo: **465393**
Fakulta: **Fakulta biomedicínského inženýrství**
Garantující katedra: **Katedra biomedicínské techniky**
Studijní program: **Biomedicínské inženýrství**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Návrh metodiky automatizované klasifikace EEG signálu s využitím adaptivní segmentace

Název diplomové práce anglicky:

Methodology proposal for EEG automated signal classification with adaptive segmentation

Pokyny pro vypracování:

Cílem této práce je vytvoření metodiky automatizované klasifikace reálného EEG záznamu. S pomocí programovacího prostředí MATLAB navrhnete metodu automatické klasifikace nativního záznamu EEG s využitím adaptivní segmentace, jež využívá analýzy v časové a frekvenční oblasti. Navrhnete příznaky, které bude možné využít pro metody umělé inteligence. Návrh metodiky otestujte a výsledky statisticky vyhodnoťte.

Seznam doporučené literatury:

- [1] Krajča V., Mohylová J., Číslíkové zpracování neurofyziologických signálů, ed. 1st, ČVUT Praha, 2011, 167 s., ISBN 978-80-01-04721-7
- [2] SANEI, Saeid a Jonathon CHAMBERS, EEG signal processing, John Wiley, c2007, ISBN 9780470025819
- [3] Mike X Cohen, Analyzing Neural Time Series Data: Theory and Practice, ed. 1st Edition, The MIT Press, 2014, ISBN 9780262019873

Jméno a příjmení vedoucí(ho) diplomové práce:

Ing. Jakub Šebek

Jméno a příjmení konzultanta(ky) diplomové práce:

Ing. Hana Děcká

Datum zadání diplomové práce: **14.02.2022**

Platnost zadání diplomové práce: **22.09.2023**

doc. Ing. Martin Rožánek, Ph.D.
vedoucí katedry

prof. MUDr. Jozef Rosina, Ph.D., MBA
děkan



PROHLÁŠENÍ

Prohlašuji, že jsem diplomovou práci s názvem „Návrh metodiky automatizované klasifikace EEG signálu s využitím adaptivní segmentace“ vypracoval samostatně a použil k tomu úplný výčet citací použitých pramenů, které uvádím v seznamu přiloženém k diplomové práci.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu § 60 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.

V Kladně dne 17. 5. 2023

.....

Bc. Jaromír Chaloupek

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu své práce, Ing. Jakubu Šebkovi, za pomoc, cenné rady a odborné vedení v průběhu řešení mé diplomové práce. Dále bych chtěl poděkovat Ing. Haně Děcké za konzultace a věcné připomínky k mojí práci. Dalším, komu bych rád poděkoval je Joe Picone s celým týmem Temple University Hospital, který mi umožnil použít EEG data z jejich databáze.

ABSTRAKT

Název práce: Návrh metodiky automatizované klasifikace EEG signálu s využitím adaptivní segmentace

Vizuální vyhodnocování EEG záznamu kontaminovaného různými druhy artefaktů je v zásadě ve všech oblastech (klinických i komerčních) časově velmi náročným a vyčerpávajícím úkolem. Proto stále přetrvává snaha o vytvoření automatizovaného systému, který by svým příspěvkem urychlil a zefektivnil práci specialistů. Prvním nezbytným krokem takového systému je efektivní adaptivní segmentace. Výsledkem této práce je návrh nové metodiky adaptivní segmentace EEG záznamu, která využívá k dosažení finálního výsledku segmentace kombinování výsledků analýz různě dlouhých oken, ve kterých jsou prováděny analýzy v časové a frekvenční oblasti signálu. Celý algoritmus navržené adaptivní segmentace je založen na individuální definici běžné činnosti mozku. Navržená metoda byla realizována v programovém prostředí MATLAB 2022a, ve kterém vykazovala hladký průběh s nízkou náročností na výpočetní výkon. Následně byla vyhodnocena skrze výkonnostní metriky Index/Score a Přesnost. Při testu na různých typech simulovaných dat dosáhla metoda hodnot přesnosti přes 90 % v obou metrikách průměrně v 15 z 32 testovaných nastavení. Při testu na anotovaných reálných datech z databáze Temple University Hospital dosáhlo 92 z 240 testovaných nastavení v obou metrikách hodnot přesnosti přes 90 % úspěšnosti. V rámci práce byly navrženy příznaky pro metody založené na kooperaci s UI, které by měly přispět k další optimalizaci nastavení parametrů navržené metody adaptivní segmentace a dále pak pro navazující úkol klasifikace získaných segmentů.

Klíčová slova

Adaptivní segmentace, NEP, AEP, příznaky pro umělou inteligenci

ABSTRACT

The title of the Thesis: Metodology proposal for EEG automated signal classification with adaptive segmentation

Visual evaluation of an EEG recording contaminated with various types of artifacts is basically a very time-consuming and exhausting task in all areas (clinical and commercial), therefore there is still an ongoing effort to create an automated system that would speed up and make the work of specialists more efficient with its contribution. The first necessary step of such a system is effective adaptive segmentation. The result of this work is a proposal of a new methodology of adaptive segmentation of the EEG recording, which to achieve the result uses combining of the results of analyses with windows of different lengths, in which analyzes are performed in the time and frequency domain of the signal. The entire algorithm of the proposed adaptive segmentation is based on an individual definition of normal brain activity. The proposed method was implemented in the MATLAB 2022a program environment, in which showed a smooth progress with a low demand on computing power and subsequently evaluated through the Index/Score and Accuracy metrics. When tested on different types of simulated data, the method achieved values over 90% in both metrics in 15 out of 32 settings tested. When tested on annotated real data, 92 out of 240 tested settings achieved over 90% success rates in both metrics. As part of the work, features were proposed for methods based on cooperation with UI, which should serve to further optimize the parameter settings of the proposed adaptive segmentation method and then for the subsequent task of classifying the obtained segments.

Keywords

Adaptive segmentation, NEP, AEP, features for artificial intelligence

Obsah

Seznam symbolů a zkratk	9
Seznam obrázků	10
Seznam tabulek	15
1 Úvod	16
2 Přehled současného stavu	18
2.1 Úvod do elektroencefalografie.....	18
2.1.1 Neurofyziologický původ.....	19
2.1.2 EEG rytmy.....	21
2.1.3 Artefakty v EEG záznamu.....	23
2.2 Způsob měření EEG v praxi.....	33
2.2.1 Přístroje.....	33
2.2.2 Elektrody pro měření EEG.....	33
2.2.3 Rozmístění elektrod na skalpu.....	35
2.2.4 Režimy zapojení elektrod.....	37
2.2.5 Simultánní měření biosignálů pro detekci artefaktů.....	39
2.2.6 Datové formáty EEG.....	40
2.3 Používané metody.....	42
2.3.1 Metody pro segmentaci EEG signálu.....	42
2.3.2 Metody pro klasifikaci EEG signálu.....	48
2.3.3 Neuronové sítě v EEG.....	53
3 Cíle práce	54
4 Metodika	55
4.1 Hypotetický základ a princip navrhované metody.....	55
4.1.1 Hypotéza 1: Komplementární efekt různě dlouhých oken.....	55
4.1.2 Hypotéza 2: Abnormalitu lze definovat pouze normalitou.....	56
4.1.3 Hypotéza 3: Hranice mezi NEP a AEP není bod, ale časový úsek....	57
4.2 Technické řešení navrhované metody adaptivní segmentace.....	57
4.2.1 Sledované parametry.....	57
4.2.2 Schéma navržené metody adaptivní segmentace.....	64
4.2.3 Fáze 1: Import dat a předzpracování.....	65

4.2.4	Fáze 2: Definice etalonu NEP v kanálu.....	66
4.2.5	Fáze 3: Detekce NEP	67
4.2.6	Sloučení výsledků jednotlivých oken – klasifikace NEP a AEP.....	68
4.3	Statistické zpracování.....	69
4.4	Použitá data	70
4.4.1	Simulovaný syntetický signál (SIM_SYNT).....	70
4.4.2	Simulovaný semisyntetický signál (SIM_SEMISYNT)	75
4.4.3	Temple University Hospital Data Corpus (TUH)	80
4.4.4	Simulovaná reálná data (SIM_REAL1 a SIM_REAL2).....	81
4.4.5	Anotovaná reálná data (REAL_ANOT).....	81
5	Výsledky.....	82
5.1	Výsledky simulovaných dat	82
5.1.1	Popis nastavení parametrů metody pro simulovaná data	82
5.1.2	Výsledky SIM_SYNT	83
5.1.3	Výsledky SIM_SEMISYNT.....	84
5.1.4	Výsledky SIM_REAL1	85
5.1.5	Výsledky SIM_REAL2	86
5.1.6	Grafická ukázka výsledků na simulacích pro nejlepší kandidáty nastavení.....	87
5.2	Výsledky anotovaných reálných dat	91
5.2.1	Sumarizace výsledků pro různá nastavení citlivosti metody (SEN) ..	92
5.2.2	Sumarizace výsledků pro jednotlivá okna a jejich sloučení.....	97
5.2.3	Analýza v prostředí Power BI pro metriku Index/Score	101
5.2.4	Grafická ukázka výsledků anotovaných reálných dat jednoho z nejlepších kandidátů nastavení	103
5.3	Výpočetní čas na testovaných signálech	105
5.4	Návrh příznaků pro UI	105
6	Diskuse.....	106
7	Závěr.....	110
8	Reference	111
	Příloha A: Obsah přiložených souborů	117
	Příloha B: Obsah přiloženého CD/DVD	118

Seznam symbolů a zkratk

Seznam zkratk

Zkratka	Význam
AEP	Abnormal Event Part (abnormální aktivita v EEG záznamu)
BIDS	Brain Imaging Data Structure
DFT	Diskrétní Fourierova transformace
EDF	European Data Format
FFT	Rychlá Fourierova transformace
FN	Chybně klasifikováno jako negativní („False Negative“)
FP	Chybně klasifikováno jako pozitivní („False Positive“)
INT	Parametr nastavení navržené metody. Násobek směrodatné odchylky při výpočtu intervalu NEP etalonu.
KURT	Koeficient špičatosti („kurtosis“)
MRI	Magnetická rezonance
NEP	Normal Event Part (klidová mozková aktivita v EEG záznamu)
NPV	Prediktivní hodnota negativního testu („Negative Predictive Value“)
PPV	Prediktivní hodnota pozitivního testu („Positive Predictive Value“)
SEN	Parametr nastavení navržené metody. Citlivost metody v %. Odráží, kolik % sledovaných parametrů musí být splněno pro klasifikaci signálu jako NEP.
SKEW	Koeficient šikmosti („skewness“)
STD	Parametr nastavení navržené metody. Násobek směrodatné odchylky při výpočtu selektivního prahu
TN	Správně klasifikováno jako negativní („True Negative“)
TP	Správně klasifikováno jako pozitivní („True Positive“)
TUAR	Temple University Hospital Artifact Corpus
TUH	Temple University Hospital

Seznam obrázků

1. Obrázek 2.1. Ilustrace měření EEG záznamů. Převzato z [50]	18
2. Obrázek 2.2. Znázornění akčního potenciálu a buňky neuronu. Obrázek převzat z [11]	19
3. Obrázek 2.3. Ilustrace běžných rytmů v EEG záznamu. Převzato z [50]....	21
4. Obrázek 2.4. Prasknutí elektrody. Převzato z [17]	24
5. Obrázek 2.5. EEG záznam zarušený síťovým šumem o frekvenci 50 Hz (nahore, červeně). EEG záznam po odstranění síťového šumu filtraceí (dole, modře). Obrázek převzat z [21].....	25
6. Obrázek 2.6. Část svalového artefaktu indikovaná morfologií vysoké frekvence a střední až vysoké amplitudy. Převzato z [17]	26
7. Obrázek 2.7. Ukázka artefaktu chvění. Převzato z [17]	27
8. Obrázek 2.8. Žvýkácký artefakt. Převzato z [17].....	28
9. Obrázek 2.9. Jedno mrknutí očí. Převzato z [17]	29
10. Obrázek 2.10. Série mrknutí očí, která může indikovat neurologickou poruchu. Převzato z [17].....	29
11. Obrázek 2.11. Jeden laterální pohyb očí. Převzato z [17]	30
12. Obrázek 2.12. Série pohybu očí. Převzato z [17]	30
13. Obrázek 2.13. EKG artefakt v EEG záznamu (červeně) spolu s reálným záznamem EKG (ECG) a EOG. Obrázek převzat z [21].....	31
14. Obrázek 2.14. Pulsní artefakt (vyznačen červeně) v EEG záznamu pod elektrodou F4 u pacienta se srdeční arytmií. Obrázek převzat z [9].....	32
15. Obrázek 2.15. Rozmístění elektrod ve standardním systému 10/20. Obrázek A reprezentuje boční pohled, obrázek B pohled shora. Obrázek převzat z [22]	35
16. Obrázek 2.16. Ukázka EEG čepice využívané při měření HD EEG. Obrázek převzat z [51]	36
17. Obrázek 2.17. Zapojení elektrod v unipolárním režimu. Obrázek převzat z [25].....	37
18. Obrázek 2.18. Zapojení elektrod v bipolárním režimu. Obrázek převzat z [25]	38
19. Obrázek 2.19. Zapojení elektrod ve zprůměrovaném režimu. Obrázek převzat z [25].....	39

20. Obrázek 2.20. Ukázka struktury datových formátů splňující standardy BIDS. Obrázek převzat z [27].....	40
21. Obrázek 2.21. Ukázka adresářového stromu datové sady BIDS, která obsahuje data EEG. Obrázek převzat z [52].....	41
22. Obrázek 2.22. Rozdíl mezi fixní (nahore) a adaptivní (dole) segmentací. Obrázek převzat z	43
23. Obrázek 2.23. Schéma algoritmu adaptivní segmentace na základě lineární predikce podle Bodensteina a Praetoriuse. Obrázek převzat z [10]	44
24. Obrázek 2.24. Schéma algoritmu adaptivní segmentace pomocí rozdílu autokorelačních funkcí v referenčním a testovacím okně. Obrázek převzat z [10].....	45
25. Obrázek 2.25. Obecné schéma metody adaptivní segmentace využívající pohyblivého spojeného dvojokna, které se posunuje po signálu. G funkce představuje výslednou diferenci mezi referenčním a testovacím oknem. Obrázek převzat z [10].....	46
26. Obrázek 2.26. Varriho algoritmus (Krajča a kol. 1991) adaptivní segmentace. Popis křivek shora: segmentovaný signál, amplitudová diference, frekvenční diference, G funkce. Obrázek převzat z [10].....	47
27. Obrázek 2.27. Hierarchické shlukování - dendrogram. Obrázek převzat z [35].....	48
28. Obrázek 2.28. Princip metody k-means. Obrázek převzat z [36].....	49
29. Obrázek 2.29. Klasifikace do tříd metodou DBSCAN. Obrázek převzat z [36]	51
30. Obrázek 2.30. Klasifikace do shluků metodou DBSCAN. Obrázek převzat z [36].....	51
31. Obrázek 4.1. Ilustrativní obrázek NEP a AEP.....	56
32. Obrázek 4.2. Porovnání algoritmů DFT a FFT. Obrázek převzat z [41].....	62
33. Obrázek 4.3. Schéma navržené metody adaptivní segmentace	64
34. Obrázek 4.4. Ukázka celého SIM_SYNT signálu s vyznačenými hranicemi syntetických artefaktů.	71
35. Obrázek 4.5. Syntetický signál. Detail přechodu mezi částmi 1 a 2.....	72
36. Obrázek 4.6. Syntetický signál. Detail přechodu mezi částmi 3 a 4.....	72
37. Obrázek 4.7. Syntetický signál. Detail přechodu mezi částmi 5 a 6.....	73
38. Obrázek 4.8. Syntetický signál. Detail přechodu mezi částmi 7 a 8.....	73

39. Obrázek 4.9. Syntetický signál. Detail přechodu mezi částmi 9 a 10.....	74
40. Obrázek 4.10. Syntetický signál. Detail přechodu mezi částmi 11, 12 a 13.	74
41. Obrázek 4.11. Semisyntetický signál s vyznačenými hranicemi syntetických artefaktů.	76
42. Obrázek 4.12. Semisyntetický signál. Detail přechodu mezi částmi 1, 2 a 3. V části 2 je syntetický svalový artefakt.	77
43. Obrázek 4.13. Semisyntetický signál. Detail přechodu mezi částmi 3, 4 a 5. V části 4 je syntetický okulární artefakt.	77
44. Obrázek 4.14. Semisyntetický signál. Detail přechodu mezi částmi 5, 6 a 7. V části 6 je syntetický síťový šum.....	78
45. Obrázek 4.15. Semisyntetický signál. Detail přechodu mezi částmi 7, 8 a 9. V části 8 je syntetický elektrodový artefakt.	78
46. Obrázek 4.16. Semisyntetický signál. Detail přechodu mezi částmi 9, 10 a 11. V části 10 je syntetický okulární artefakt s nízkou amplitudou.....	79
47. Obrázek 4.17. Semisyntetický signál. Detail přechodu mezi částmi 11, 12 a 13. V části 12 je syntetický svalový artefakt s nízkou amplitudou.....	79
48. Obrázek 5.1. Grafické výsledky SIM_SYNT pro SEN = 70 % v kombinaci s Nastavením 1.....	87
49. Obrázek 5.2. Grafické výsledky SIM_SYNT pro SEN = 80 % v kombinaci s Nastavením 2.....	87
50. Obrázek 5.3. Grafické výsledky SIM_SEMISYNT pro SEN = 70 % v kombinaci s Nastavením 1.....	88
51. Obrázek 5.4. Grafické výsledky SIM_SYNT pro SEN = 80 % v kombinaci s Nastavením 2.....	88
52. Obrázek 5.5. Grafické výsledky SIM_REAL1 pro SEN = 70 % v kombinaci s Nastavením 1.....	89
53. Obrázek 5.6. Grafické výsledky SIM_REAL1 pro SEN = 80 % v kombinaci s Nastavením 2.....	89
54. Obrázek 5.7. Grafické výsledky SIM_REAL2 pro SEN = 70 % v kombinaci s Nastavením 1.....	90
55. Obrázek 5.8. Grafické výsledky SIM_REAL1 pro SEN = 80 % v kombinaci s Nastavením 2.....	90

56. Obrázek 5.9. Sumarizace výsledků metody pro citlivost metody SEN = 60 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.	92
57. Obrázek 5.10. Sumarizace výsledků metody pro citlivost metody SEN = 70 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100 %) se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.....	93
58. Obrázek 5.11. Sumarizace výsledků metody pro citlivost metody SEN = 80 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.....	94
59. Obrázek 5.12. Sumarizace výsledků metody pro citlivost metody SEN = 90 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.....	95
60. Obrázek 5.13. Sumarizace výsledků metody pro citlivost metody SEN = 100 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.....	96
61. Obrázek 5.14. Sumarizace výsledků metody pro okno 256 vzorků. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.	97
62. Obrázek 5.15. Sumarizace výsledků metody pro okno 512 vzorků. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.	98
63. Obrázek 5.16. Sumarizace výsledků metody pro okno 1024 vzorků. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.	99
64. Obrázek 5.17. Sumarizace výsledků metody pro sloučení oken. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0- 100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.	100

65. Obrázek 5.18. Výsledky analýzy v prostředí Power BI ve vztahu k metrice Index/Score. Klíčové influencery a identifikované segmenty různých nastavení.	101
66. Obrázek 5.19. Detailní popis identifikovaných segmentů v prostředí Power BI	102
67. Obrázek 5.20. Grafická ukázka výsledků metody na anotovaných reálných datech.	103
68. Obrázek 5.21. Grafická ukázka výsledků metody na anotovaných reálných datech.	104
69. Obrázek 5.22. Grafická ukázka výsledků metody na anotovaných reálných datech.	104

Seznam tabulek

1. Tabulka 2.1. Porovnání vlastností algoritmů k-means a DBSCAN podle [36]..	52
2. Tabulka 4.1. Ukázka rozdílů ve výpočetní náročnosti DFF a FFT v závislosti na velikosti vstupního signálu.....	61
3. Tabulka 4.2. Přehled simulace pro bezartefaktový simulovaný syntetický signál	70
4. Tabulka 4.3. Přehled simulace syntetických artefaktů	70
5. Tabulka 4.4. Přehled částí simulovaného syntetického signálu.....	71
6. Tabulka 4.5. Přehled simulovaných syntetických artefaktů	75
7. Tabulka 4.6. Přehled částí simulovaného semisyntetického signálu	76
8. Tabulka 5.1. Výsledky metody na SIM_SYNT pro testovaná nastavení.	83
9. Tabulka 5.2. Výsledky metody na SIM_SEMISYNT pro testovaná nastavení..	84
10. Tabulka 5.3. Výsledky metody na SIM_REAL1 pro testovaná nastavení.	85
11. Tabulka 5.4. Výsledky metody na SIM_REAL2 pro testovaná nastavení.	86
12. Tabulka 5.5. Výsledky 40 nejlepších nastavení na REAL_ANOT datech, která dosáhla v metrikách Index/Score i Přesnost více než 94 % úspěšnosti	91
13. Tabulka 5.6. Výpočetní čas testovaných signálů.....	105

1 Úvod

EEG (elektroencefalografie) měří napětí mezi elektrodami na skalpu hlavy, a tím zaznamenává elektrickou aktivitu pocházející z mozku. V současnosti je EEG běžnou součástí nejen klinické medicíny, ale i nástrojem pro analýzu behaviorálního chování pro široké komerční využití (např. spotřebitelské preference), analýzu spánkových poruch, uplatňuje se při vývoji BCI (brain computer interface) a v mnoha dalších oblastech. Z toho vyplývá, že analýza EEG záznamu je dnes používána rozdílně v mnoha oblastech a k různým účelům, jejichž cíle jsou naprosto odlišné. [1; 2; 3; 4; 5; 6; 7]

Pořízený EEG záznam je vždy kontaminován artefakty biologického i technického charakteru, kterým se nelze stoprocentně bránit již při měření. Nalezení a následná klasifikace všech artefaktů a jiných abnormalit v EEG záznamu je velice důležitá pro všechny možné oblasti využívání EEG. Identifikace abnormalit znamená, že potřebujeme oddělit normální činnost mozku od abnormální, kterou ale v záznamu samotném nelze generalizovaně identifikovat kvůli jejím individuálním projevům napříč hodnocenými subjekty. Dalším důležitým problémem k řešení, který s identifikací artefaktů souvisí, je identifikace na časové rovině záznamu, kdy taková událost začíná a končí. [1; 2; 3; 8; 9]

Vizuální vyhodnocení záznamu člověkem, které je dosud nejvíce využíváno ve většině aplikací v praxi, jako nejspolehlivější metoda a je časově velmi náročné. I touto metodikou dochází k častým neshodám napříč hodnotícími v závislosti na jejich subjektivním hodnocení. Z tohoto důvodu přetrvává snaha o vytvoření systému pro automatizované vyhodnocování EEG záznamu. [3; 9]

V medicínské praxi je jedinou entitou, která je schopna stoprocentně identifikovat a klasifikovat všechny události v EEG záznamu zkušený specialista-člověk. Všechny ostatní strojově-automatizované analýzy budou z důvodu obrovského vlivu na medicínskou diagnostiku vždy sloužit pouze jako nástroj diagnostika. Tento fakt je důležitý zejména pro vývoj softwarových metod identifikující artefakty v EEG záznamu založených na umělých inteligencích (UI), autonomních neuronových sítích apod. Současný trend vývoje nových metod často směřuje k použití těchto UI. Je třeba si uvědomit, že tyto systémy jsou založeny na obrovském nekontrolovaném každodenním sběru dat všeho druhu a interpretují je ze zdrojů a způsobem, který neznáme, tedy je pro nás netransparentní. Při návrzích nových metodik automatizované klasifikace EEG signálu využívajících UI je potřeba navrhovat takové příznaky dotazovaných parametrů, pro které máme validované referenční příznaky z jiného zdroje.

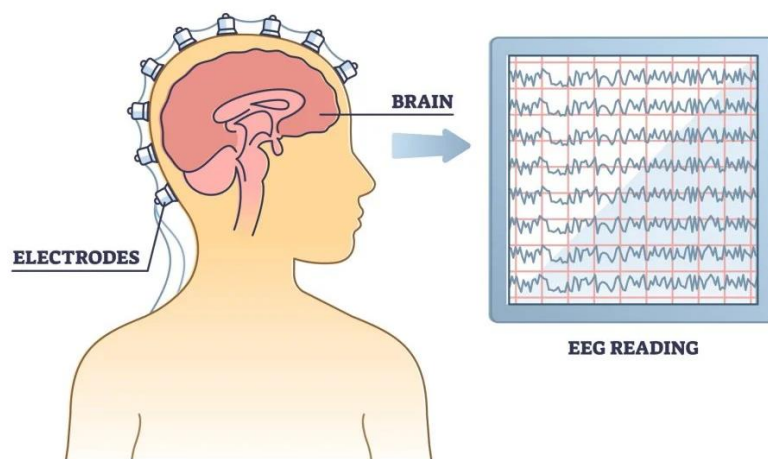
Většina současných, oproti UI v mnoha ohledech transparentních, automatizovaných systémů pro vyhodnocení EEG záznamu pracuje ve třech krocích: segmentace, extrakce příznaků, klasifikace. Segmentace má za cíl rozdělit EEG signál na časové ose tak, aby byly jednotlivé události v záznamu odděleny pro následující analýzy, k čemuž se využívá různých typů adaptivních segmentací. Extrakce příznaků spočívá v sofistikovaných matematických analýzách výsledných segmentů. Klasifikace má za úkol na základě příznaků rozdělit události do stejnorodých skupin a případně je v záznamu viditelně označit pro další potřeby. [1; 2; 3; 9; 10]

Cílem této diplomové práce je navrhnout metodiku pro automatizovanou klasifikaci reálného EEG záznamu. Součástí metodiky bude návrh adaptivní segmentace, která bude využívat analýzu v časové a frekvenční oblasti. Metodika bude dále realizována, otestována a statisticky vyhodnocena v programovém prostředí MATLAB na simulovaných a reálných datech. Na základě výsledků budou navrženy příznaky a procesy pro klasifikaci pomocí umělé inteligence.

2 Přehled současného stavu

2.1 Úvod do elektroencefalografie

Elektroencefalografie (EEG) je metoda měření elektrické aktivity mozku, která slouží pro diagnostiku a výzkum činnosti mozku. Informace o elektrické aktivitě mozku získané pomocí měření EEG na skalpu hlavy je kontaminována řadou artefaktů biologického i technického původu. Primární snahou při softwarových (SW) analýzách je oddělení těchto nežádoucích složek od mozkové aktivity. Všeobecně se předpokládá, že neexistuje metoda automatizované klasifikace EEG záznamu, která by byla natolik spolehlivá, aby se na ni mohli spolehnout lékaři v klinické praxi. V současné době musí diagnostik záznam kontrolovat a klasifikovat manuálně, což vyžaduje obrovské množství času, protože např. spánkové záznamy mohou trvat i více než 24 hodin. Z tohoto důvodu je stále vyvíjeno obrovské úsilí na vytvoření SW pro automatizovanou klasifikaci EEG záznamů, nebo systému, který by provedl předvýběr významných úseků v signálu tak, aby diagnostik nemusel procházet ručně celý záznam a mohl se zaměřit pouze na části signálu relevantní pro stanovení diagnózy. Vzhledem k tomu, že EEG je stochastický signál, mělo by být jeho SW zpracování založeno zejména na statistických charakteristikách. [1; 2; 3; 9]

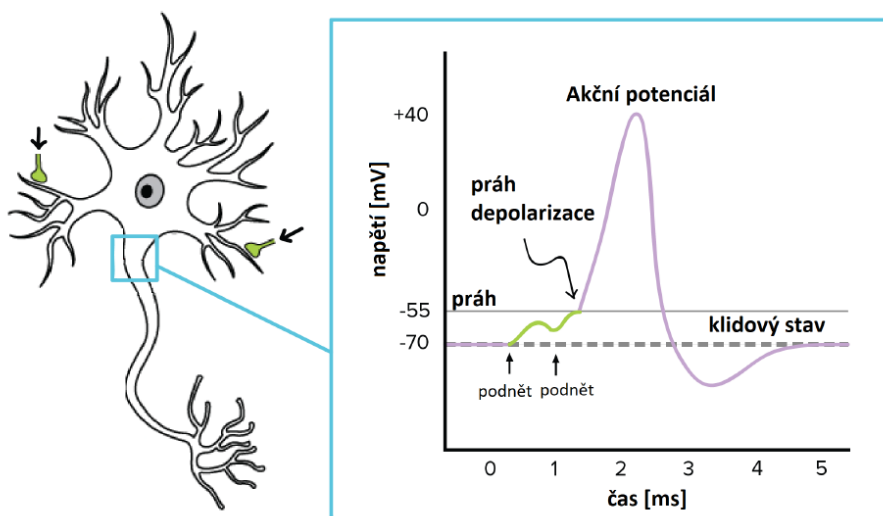


Obrázek 2.1. Ilustrace měření EEG záznamů. Převzato z [50]

2.1.1 Neurofyziologický původ

Neurofyziologie je medicínským oborem, který se zabývá vznikem a přenosem informace v nervovém systému, který se skládá z mozku, míchy a periferního nervového systému. Základní buňkou nervové soustavy je neuron, který se skládá z buněčného těla s jediným jádrem, odstředivého přenášejícího elektrický impuls a proměnlivého počtu dostředivých dendritů, které obsahují velké množství receptorů, které jsou klíčové pro chemický přenos informace pomocí tzv. neurotransmiterů. Buněčná membrána neuronů má omezenou propustnost pro většinu chemických látek, což v důsledku zapříčiňuje přítomnost klidového membránového potenciálu. Klidový membránový potenciál vzniká kvůli rozdílné koncentraci Na^+ a K^+ iontů uvnitř a vně buněčné membrány (rozdíl jejich koncentrací udržují sodno-draselné pumpy) a dosahuje hodnot -70 až -90 mV. Jedinečnost neuronů spočívá v jejich rychlosti přenosu informace ve formě podráždění. V lidském mozku je každý neuron spojen s přibližně dalšími 10 000 jinými neurony, většinou prostřednictvím dendritických spojení. [1; 2; 9]

Pokud je neuron s klidovým membránovým potenciálem stimulován, dochází k ovlivnění aktivity napěťově řízených iontových kanálů a v důsledku dojde k porušení rovnováhy klidového membránového potenciálu. Je-li dosaženo dostatečné změny polarity membrány, změnou napětí dojde k otevření kanálů pro Na^+ a Na^+ ionty začnou prostupovat ve směru koncentračního gradientu do buňky, která takto získává kladný náboj a dochází k její depolarizaci. Se vzrůstající depolarizací se Na^+ kanály postupně zavírají, a tím přestanou být sodíkové ionty vtahovány do buňky. Současně s otevřením Na^+ kanálů jsou otevírány i kanály (rovněž řízené napětím) pro K^+ ionty proudící z buňky ven ve směru koncentračního spádů. K^+ kanály zůstávají otevřeny déle než Na^+ kanály a tím nakonec napomáhají k repolarizaci buňky a uvedení buňky zpět do původního stavu klidového membránového potenciálu. [1; 2; 9; 11]



Obrázek 2.2. Znázornění akčního potenciálu a buňky neuronu. Obrázek převzat z [11]

Základem snímané aktivity pomocí EEG je generování akčních potenciálů nervovými buňkami, které vzniká prouděním iontů skrze membránu, a to následně vede k vzniku excitačních/inhibičních postsynaptických potenciálů, které jsou v naměřeném EEG signálu majoritním zdrojem aktivity. Samotné akční potenciály, které byť mají vyšší amplitudu ale kratší čas trvání než excitační/inhibiční postsynaptické potenciály, přispívají do EEG jen minimálně.

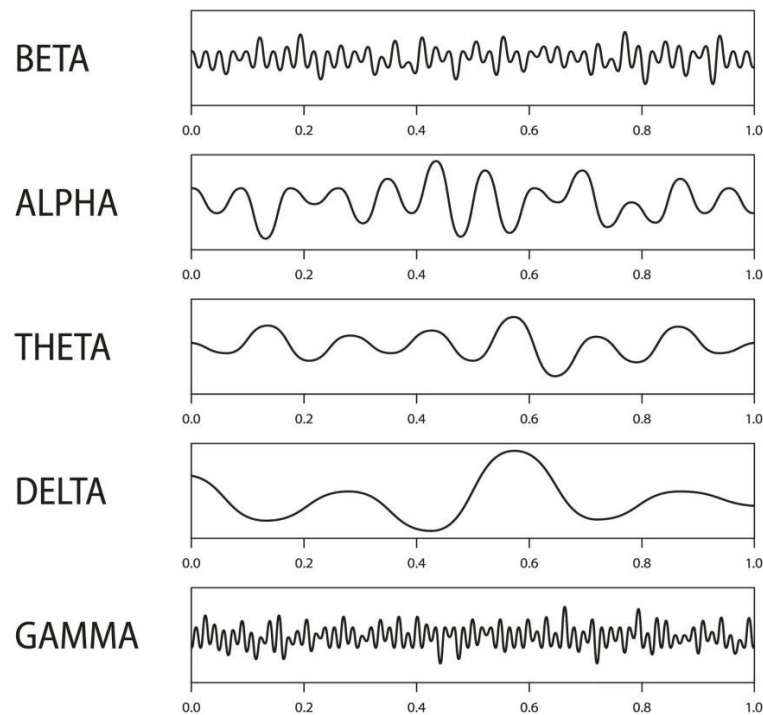
Jeden samotný neuron je schopen vytvořit pouze velice malý elektrický potenciál, který není možné měřit elektrodami na skalpu hlavy, protože tato aktivita je překryta aktivitou sousedních neuronů, nebo je natolik malá, že je utlumena při prostupu zejména lebkou. Z toho důvodu není dosaženo takové sumace potenciálů, která by vedla k propagaci tohoto vzruchu až k elektrodám umístěným na skalpu hlavy. V případě, že dojde k současné aktivitě větší skupiny neuronů, vzniká elektrické potenciálové pole, které je dostatečně velké na to, aby i přes utlumení lebkou a dalšími faktory bylo měřitelné až na skalpu ve formě signálu EEG. [1; 2; 9]

Mozková kůra je uspořádána do šesti vrstev, ve kterých jsou jednotlivé nervové buňky uspořádány do sloupce a jsou uloženy paralelně s hlavní osou axonu a dendritického stromu, přičemž celé toto uspořádání je kolmé k povrchu mozkové kůry. Takovéto uspořádání je primární podmínkou pro snímání elektrické aktivity nervových buněk chovajících se jako dipóly. [2; 9; 11]

Výsledný signál EEG snímaný na elektrodě umístěné na skalpu je převážně složen ze sumace excitačních a inhibičních postsynaptických potenciálů vznikajících ve sloupcové oblasti mozkové kůry pod danou elektrodou. [1; 2; 9]

2.1.2 EEG rytmy

Jak bylo zmíněno dříve, EEG signál vzniká sumací elektrické aktivity velkého počtu neuronů, která je snímána na skalpu a vyhodnocována zkušeným specialistou pro diagnostiku různých poruch, např. epilepsie, zranění hlavy, poruchy spánku, psychické poruchy, neurologické poruchy atd. Hodnoty amplitudy a frekvence v signálu jsou pro každého pacienta unikátní a mění v závislosti na aktuálním stavu, jako jsou: bdění, spánek, únava, věk apod. Signál EEG má velmi široké frekvenční spektrum a rozlišujeme v něm několik základních frekvenčních pásem, tzv. rytmů, které jsou významné pro klinickou praxi, a to konkrétně pásma: delta, theta, alfa, beta a gama (viz obrázek 2.3). Jednotlivá frekvenční pásma jsou spojována se speciálním typem aktivity, přičemž jejich rozdíl nespočívá pouze ve frekvencích, ale také ve velikosti jejich amplitudy. [1; 3; 12]



Obrázek 2.3. Ilustrace běžných rytmů v EEG záznamu. Převzato z [50]

Delta rytmus

Delta rytmus se pohybuje ve frekvencích 0,5–4 Hz a jeho amplituda dosahuje hodnot 20–200 μV . Primárně je spojován s hlubokým spánkem, nebo u kojenců v bdělém stavu. Pokud se delta rytmus vyskytuje i u bdělého dospělého pacienta, jedná se pravděpodobně o patologii, např. nádor. Delta rytmus je významným pásmem pro určování fází spánku. [1; 2; 3; 13]

Theta rytmus

Theta rytmus se nachází ve frekvenčním pásmu 4–8 Hz a jeho amplituda dosahuje hodnot 0–30 μV . Vyskytuje se především ve frontální a temporální oblasti. S tímto rytmem se můžeme setkat v EEG záznamu dětí, nebo u dospělých pacientů při ospalosti, usínání, hluboké meditaci a v některých fázích spánku. Nadměrný výskyt theta rytmu u bdělého dospělého pacienta je způsoben různými patologickými problémy. [1; 2; 12; 13]

Alfa rytmus

Alfa rytmus se pohybuje ve frekvenčním pásmu 8–14 Hz s amplitudou a jeho amplituda dosahuje hodnot 30–80 μV . Vyskytuje se v okcipitální a parietální oblasti. Alfa rytmus lze snadno detekovat u zdravých pacientů ve stavu bdělosti, kdy má subjekt zavřené oči. Předpokládá se, že alfa rytmus provází uvolněné vědomí. Při otevření očí, usínání/ospalosti a soustředěné mentální činnosti se tento rytmus z EEG signálu vytrácí. [1; 2; 12; 13]

Beta rytmus

Beta rytmus se typicky pohybuje ve frekvenčním pásmu 14–30 Hz a jeho amplituda dosahuje hodnot 10–30 μV . Vyskytuje se ve frontální a centrální oblasti. Beta rytmus je spojován s aktivním myšlením/pozorností zaměřené na vnější svět, nebo řešení konkrétních problémů. K blokování beta rytmu dochází při pohybu, taktilní stimulaci, nebo vlivem omamných látek. [1; 2; 12; 13]

Gama rytmus

Gama rytmus se vyskytuje ve frekvenčním rozsahu vyšším než 30 Hz s amplitudou 2–10 μV a vyskytuje se převážně nad frontální a centrální oblastí. Gama rytmus souvisí s bdělým stavem pacienta, který v danou chvíli aktivně zpracovává informace. I přes to, že amplituda tohoto rytmu bývá velmi nízká, může být využit pro potvrzení některých mozkových onemocnění. [1; 2; 12; 13]

2.1.3 Artefakty v EEG záznamu

Artefakty v EEG záznamu se rozlišují podle prostředí jejich vzniku na dvě základní skupiny, a to: biologické (endogenní) a technické (exogenní). Biologickým artefaktům, které jsou fyziologickými projevy činnosti těla pacienta, nelze nijak účinně zabránit. V případě velice krátkodobého měření je možno pokusit se o naprostý pohybový klid a pokusit se o tzv. „bezartefaktové chování“. V této práci byly biologické artefakty rozděleny do třech skupin podle jejich zdroje: svalové, okulární a kardiovaskulární. Naproti tomu, technické artefakty je možné z velké části omezit zabezpečením vhodného prostředí při měření (např. Faradayova klec, správné umístění a upevnění elektrod, vhodná izolace, dostatečné množství gelu/roztoku pro správnou vodivost). [1; 9; 13]

Seznam artefaktů popsaných v následujících kapitolách:

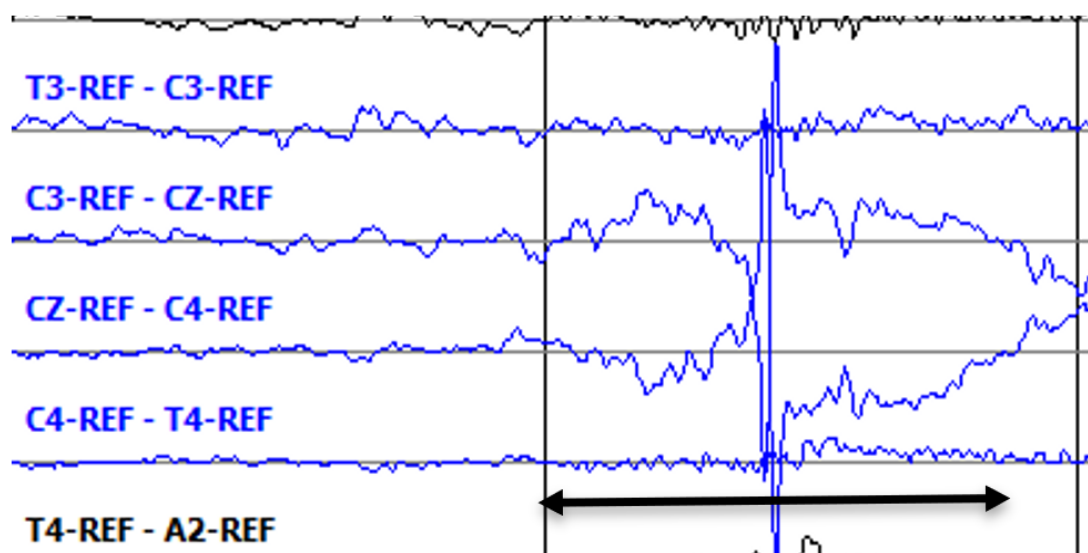
- Technické artefakty
 - Artefakt prasknutí elektrody
 - Síťový šum
- Biologické artefakty
 - Svalové
 - Svalový artefakt
 - Artefakt z chvění
 - Žvýkací artefakt
 - Okulární
 - Artefakt mrkání
 - Artefakt pohybu očí
 - Kardiovaskulární
 - EKG artefakt
 - Pulsní artefakt

Technické artefakty

Technické artefakty, zahrnují velmi široké spektrum možností projevu v EEG záznamu. Jedním z jejich největších problémů je, že není možné generalizovat jejich způsob vzniku a jejich morfologii, protože se na jejich vzniku podílí nesčetné množství faktorů ze samotného přístroje a prostředí, ve kterém byl záznam pořizován. [9; 14; 15; 16]

Artefakt prasknutí elektrody

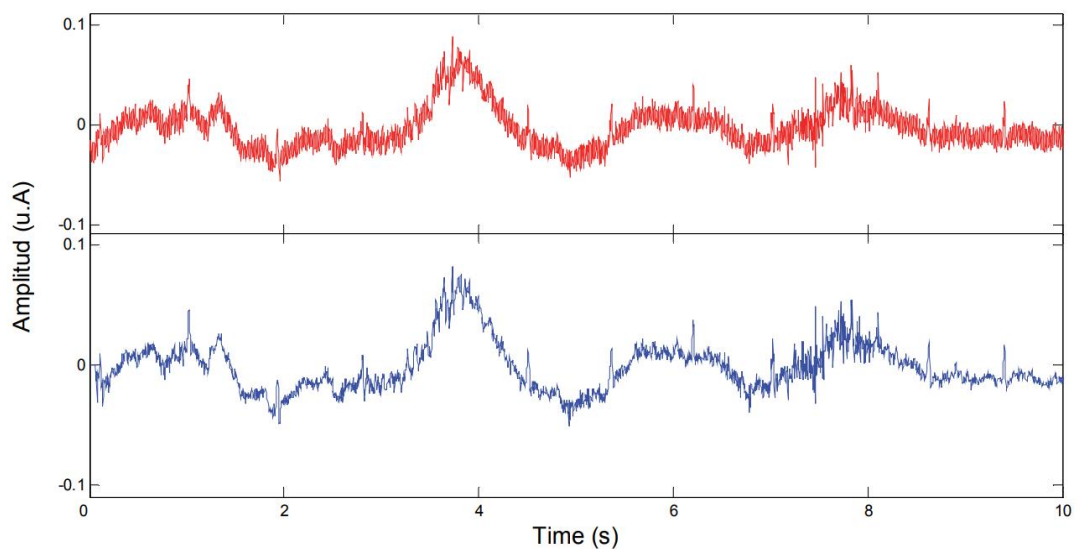
Artefakt prasknutí elektrody vzniká závadou na elektrodě a může být způsoben např. nesprávnou adhezí, nebo bublinkami v použitém gelu. Projevuje se velmi rychlou změnou amplitudy na pouze jedné elektrodě (viz obrázek 2.4), která je ostrá a je možné ji snadno odlišit od záznamu pocházejícího z mozku. [9; 14; 17]



Obrázek 2.4. Prasknutí elektrody. Převzato z [17]

Síťový šum

Síťový šum je jedním z nejčastěji se vyskytujících artefaktů v EEG záznamech. Jedná se o artefakt síťového kmitočtu, který je 50 Hz v České republice, nebo 60 Hz v USA, přičemž v EEG záznamu se objevují i jejich vyšší harmonické složky. Může být způsoben interferencí z okolních přístrojů nebo zdrojů. V případě, že dojde např. k odpojení zemnicí elektrody, objeví se síťový šum ve všech svodech. Pokud se síťový šum vyskytuje pouze u jedné elektrody, je pravděpodobně způsobem velmi vysokým odporem elektrody. Praktickým řešením pro omezení vlivu síťového šumu na EEG záznam je měření ve Faradayově kleci a po naměření záznamu je rovněž vhodné použití filtru typu pásmová zádrž. [9; 18]

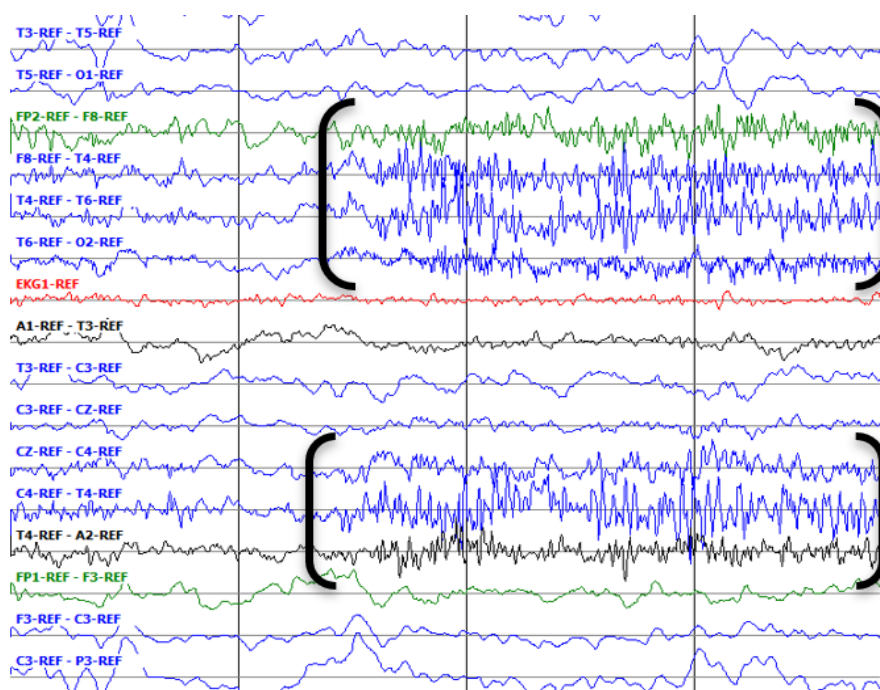


Obrázek 2.5. EEG záznam zarušený síťovým šumem o frekvenci 50 Hz (nahore, červeně). EEG záznam po odstranění síťového šumu filtrací (dole, modře). Obrázek převzat z [21]

Biologické artefakty – svalové

Svalový artefakt

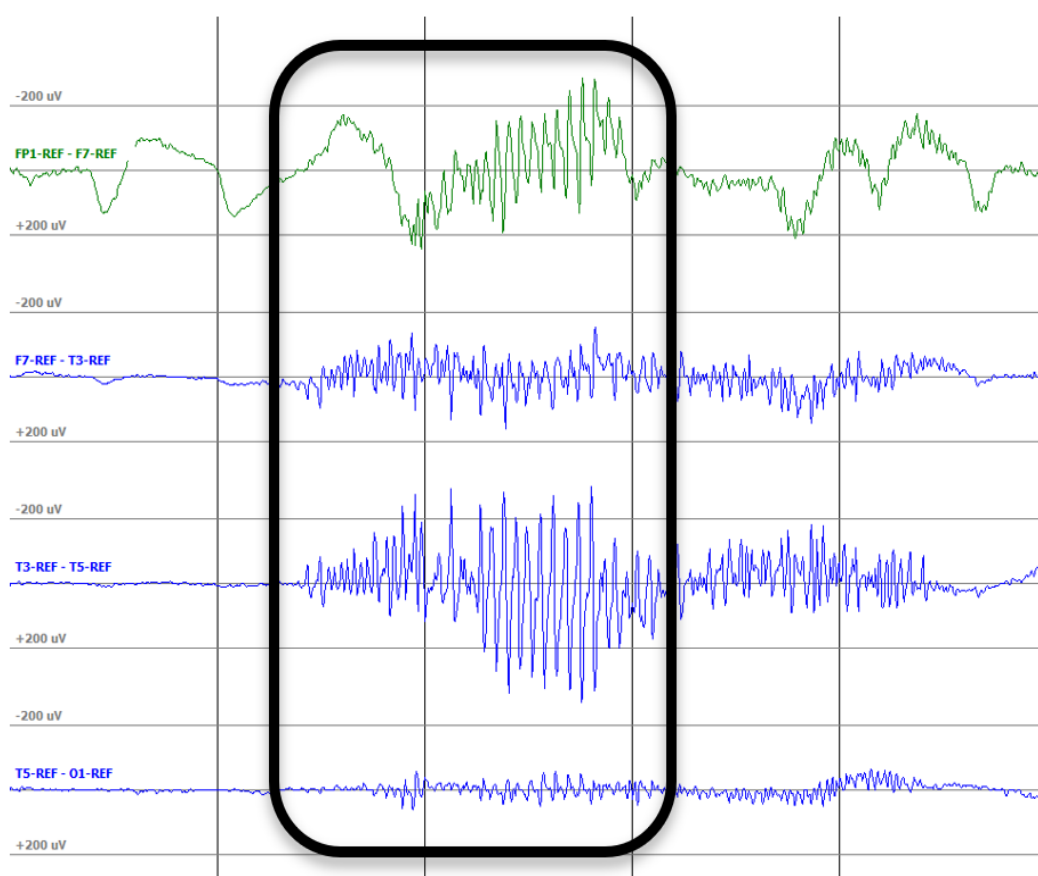
Svalový artefakt (viz obrázek 2.6) je spojován s jakýmkoli typem pohybu pacienta při měření, např. polykání, žvýkání, mluvení, sání, škytání, kašláním apod. Svalový artefakt je charakterizován zejména vysokofrekvenční aktivitou (obvykle vyšší než 30 Hz), přičemž jeho amplituda je velice proměnlivá v rozmezí velmi nízké až velmi vysoké [17; 19]. Kvůli jejich vysokofrekvenčnímu projevu jsou tyto artefakty relativně snadno rozpoznatelné, ale na druhou stranu je velice složitým problémem je ze záznamu odstranit. Odstranění svalových artefaktů s využitím filtrace vysokých frekvencí se nedoporučuje, protože kvůli nedostatečné efektivitě filtrace může být výsledek takovéto úpravy záznamu snadno zaměnitelný s výskytem hrotů mozkového původu, což může v koncové klasifikaci vést např. k záměně s epileptiformním výskytem [9; 20]. Vzhledem k tomu, že svalové artefakty vznikají kvůli elektrické aktivitě generované svalovou kontrakcí, je možné jim částečně předcházet již při samotném měření tím, že měřený pacient se snaží být co nejvíce relaxován bez mluvení, pohybů apod. [1; 9; 14; 17; 19]



Obrázek 2.6. Část svalového artefaktu indikovaná morfologií vysoké frekvence a střední až vysoké amplitudy. Převzato z [17]

Artefakt z chvění pacienta

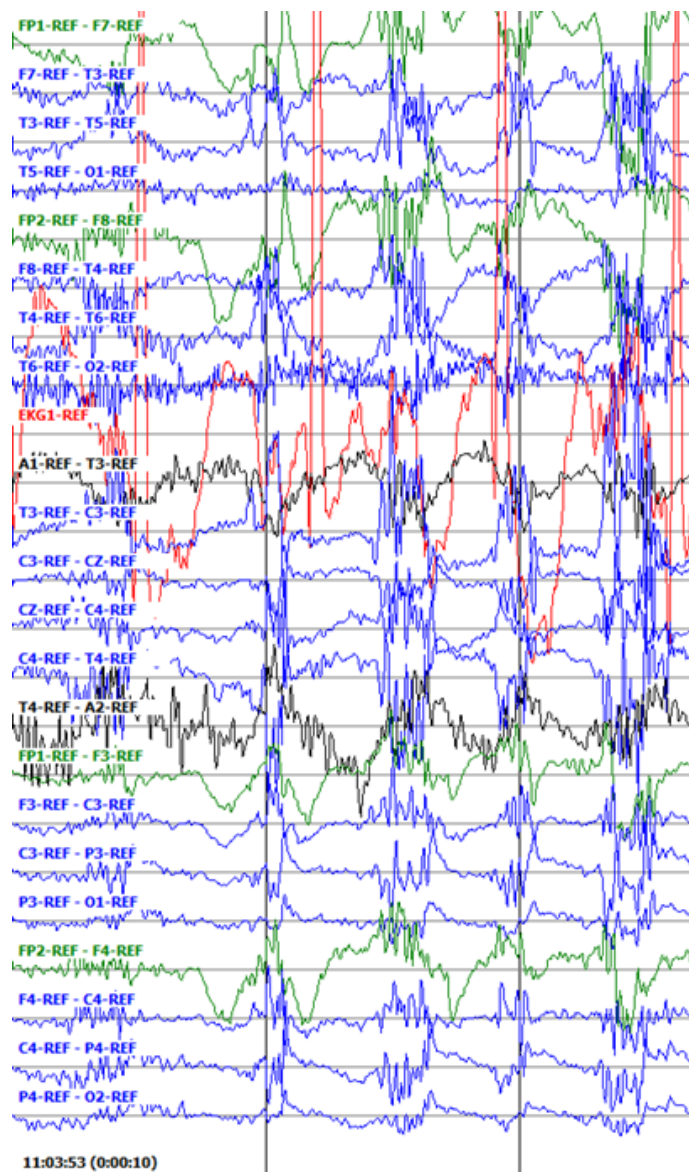
Tento typ artefaktu je podmnožinou rozsáhlejší skupiny svalových artefaktů a dochází k němu při třesu pacienta. Tento artefakt je poměrně vzácný, ale v případě, že k němu dojde, stává se značným zdrojem chyb, protože v mnoha případech může být podobný špičce a pomalé vlně, které indikují záchvatovou událost. Projevuje se charakteristickým komplexem složeným ze střední ostré amplitudy, která se přenáší do špičky vyšší amplitudy, než dojde k ponoření křivky pod základní linii, poté následuje vznik dalšího komplexu. Chvějící se artefakt se pohybuje zejména ve frekvencích beta pásma a vyskytuje se na většině měřených kanálů. Ukázku chvějícího se artefaktu můžete vidět na obrázku 2.7 níže. [1; 14; 17]



Obrázek 2.7. Ukázka artefaktu chvění. Převzato z [17]

Žvýkáci artefakt

Žvýkáci artefakt je rovněž podmnožinou svalových artefaktů a je způsoben zatínáním a uvolňováním čelistních svalů. Projevuje se vysokofrekvenční aktivitou podobně jako nadřazená skupina svalových artefaktů, ale mezi výskyty těchto artefaktů se vyskytuje přibližně půl sekundy základní linie. Žvýkáci artefakt se vyskytuje zejména v temporálních kanálech s velkým rozpětím a někdy může v jedné hemisféře vykazovat vyšší aktivitu. Žvýkáci artefakt je poměrně běžným artefaktem v EEG záznamech spících pacientů, kteří během spánku zatínají a uvolňují čelistní svalstvo a v některých případech může být žvýkáci artefakt spojen se záchvatovou událostí. [14; 17]



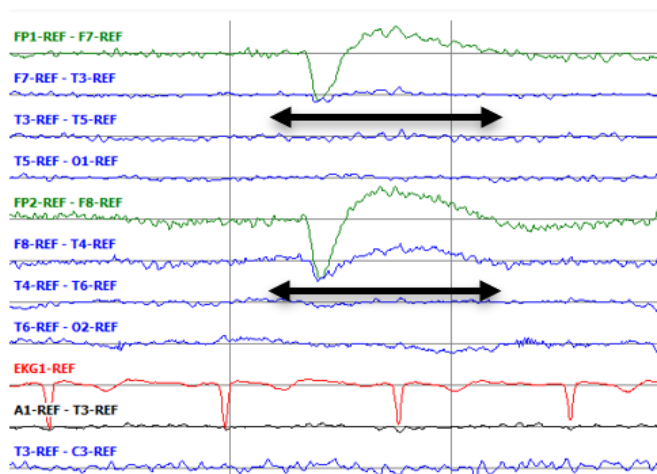
Obrázek 2.8. Žvýkáci artefakt. Převzato z [17]

Biologické artefakty – okulární

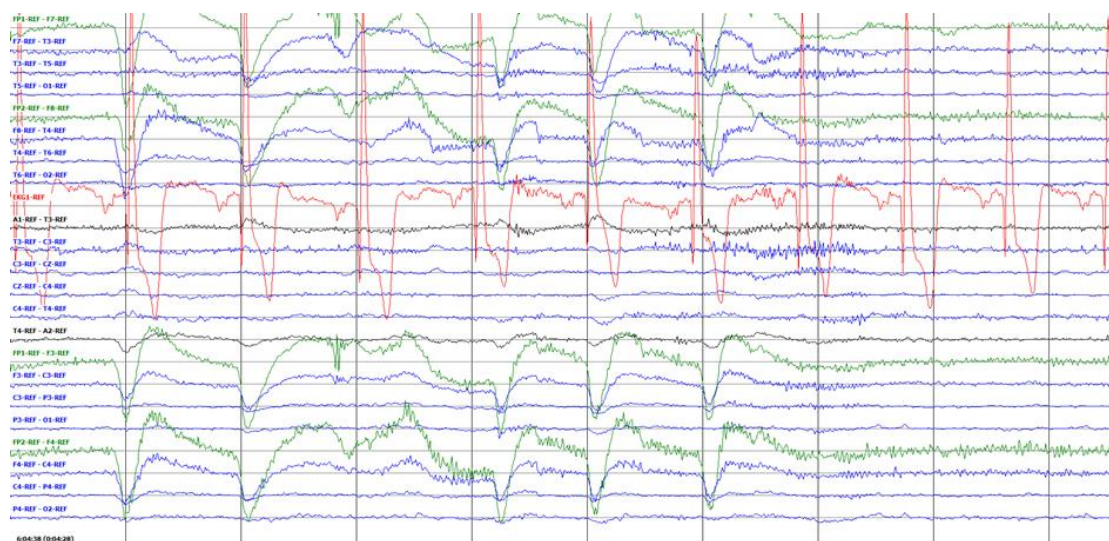
Další skupinou nejčastěji se vyskytujících artefaktů v EEG záznamech jsou artefakty okulární. Vyskytují se zejména ve frontálních svodech a nelze jim účinně zamezit. Tuto skupinu artefaktů je možné dále rozlišovat na artefakty způsobené mrkáním a pohyby očních bulbů.

Artefakty mrkání

Artefakty mrkání (viz obrázek 2.9 a 2.10) jsou vyvolány pohybem elektrod na čele pacienta, kdy se tento artefakt v záznamu projevuje jako jedna ostrá vlna s vysokou amplitudou, kterou zpravidla následuje vlna opačné polarity. V případě, že se v záznamu objevuje mrkání velmi často, může být tento jev spojen s neurologickou dysfunkcí. [1; 14; 17]



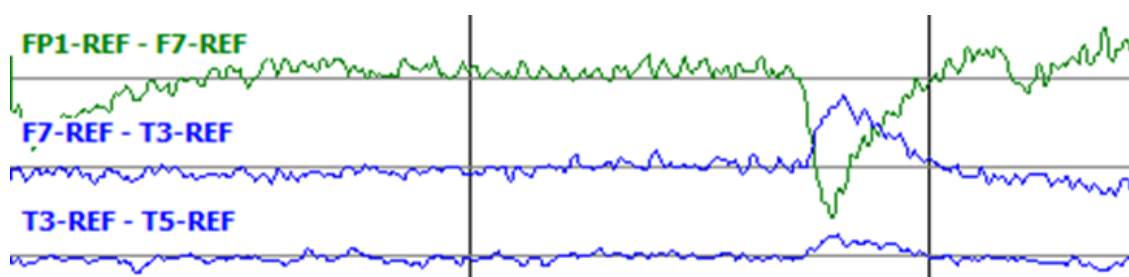
Obrázek 2.9. Jedno mrknutí očí. Převzato z [17]



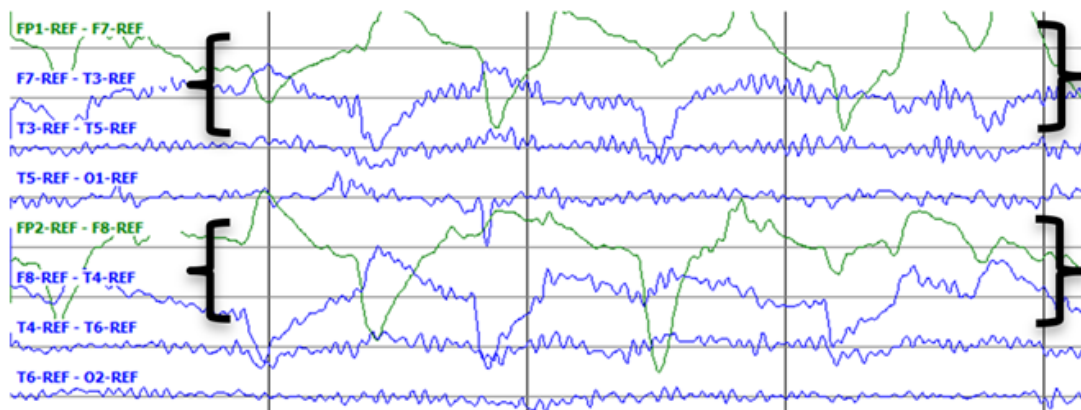
Obrázek 2.10. Série mrknutí očí, která může indikovat neurologickou poruchu. Převzato z [17]

Artefakty z pohybu očí

Tento typ artefaktu se vyskytuje převážně ve frontálních svodech EEG a vyskytují se téměř v každém EEG záznamu. Oko se při měření chová jako dipól, kdy je rohovka vůči retině kladně nabitá, tedy při jeho pohybu vzniká elektrické napětí na elektrodách měřicí soustavy EEG, přičemž nejvyšší změna potenciálu (až mV) je propagována do frontálních elektrod, které jsou umístěny nad očima. Vertikální pohyby se díky svému velkému výkonu projevují i v ostatních svodech, avšak s nižší amplitudou a často i opačnou polaritou [9]. Artefakty pohybu očí mají velkou variabilitu, což zvyšuje chybovost automatizované detekce záchvatů, neboť tyto opakující se vlny mohou napodobovat morfologii záchvatů. [1; 14; 17]



Obrázek 2.11. Jeden laterální pohyb očí. Převzato z [17]



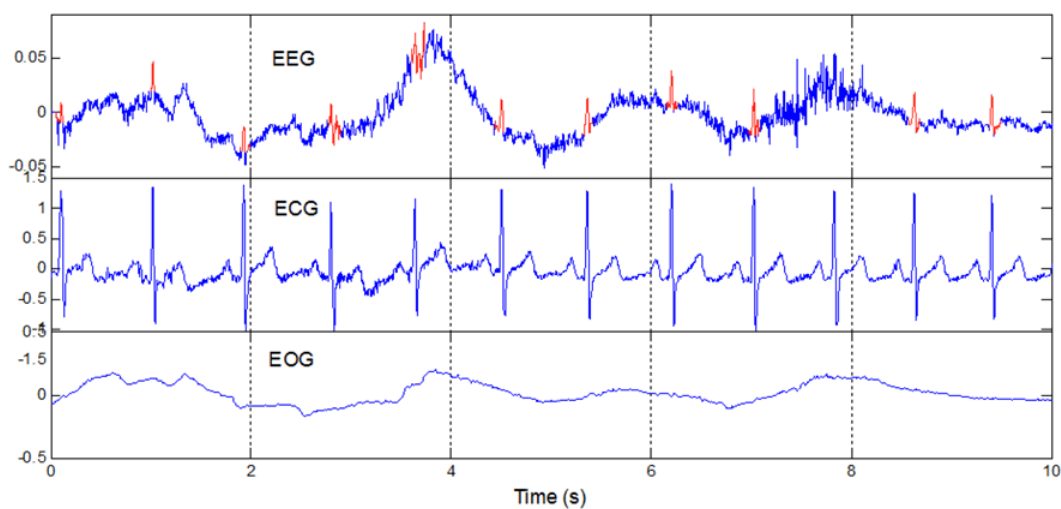
Obrázek 2.12. Série pohybu očí. Převzato z [17]

Biologické artefakty – kardiovaskulární

Mezi kardiovaskulární artefakty patří dva typy artefaktu (EKG a pulsní). EKG artefakt má původ v převodním systému srdečním, tedy je synchronní s EKG křivkou. Pulsní artefakt pochází z mechanické činnosti srdce a odpovídá tepennému pulsu.

EKG artefakt

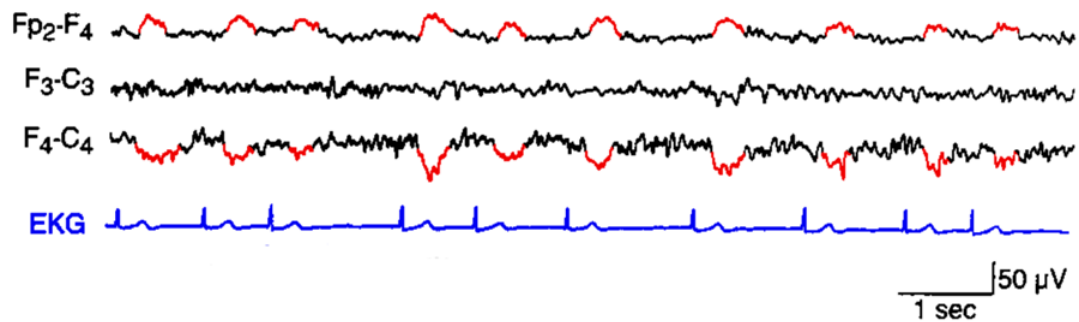
I přes poměrně značnou vzdálenost mezi srdcem a elektrodami EEG umístěnými na skalpu, se u specifických tělesných konstitucí (např. nadváha, krátký krk) může elektrická aktivita srdce propagovat do EEG jako EKG artefakt (viz obrázek 2.13). Bohužel se tomuto artefaktu nedá nijak účinně zabránit. Hlavním projevem v EEG záznamu je výskyt R vlny QRS komplexu. Vzhledem k jeho pravidelnosti a simultánnímu nahrávání EKG a EEG je tento artefakt poměrně snadno identifikovatelný. Problém při vyhodnocování však může nastat při srdečních patologiích, kdy srdce bije nepravidelně, což působí potíže při jeho rozpoznání. [9; 14; 21]



Obrázek 2.13. EKG artefakt v EEG záznamu (červeně) spolu s reálným záznamem EKG (ECG) a EOG. Obrázek převzat z [21]

Pulsní artefakt

Pulsní artefakt pochází z činnosti tepny, která se nachází pod danou elektrodou EEG. Pulsní artefakt odpovídá tepové frekvenci, v EEG se projevuje jako pomalé vlny a může být zaměněn s aktivitou delta, což by mohlo vést k falešné indikaci patologie. Je způsoben periodickými změnami kontaktu mezi elektrodou a skalpem a vyskytuje se většinou jen v jedné svodu. Pulsní artefakt lze účinně odstranit přesunutím elektrody. [1; 9; 14]



Obrázek 2.14. Pulsní artefakt (vyznačen červeně) v EEG záznamu pod elektrodou F4 u pacienta se srdeční arytmií. Obrázek převzat z [9]

2.2 Způsob měření EEG v praxi

2.2.1 Přístroje

Pro měření EEG existuje obrovské množství různých HW sestav, které se od sebe vzájemně liší už na úrovni základních použitých prvků. Mohou se lišit např. v počtu snímaných kanálů, režimu zapojení, zesilovačích, elektrodách, A/D převodnicích, vzorkovacích frekvencích apod. [2; 18]

2.2.2 Elektrody pro měření EEG

Elektrody jsou velmi důležitou HW částí elektroencefalografu, které nám umožňují snímat elektrickou aktivitu mozku jejím převodem na zesilovač. Pro správnou funkčnost elektrod, tedy měření rychlých změn v signálu, nesmí být elektrody polarizovatelné. Charakteristika různých druhů elektrod se odvíjí od použitého materiálu (např. stříbro) a vodivého média. Podle způsobu aplikace je rozdělujeme na povrchové, podkožní a speciální. [15; 18]

Povrchové elektrody

Povrchovými elektrodami měříme elektrickou aktivitu mozku na skalpu hlavy, proto se můžeme v literatuře setkat i s názvem skalpové elektrody. Jejich průměr se pohybuje v rozmezí cca 4-10 mm a nejčastěji používaným materiálem je stříbro. Ve většině případů jsou ploché s otvorem pro nanesení vodivého média, přičemž udržení polohy na skalpu je realizováno pomocí pásků, past apod. Nejrozšířenějším typem elektrody, který je používán v praxi jsou elektrody druhého druhu Ag/AgCl. [15; 18]

- Diskové elektrody

Diskové elektrody jsou ploché s kruhovým tvarem o průměru do 10 mm. Na skalp jsou připevněny pomocí speciální pasty. Nutnou přípravou pro použití tohoto typu elektrody je očištění a připravení kůže před upevněním elektrody. U tohoto typu elektrody nastává často problém ve formě zlomení kabelu v místě, ve kterém je spojen s elektrodou. [15; 18]

- Elektrody s fyziologickým roztokem

Upevnění tohoto typu elektrody je realizováno zašroubováním do držáku z umělé hmoty a jsou přidržovány maskou z gumových pásků pošitých tkaninou z bavlny. Před aplikací elektrod je nutné je vložit do fyziologického roztoku a během měření pohlídat, aby roztok nevyschnul a nedocházelo tak ke ztrátě kontaktu. Na druhou stranu, roztoku nesmí být aplikováno ani příliš moc, aby nestékal po skalpu a nedošlo tak ke zkratu mezi elektrodami. Výhodou tohoto typu je rychlá aplikace a to, že kůže před aplikací nemusí být nijak zvlášť důkladně připravena. [15; 22]

- Elektrody s koloidem

Tyto elektrody jsou opět ploché s kruhovým tvarem, který v sobě má prohlubeň vyplněnou vodivou pastou. Aplikace tohoto typu je zdlouhavá, neboť je nutné důkladně očistit pokožku, poté na ni nanést trochu pasty, přichytit elektrodu s využitím speciálního držáku a v posledním kroku ji zakápnout koloidem. Tento typ elektrody se používá pouze při dlouhodobém snímání kvůli dlouhé přípravě a pro pacienta nepříjemnému odstraňování elektrod z pokožky. [15]

- Ag/AgCl elektrody

Tyto elektrody se upevňují opět s využitím pasty na holých částech pokožky. Jejich upevnění na skalpu je zajištěno pomocí oboustranně lepících kroužků z PVC fólie a vodivost je zprostředkována pomocí elektrolytu, který obsahuje ionty chloru, proto se využívá roztok KCl (většinou zahuštěný do gelové konzistence). [15; 18; 23]

Podkožní elektrody

Podkožní elektrody se zavádí pod kůži, jsou tenké a tvarem připomínají injekční jehlu o velikosti 20 mm. U měřeného subjektu vyvolávají nepříjemný pocit způsobený vpichem, při aplikaci hrozí nebezpečí infekce, a proto se v dnešní době v praxi již nepoužívají. [15; 18]

Speciální elektrody

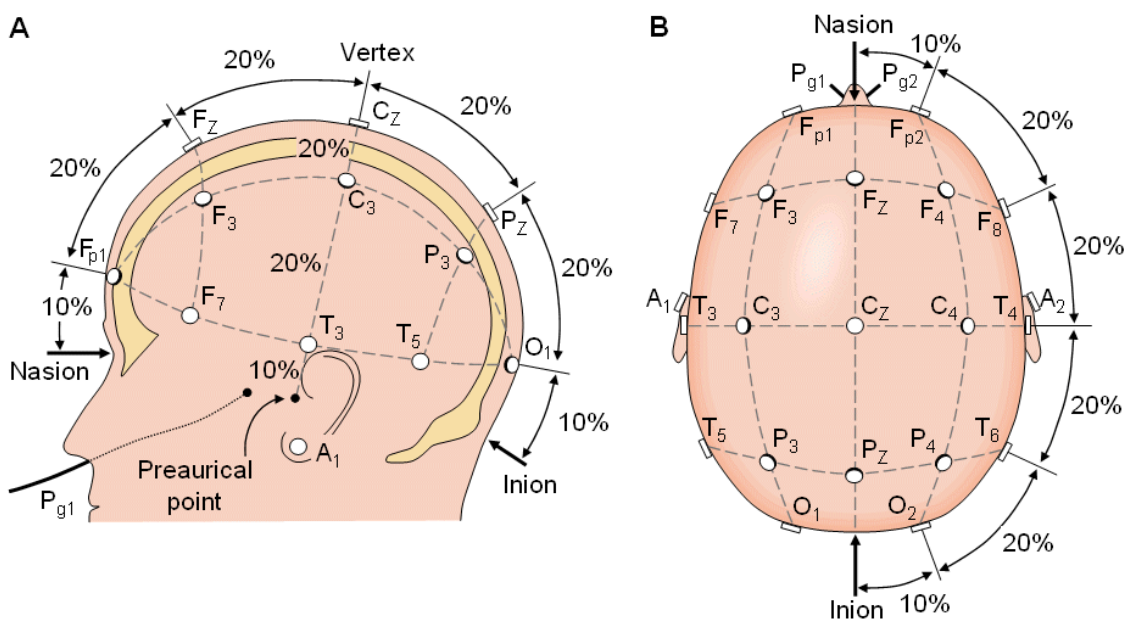
Speciální elektrody jsou využívány k zavedení do konkrétní dutiny lidského těla, popř. pro vpichování. Mezi speciální elektrody patří:

- Sfenoidální – aplikace vpichem mezi mandibulu a jařmovou kost
- Nasofaryngeální – zavádí se do nosní dírky
- Tympanální – aplikace do zvukovodu
- Kortikální – pořizování záznamu z povrchu mozkové kůry
- Intracerebrální – zavádění s využitím jehly do hlubších mozkových struktur

2.2.3 Rozmístění elektrod na skalpu

Jedním z nejrozšířenějších uspořádání elektrod na skalpu hlavy je mezinárodně uznávaný systém 10/20, kdy je na skalpu rozmístěno 21 elektrod, přičemž vzdálenost mezi jednotlivými elektrodami je buď 10 % nebo 20 % z celkové vzdálenosti od kořenu nosu (nasion) ke kostnímu výrůstku na týlní kosti (inion) (viz obrázek 2.15). V praxi se objevují i další rozšíření tohoto standardního zapojení, kdy výsledná konfigurace elektrod může být např. 10/10, nebo 10/5, přičemž v těchto konfiguracích je samozřejmě na skalpu hlavy rozmístěno více elektrod tak, aby byly dodrženy procentuální rozestupy elektrod dle používané konfigurace podle výše zmíněného pravidla. Ve výzkumných zařízeních se můžeme ještě setkat i s tzv. HD (high density) EEG, které používá mnohem vyšší prostorové uspořádání elektrod (128 až 256 elektrod), které je i přes své vyšší náklady na pořízení, provoz, údržbu a vyhodnocení záznamu výhodné zejména při lokalizaci zdrojů aktivity v mozku např. zdroje epileptické aktivity. [2; 15; 22; 24]

Měrné elektrody jsou v každé z výše zmíněných konfigurací označovány písmenem/písmeny a číslem podle jejich umístění na skalpu hlavy, a to: Fp (prefrontální), F (frontální), C (centrální), P (parietální), T (temporální), O (okcipitální), A (aurikulární). Číselné označení elektrod se řídí pravidlem, kdy lichá čísla jsou přiřazena levé hemisféře a sudými čísly jsou označovány elektrody pravé hemisféry (viz obrázek 2.15). [1; 2; 15; 22]



Obrázek 2.15. Rozmístění elektrod ve standardním systému 10/20. Obrázek A reprezentuje boční pohled, obrázek B pohled shora. Obrázek převzat z [22]

K upevnění výše zmíněných systémů elektrod na skalpu hlavy se používá EEG čepice, která je považována za nejpraktičtější a nejefektivnější řešení tohoto problému. Čepice je zpravidla výrobcí dodávána v několika různých velikostech tak, aby bylo možné pořizovat záznam EEG u všech pacientů od novorozeneckého až po dospělý věk. Jedná se o čepici vyrobenou z elastického materiálu, jejíž součástí jsou zabudované elektrody v dané konfiguraci a je připravená k tomu, aby se pouze nasadila na hlavu pacienta a obsluha nemusela řešit umístování jednotlivých elektrod. Jediné, co musí obsluha při měření zajistit, je dostatek vodivého média mezi každou elektrodou a skalpem tak, aby byla snížena měrná elektrická impedance, čehož je možné snadno dosáhnout použitím vhodných vodivých gelů nebo roztoků. Roztoky se používají spíše u HD EEG, kdy by bylo velice časově náročné nanášet ručně gel pod všechny elektrody v této konfiguraci (viz obrázek 2.16). [2; 3; 22]



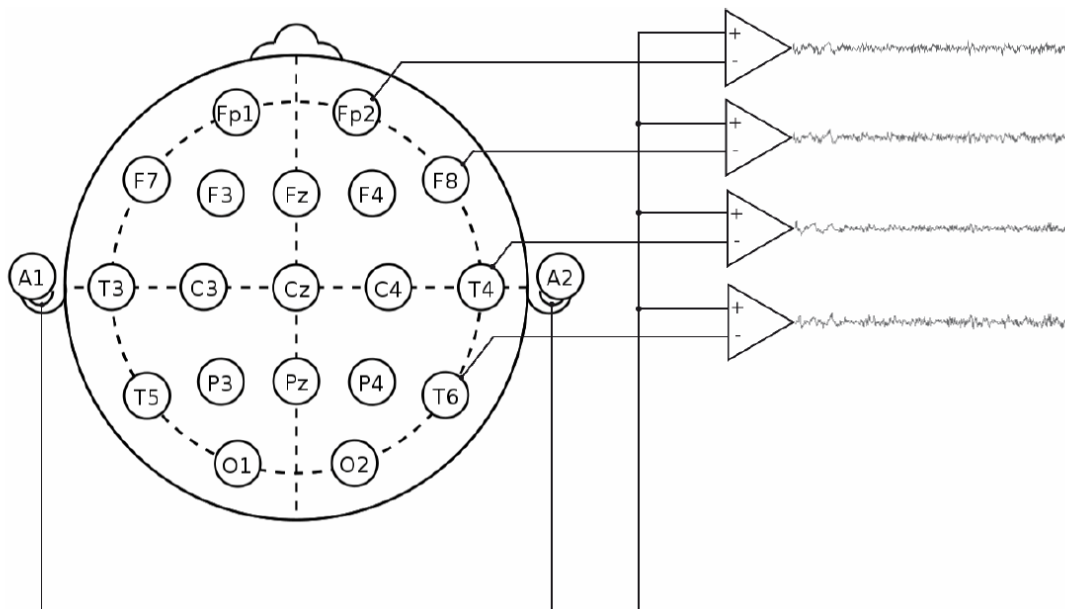
Obrázek 2.16. Ukázka EEG čepice využívané při měření HD EEG. Obrázek převzat z [51]

2.2.4 Režimy zapojení elektrod

Pro snímání EEG záznamu se v praxi nejčastěji používají 3 režimy zapojení elektrod: unipolární, bipolární, zprůměrovaný. Samotná volba zapojení zásadně ovlivňuje snímání EEG signálu a vede k jisté „analogové předfiltraci“ signálu, která je považována za bezeztrátovou (např. zprůměrovaný režim vyrovnává pohybové artefakty). [3]

Unipolární režim

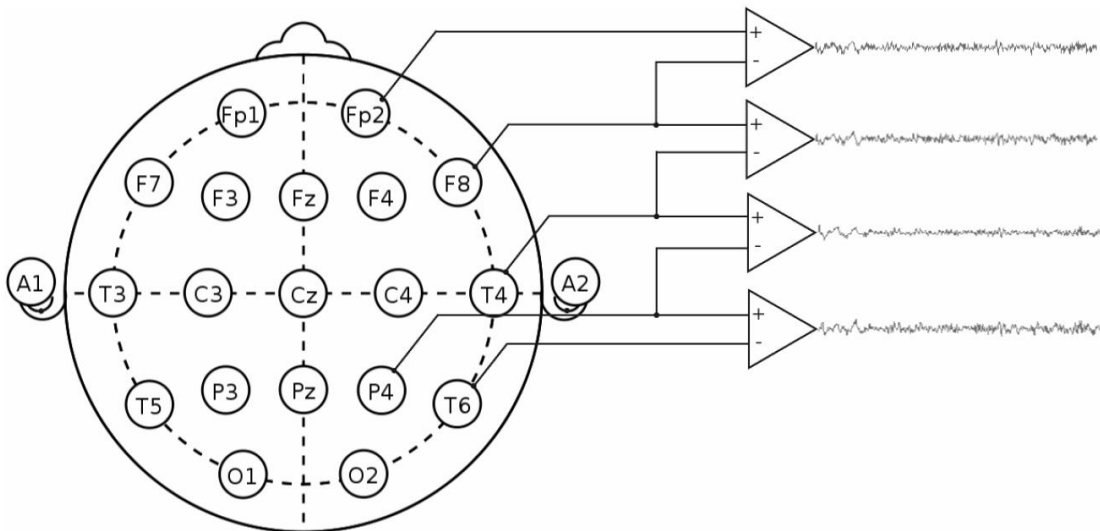
Unipolární montáž (viz obrázek 2.17) označuje rozdíl mezi elektrickým potenciálem naměřeným na měrné elektrodě a referenčním uzlem. Referenční elektroda bývá nejčastěji umístěna na ušním lalůčku, ale můžeme se setkat i se zapojením, kdy se oba ušní lalůčky spojují. V tomto režimu je možné lépe pozorovat velikost a tvar vln a signál má vyšší úroveň. [3; 15; 25]



Obrázek 2.17. Zapojení elektrod v unipolárním režimu. Obrázek převzat z [25]

Bipolární režim

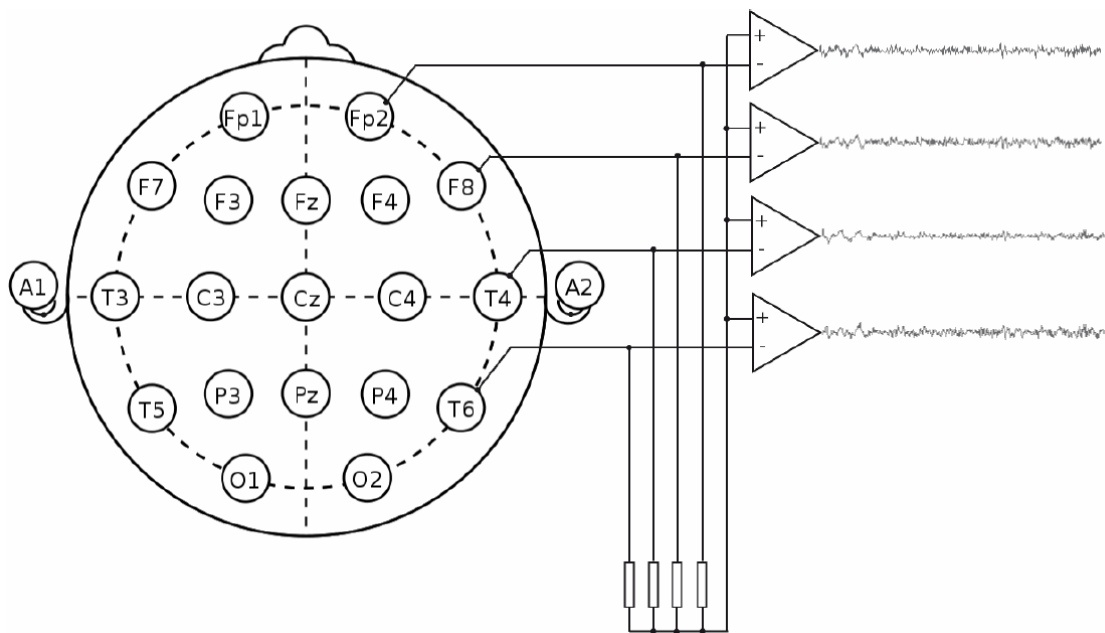
V bipolárním režimu (viz obrázek 2.18), na rozdíl od unipolárního režimu, není použita referenční elektroda. Bipolární režim se získává virtuálně odečtením hodnot naměřených na dvou sousedních elektrodách. Invertující vstup zesilovače je připojen na neinvertující vstup zesilovače následujícího. Elektrický signál v bipolárním režimu, kdy se jedná o signál mezi přilehlými elektrodami neutralizuje šum a artefakty, které jsou způsobovány společným referenčním bodem a výsledek je často snáze interpretovatelný, avšak toto zapojení je náchylnější na ovlivnění vůči specifickým artefaktům. [3; 25; 26]



Obrázek 2.18. Zapojení elektrod v bipolárním režimu. Obrázek převzat z [25]

Zprůměrovaný režim

Zprůměrovaný režim (viz obrázek 2.19) vzniká virtuálně tak, že v daném okamžiku ze všech elektrod následně určíme hodnotu aritmetického průměru (hodnota hypotetického společného potenciálu) a tuto hodnotu odečteme od hodnot naměřených ve všech kanálech. Tento typ zapojení například vyrovnává pohybové artefakty. [3; 25]



Obrázek 2.19. Zapojení elektrod ve zprůměrovaném režimu. Obrázek převzat z [25]

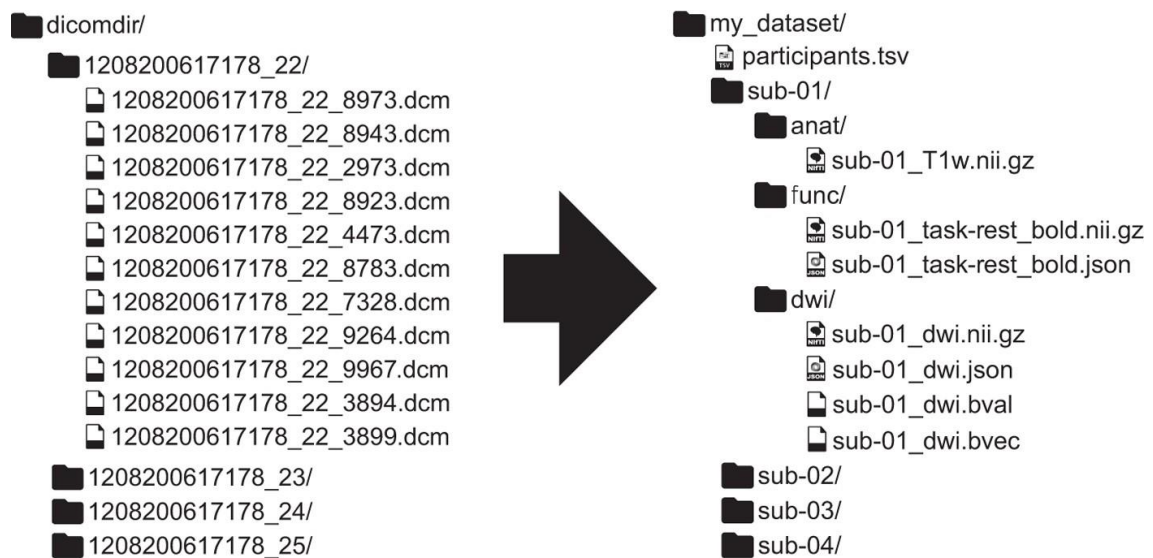
2.2.5 Simultánní měření biosignálů pro detekci artefaktů

Pro lepší identifikaci artefaktů v EEG záznamu je možné spolu s EEG nahrávat i další biosignály, jako např. EKG, EOG, EMG, EYE TRACKER. Tyto signály mohou dobře posloužit při následné detekci artefaktů, např. zkoumáním korelace s EEG. [9; 11]

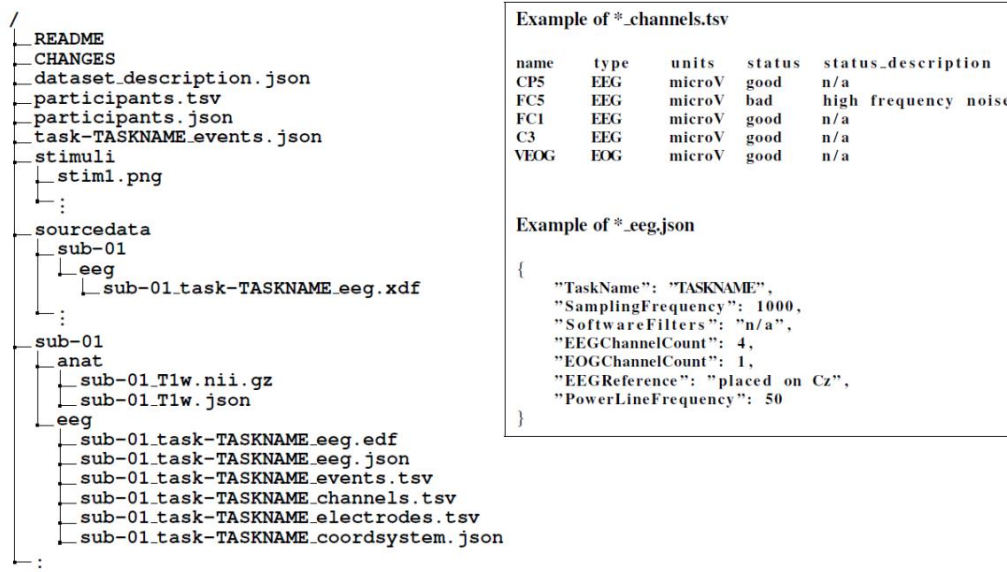
2.2.6 Datové formáty EEG

Často se v praxi stává, že jednotliví výrobci neurofyziologického vybavení a SW používají své proprietární datové formáty, místo univerzálních, pro ukládání výstupních dat měření. Výsledkem každé studie jsou komplexní data zahrnující mnoho informací v různých formátech od jednoduchých textových souborů až po složitá vícerozměrná obrazová data (např. kombinace EEG s MRI). Nedostatek konsensu v oboru vede ke zbytečnému plýtvání časem při přeskupování dat před analýzami, nebo přepisování hotových algoritmů tak, aby bylo možné otestovat data jiného formátu. Přijetí společného datového standardu celosvětově by přineslo řadu výhod, jako jsou: redukce chyb při nepochopení datové informace, rychlejší import, optimalizace SW, vývoj SW pro automatickou kontrolu konzistence datových souborů, jednotnost anotací a další. Cílem dřívějších přístupů založených na databázích bylo zejména efektivní shromažďování datových souborů oproti vytváření formátu pro jejich popis. [27]

Pro tyto účely byl v [27] představen standard BIDS (Brain Imaging Data Structure). V rámci BIDS jsou zavedeny postupy v kódování dat, např. účel, popisy kanálů/zapojení, anotace atd. Systém využívá datovou strukturu založenou na jednoduchých formátech vnořených do souboru složek a podsložek. Příklad uspořádání BIDS můžete vidět na obrázku 2.20, příklad datové sady BIDS s EEG daty na obrázku 2.21.



Obrázek 2.20. Ukázka struktury datových formátů splňující standardy BIDS. Obrázek převzat z [27]



Obrázek 2.21. Ukázka adresářového stromu datové sady BIDS, která obsahuje data EEG. Obrázek převzat z [52]

European Data Format (.edf)

Datový formát *.edf*, který byl představen v roce 1992, dodržuje obecné specifikace BIDS, kdy každý subjekt má vlastní adresář nezpracovaných dat, ve kterém jsou další adresáře pro každou relaci, popř. modalitu. K tomuto souboru jsou připojeny další soubory ve formátu *.json* nebo *.csv*, které obsahují, např: popis HW, anotace a další metadata. EDF formát je vhodný ke zpracování v široce používaných jazycích jako je MATLAB a Python, a proto je nejvhodnějším datovým formátem, pro implementaci metody testované v této diplomové práci. [28]

2.3 Používané metody

2.3.1 Metody pro segmentaci EEG signálu

Všechny metody adaptivní segmentace, ač pracují na různých principech, mají jedno společné, a to vstupní parametry, kterými jsou:

- velikost okna – počet vzorků používaného okna
- posun okna – počet vzorků, o který se okno posouvá po signálu
- prahová hodnota – hodnota hranice pro porovnání

Tyto parametry zásadním způsobem ovlivňují konečný výsledek segmentace a slouží též k ladění a optimalizaci výsledků. Správným nastavením těchto parametrů je možné dosáhnout smysluplné segmentace EEG záznamu. Použitím krátkého okna získáme vyšší časové rozlišení, ale zaplatíme za něj daň ve formě malého frekvenčního rozlišení při analýze frekvenčních parametrů. Malé okno může vytvářet výsledné segmenty s velmi malou délkou trvání a může být vhodné, např. pro odlišení samostatných epileptických hrotů v záznamu, které mají typickou délku okolo 100 ms [10]. Velké okno naopak nabízí výhodu ve formě vyššího frekvenčního rozlišení za cenu nižšího rozlišení v čase. Větší okno může být výhodné např. pro odlišení celé epileptické aktivity (trvání cca několik sekund) od aktivity pozadí. Posun okna je parametr, který udává o kolik vzorků se okno posune po signálu při další iteraci algoritmu. Posun okna má vliv zejména na konečný počet výpočtů, které musí SW provést, aby dospěl k výsledku, ale umožňuje také jistým způsobem zpřesňovat rozlišení v čase.

V kapitolách níže jsou podrobněji popsány tyto druhy segmentací EEG data:

- Fixní segmentace
- Adaptivní segmentace
 - Adaptivní segmentace na základě lineární predikce
 - Adaptivní segmentace na základě autokorelační funkce
 - Adaptivní segmentace s využitím dvou spojených oken
 - Adaptivní segmentace na základě jednoduché míry diference

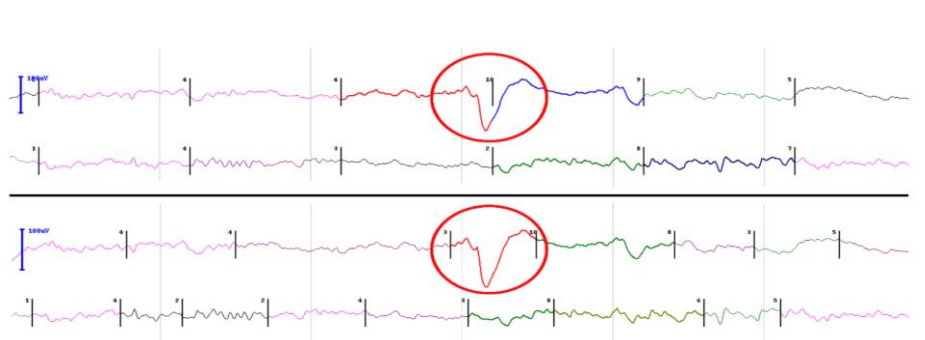
Fixní segmentace

Fixní segmentace (viz obrázek 2.22) EEG signálu je nejjednodušším přístupem z pohledu metodiky i ve vztahu k výpočetní náročnosti, kdy výsledné segmenty vždy sousedí a mají stejný počet vzorků. Tato metoda ovšem není ve vztahu k EEG signálu relevantní, protože nijak nerespektuje žádné charakteristiky signálu – jediným parametrem, kterým se zabývá, je čas. Z pravidla dochází ke vzniku tzv. hybridních segmentů, které obsahují více než jednu spojitou událost v signálu, což v následné analýze znemožňuje klasifikaci tohoto hybridního segmentu, protože při jeho parametrickém popisu obsahuje znaky všech událostí obsažených v tomto časovém řezu signálu. V druhém případě dochází k rozdělení jednoho v čase delšího artefaktu do více segmentů, což opět znesnadňuje následující klasifikaci jednotlivých segmentů. [10; 15; 29]

Adaptivní segmentace

Adaptivní segmentace (viz obrázek 2.22) kompenzuje nedostatky konstantní segmentace tím, že signál rozděljuje na segmenty proměnlivé délky v závislosti na sledovaných charakteristikách signálu. Adaptivní segmentace je velice důležitým nástrojem segmentace signálu, který byl od svého prvního představení v roce 1977 [30] nespočetněkrát modifikován. Samotné modifikace se liší zejména v různém přístupu k práci s okny (referenčním a testovacím). Druhou odlišností jsou testované parametry, které se při analýze signálu porovnávají jak mezi sebou, tak v čase. Výsledek těchto analýz je ve většině těchto modifikací porovnáván s nastavenou prahovou hodnotou, a tímto způsobem je určována finální hranice segmentu. Níže můžete najít popis modifikací adaptivní segmentace pro pochopení různých vyvinutých přístupů, a to konkrétně:

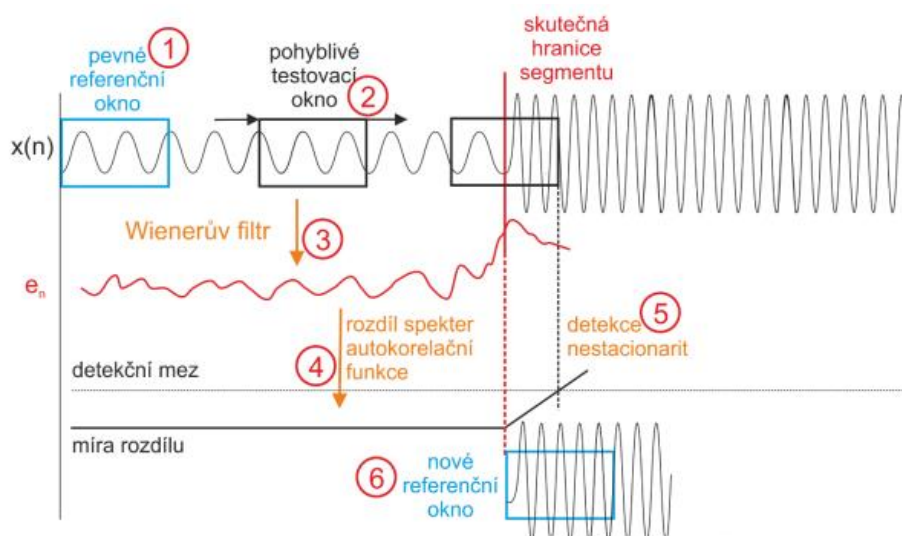
1. Adaptivní segmentace na základě lineární predikce
2. Adaptivní segmentace na základě autokorelační funkce
3. Adaptivní segmentace s využitím dvou spojených oken
4. Adaptivní segmentace na základě jednoduché míry difference



Obrázek 2.22. Rozdíl mezi fixní (nahore) a adaptivní (dole) segmentací.
Obrázek převzat z

Adaptivní segmentace na základě lineární predikce

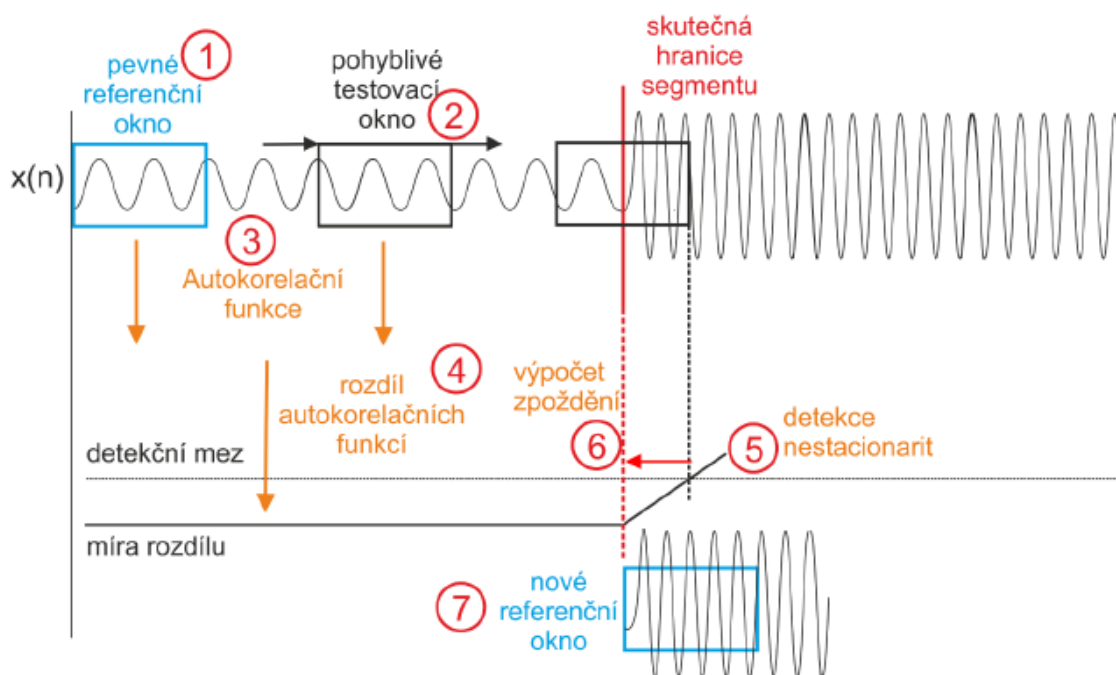
První vyvinutou metodou adaptivní segmentace (viz obrázek 2.23), je metoda z roku 1977, která využívá Wienerův algoritmus lineární predikce [30], který transformuje vstupní posloupnost do posloupnosti nekorelovaných náhodných veličin. Tento algoritmus využívá dvou oken. První okno je nepohyblivé a slouží jako reference a to tak, že jsou v něm odhadnuty koeficienty Wienerova filtru. Druhé okno postupně prochází přes signál a uvnitř tohoto okna je signál filtrován inverzním filtrem k filtru odhadnutému v pevném referenčním okně a zároveň se sleduje chyba predikce. V případě, že je signál stacionární (adaptován na signál), bude chyba predikce nekorelovaná. Pokud ovšem dojde k podstatným změnám stacionarity, filtr přestane být adaptován na signál, tzn. dojde k detekci nestacionarity. V dalším kroku je určena autokorelační funkce chyby predikce a v obou oknech se určí míra difference mezi spektry. Míra difference mezi spektry se porovnává s rozhodovací prahovou hodnotou a slouží pro umístění značky signalizující hranici segmentu doprostřed testovacího okna. Referenční okno se dále přesune na začátek nového segmentu a celý proces se opakuje. [3; 10; 29; 31]



Obrázek 2.23. Schéma algoritmu adaptivní segmentace na základě lineární predikce podle Bodensteina a Praetoriuse. Obrázek převzat z [10]

Adaptivní segmentace na základě autokorelační funkce

Metoda adaptivní segmentace na základě autokorelační funkce (viz obrázek 2.24) využívá dvojici oken podobným způsobem jako u předchozí metody. Jediným rozdílem je, že se zde nezjišťují koeficienty Wienerova filtru, ale určuje se míra rozdílnosti signálů obsažených v referenčním a testovacím okně s využitím autokorelační funkce. V případě, že míra rozdílnosti těchto dvou signálů překročí nastavenou prahovou hodnotu, pak je toto místo identifikováno jako změna stacionarity. Výsledná poloha je pak určena díky předpokladu, že k překročení nastaveného thresholdu dochází o chvíli později než ke skutečné změně stacionarity signálu. [3; 10; 29]



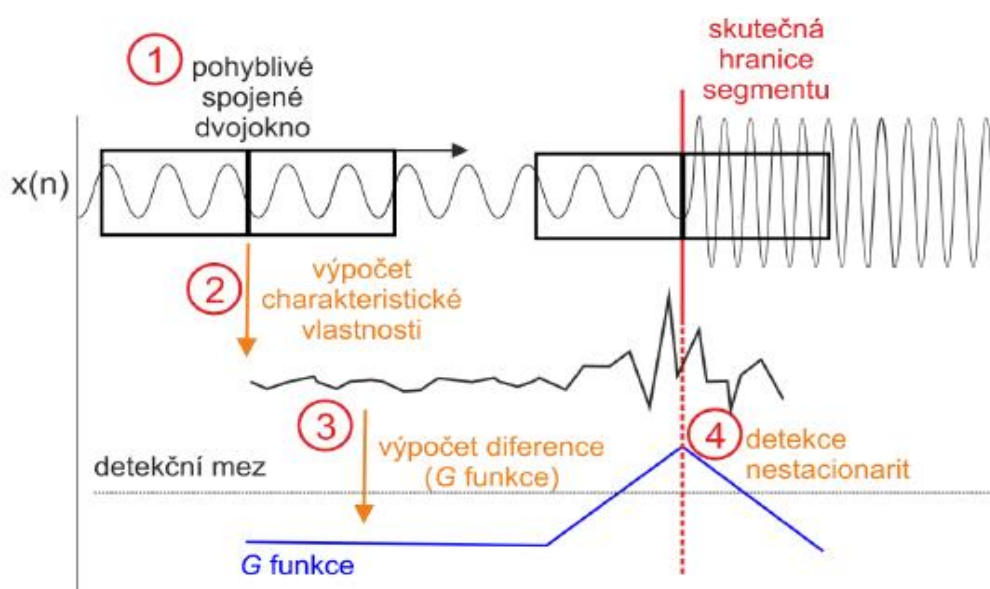
Obrázek 2.24. Schéma algoritmu adaptivní segmentace pomocí rozdílu autokorelačních funkcí v referenčním a testovacím okně. Obrázek převzat z [10]

Metoda adaptivní segmentace na základě autokorelační funkce byla později zdokonalena novým přístupem k referenčnímu oknu, kdy jeho velikost nezůstává stacionární, ale s postupným pohybem testovacího okna narůstá do doby, než je detekována změna stacionarity signálu. Tato změna umožňuje při porovnávání signálu těchto dvou oken zachovat maximální možnou informaci týkající se signálu v referenčním okně. [10]

Obě metody tohoto typu adaptivní segmentace nejsou vhodné pro multikanálovou segmentaci, protože napříč kanály dochází k desynchronizaci a různému posunu signálových křivek na časové ose, tzn. provedení „řezu“ napříč všemi kanály by způsobilo nevratnou chybu segmentace. [10; 29]

Adaptivní segmentace s využitím dvou spojených oken

Dalším přístupem k adaptivní segmentaci je novější přístup k práci s referenčním a testovacím oknem (viz obrázek 2.25), který poprvé představil Skrylev v roce 1984 a později ho v roce 1991 upravil Varri. Tento přístup je od předchozích odlišný v tom, že referenční okno klouže spolu s testovacím oknem a pro každé okno se zjišťuje charakteristická vlastnost signálu. V dalším kroku se určuje diference těchto dvou oken někdy nazývána G funkce [32], nebo Totální diference [3; 33] a při překročení nastavené prahové hodnoty je umístěna značka signalizující hranici segmentu v místě, kde jsou tato okna spojena. Tento přístup práce s okny zamezuje desynchronizaci napříč kanály, a proto je zatím nejvhodnější ze zmíněných metod. [3; 10]



Obrázek 2.25. Obecné schéma metody adaptivní segmentace využívající pohyblivého spojeného dvojokna, které se posunuje po signálu. G funkce představuje výslednou diferenci mezi referenčním a testovacím oknem. Obrázek převzat z [10]

Adaptivní segmentace na základě jednoduché míry diference – Varriho metoda

Varriho metoda adaptivní segmentace (viz obrázek 2.26) využívá výše zmíněného principu spojených dvou oken. Existuje však více možných parametrů, díky kterým se určuje míra diference mezi okny, jako např. amplitudová a frekvenční diference signálu [33], nebo nelineární energetický operátor [32], nebo vylepšený nelineární energetický operátor s předcházející vlnkovou transformací [34]. V tomto přístupu se z míry diference určuje totální diference a značka pro hranici segmentu se umísťuje tam, kde totální diference mezi okny přesáhla stanovenou prahovou hodnotu. [10]

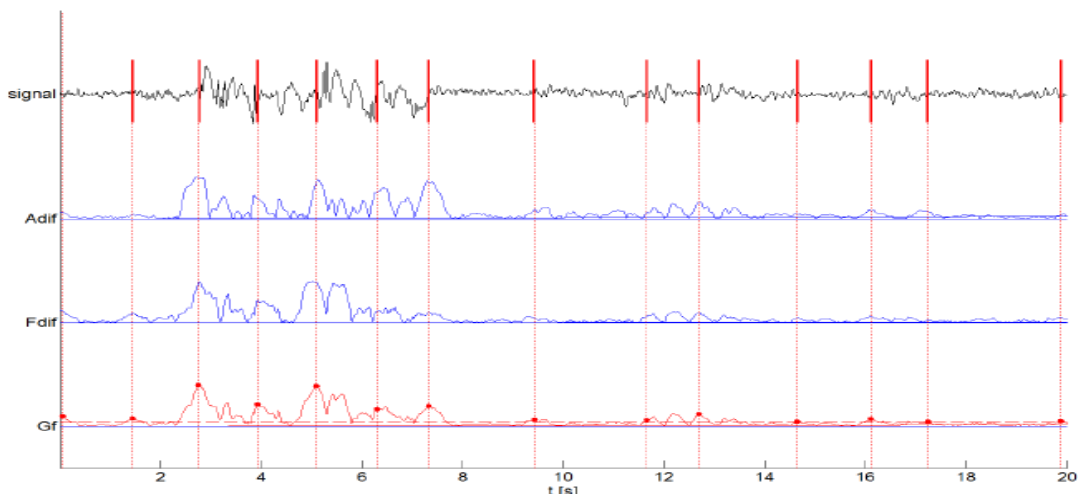
Algoritmus, který navrhl Krajča v roce 1991 využívá výše zmíněnou amplitudovou a frekvenční diferenci signálu, a z nich poté zjišťuje takzvanou totální diferenci (jinde nazývanou: globální funkce, G funkce, Gf) podle následujících rovnic 2.1-2.3. Krajčův algoritmus využívá velice jednoduché matematické operace, což ho činí výpočetně nenáročným systémem pro segmentaci dat, který je vhodný i pro online vyhodnocování EEG záznamů.

$$Adif = \frac{\sum_{i=1}^N |x_i|}{N} \quad (2.1)$$

$$Fdif = \frac{\sum_{i=1}^N |x_i - x_{i-1}|}{N} \quad (2.2)$$

$$Gf = k_1 \cdot |Adif_L - Adif_R| + k_2 \cdot |Fdif_L - Fdif_R|, \quad (2.3)$$

kde N je velikost okna, koeficienty L a R značí levé/pravé (left/right) okno, k_1 a k_2 jsou empiricky nastavené koeficienty. Koeficient k_1 je nastaven na hodnotu 1 a koeficient k_2 je nastaven na hodnotu 7.



Obrázek 2.26. Varriho algoritmus (Krajča a kol. 1991) adaptivní segmentace. Popis křivek shora: segmentovaný signál, amplitudová diference, frekvenční diference, G funkce. Obrázek převzat z [10]

2.3.2 Metody pro klasifikaci EEG signálu

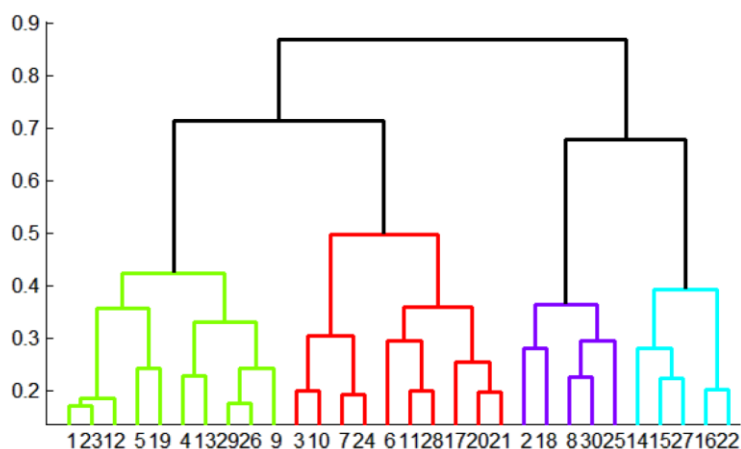
Klasifikace EEG spočívá v rozpoznání událostí vyskytujících se v segmentech záznamu a jejich rozřazení do tříd (shluků) podle podobných vlastností, např. fáze spánku, různé typy artefaktů, epilepsie atd. Klasifikační metody můžeme rozdělit na metody s učitelem a metody bez učitele.

Klasifikátor s učitelem je závislý na velmi robustní množině trénovacích dat (etalon), na kterých se učí vzory jednotlivých shluků, do kterých později na základě podobnosti přiřazuje segmenty zkoumaného EEG záznamu. Mezi tyto metody se řadí např. neuronové sítě.

Klasifikátory bez učitele, tedy bez trénovací množiny, jsou založeny pouze na vnitřní struktuře dat a jejich matematickém vztahu. Mezi tyto metody se řadí např. hierarchické shlukování, k-means, DBSCAN.

Hierarchické shlukování

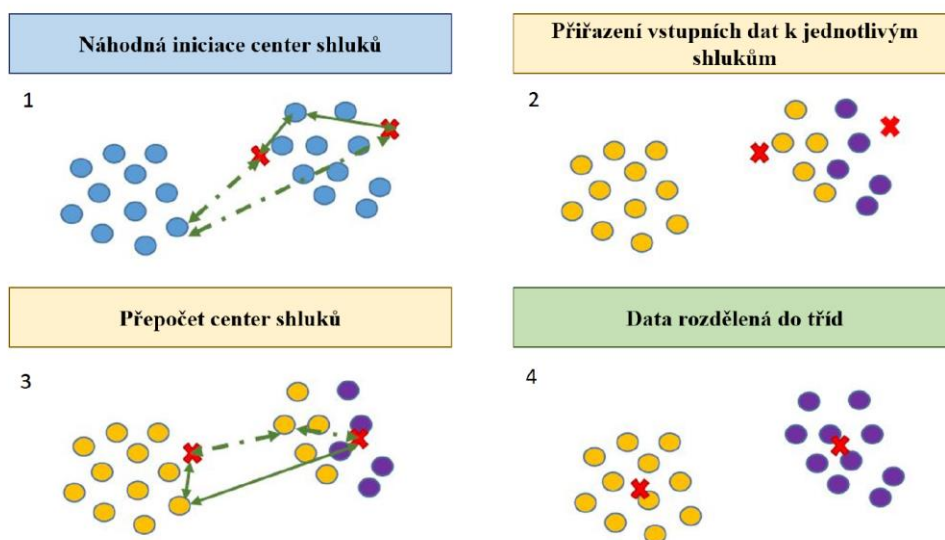
Principem hierarchického shlukování je stromová struktura, tzv. dendrogram (viz obrázek 2.27), kde jednotlivé uzly představují shluky, vodorovné úrovně dendrogramu jsou stupně rozkladu shluků a svislý směr reprezentuje vzdálenost mezi shluky. Na začátku celého procesu (viz obrázek 2.27) je každý objekt přiřazen do samostatného shluku, tedy dostáváme stejný počet shluků jako vstupních objektů a je nutné definovat pravidla hodnocení podobnosti vztahů mezi shluky. Ve druhém kroku algoritmus hledá nejpodobnější pár shluků, který spojí, a tím klesne celkový počet shluků o jeden. Každé seskupení vzniká ve větší vzdálenosti od shluků z předchozího uskupení. Algoritmus takto pokračuje, dokud nedosáhne $N-1$ kroku, kdy N je počet segmentů vstupujících do analýzy. Shlukování může být ukončeno na základě vzdálenosti mezi shluky, nebo při získání požadovaného počtu shluků. [35]



Obrázek 2.27. Hierarchické shlukování - dendrogram. Obrázek převzat z [35]

K-means

Metoda k-means (viz obrázek 2.28) patří mezi nehierarchické metody shlukování, kdy je nutné na vstupu algoritmu zadat požadovaný počet shluků, do kterých chceme objekty roztrždit. Princip algoritmu je založen na postupném přesouvání objektů na základě výpočtu vzdálenosti (např. Euklidovské, Mahalanobisovi, Hammigovi). V kroku 1 dochází k náhodné iniciaci center shluků (tzv. centroidů) v prostoru, kdy každý centroid reprezentuje jeden z výsledných shluků. V kroku 2 je vypočítána vzdálenost zkoumaných objektů od centroidů a objekty jsou na základě této vzdálenosti přiřazeny k nejbližšímu centroidu. V rámci takto vzniklých shluků jsou v kroku 3 vypočítána těžiště a do těchto těžišť jsou umístěny příslušné centroidy. Proces se takto opakuje, dokud se poloha centroidů neustálí, nebo není dosaženo uživatelem předem nastaveného počtu iterací – krok 4. [13; 36]



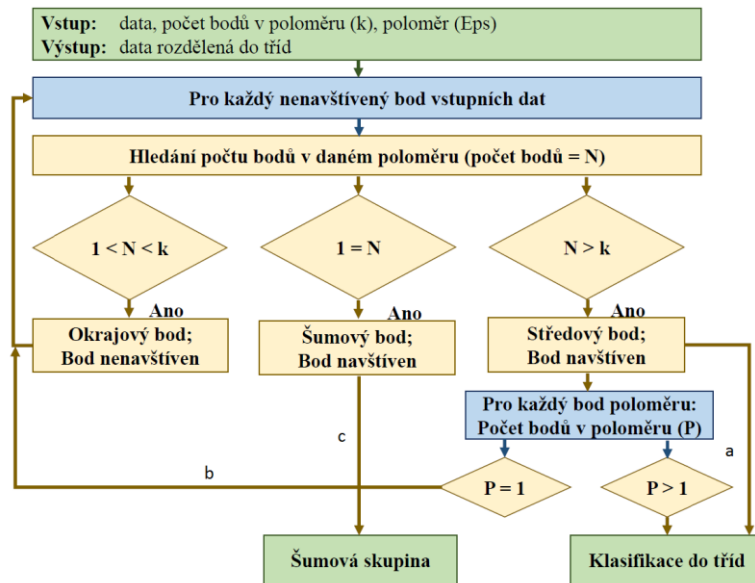
Obrázek 2.28. Princip metody k-means. Obrázek převzat z [36]

DBSCAN

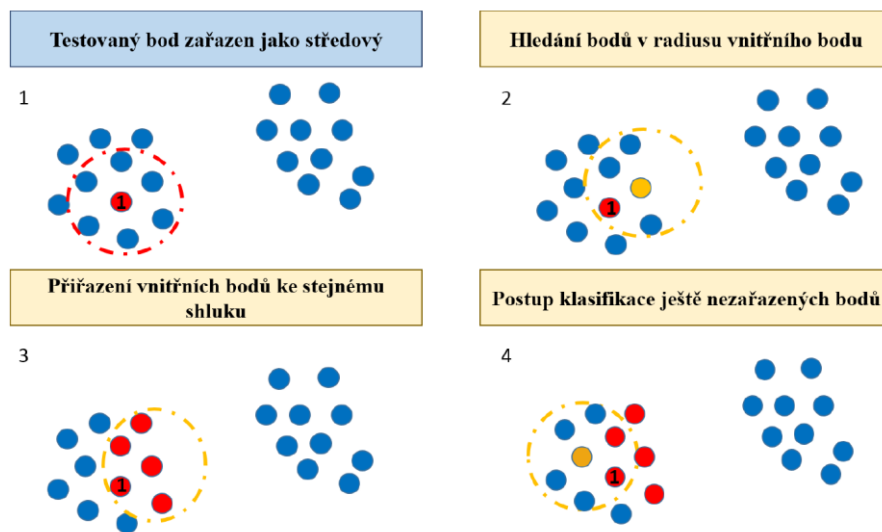
DBSCAN je neparametrická shluková analýza založená na hustotě (viz obrázek 2.29 a 2.30), která vzhledem ke svému unikátnímu přístupu byla několikrát modifikována (např. DMDBSCAN, GRIDBSCAN). Algoritmus rozřazuje body do výsledných shluků na základě jejich hustoty v prostoru, přičemž definuje shluk jako oblast s vysokou hustotou bodů, která je ohraničená oblastí s nízkou hustotou bodů. Na rozdíl od výše zmíněné metody k-means, DBSCAN nepotřebuje na vstupu zadat výsledný počet shluků, protože nezávisí na iniciačním místě, tedy je výsledný počet shluků a rozřazení dat výstupem algoritmu DBSCAN. DBSCAN začíná s náhodně vybraným bodem dat, které chceme klasifikovat a na základě jeho blízkého okolí hledá počet bodů, které k němu spadají, proto je nutné na vstupu nastavit parametry: poloměr (Eps), počet bodů v poloměru (k). Na základě těchto vstupních parametrů DBSCAN rozděluje jednotlivé body do tří skupin [36]:

- Okrajový bod = Pokud je počet bodů v zadaném poloměru shodný s nastaveným počtem bodů v poloměru, označí se okraj shluku
- Středový bod = Pokud je počet bodů v zadaném poloměru vyšší než nastavený počet bodů v poloměru.
- Šumový bod = Pokud je počet bodů v zadaném poloměru nižší než nastavený počet bodů v poloměru, označí se jako odlehlý (šumový) bod

Klíčovým rozhodovacím kritériem je počet bodů v zadaném poloměru (viz obrázek 2.30). Je-li počet bodů v poloměru dostatečný, pak je bod přiřazen do shluku (trasa a). Poté jsou testovány jednotlivé body v tomto poloměru, které jsou stejným kritériem buď přiřazeny do shluku, nebo znovu testovány (trasa b). Lokalizované šumové body jsou algoritmem zařazeny do speciálního shluku (trasa c). [36]



Obrázek 2.29. Klasifikace do tříd metodou DBSCAN. Obrázek převzat z [36]



Obrázek 2.30. Klasifikace do shluků metodou DBSCAN. Obrázek převzat z [36]

Porovnání algoritmů DBSCAN a k-means

V tabulce 2.1 můžete vidět srovnání výhod a nevýhod algoritmu DBSCAN a k-means. U klasifikace algoritmem k-means je zásadní problém při nastavení počtu tříd na vstupu algoritmu, protože bez manuálního procházení EEG záznamu nejsme schopni určit počet různých událostí, tedy shluků, do kterých k-means segmenty roztrídí. Výhodu variabilního počtu shluků, kterou má DBSCAN, ale provází vyšší výpočetní náročnost.

Tabulka 2.1. Porovnání vlastností algoritmů k-means a DBSCAN podle [36]

K-means	DBSCAN
+ rychlost	+ detekce prolnutých shluků
+ komplexnost – různá vstupní data	+ separace šumu
+ variabilita změnou výpočtu vzdáleností	+ velký počet dostupných modifikací
+ jednoduchost	+ automatický počet shluků
- vstupní parametr: počet shluků	- vysoká výpočetní náročnost
- neseparuje prolnuté shluky	- chyba při nerovnoměrné hustotě

2.3.3 Neuronové sítě v EEG

Neuronová síť jako taková vzniká propojením jednotlivých neuronů mezi sebou. Nejčastěji jsou v síti uspořádány do několika vrstev, kdy sousedící vrstvy jsou propojeny a poslední výstupní vrstvy mají pouze výstup. Jejich hlavní výhodou je schopnost hlubokého učení a adaptace na vstupní data, čehož dosahuje změnami vlastních vnitřních nastavení parametrů v kontextu dosažené úspěšnosti predikce. [37; 38]

Aplikace neuronových sítí při klasifikaci EEG s celkovým rozvojem UI exponenciálně stoupá i přes jejich nevýhodu ve formě delší doby výpočtu a nutné velké množiny trénovacích dat pro jejich nastavení. UI je v současnosti využívána v mnoha oblastech, jako jsou: rozpoznávání emocí, řešení úkolů motorického zobrazování, měření duševní zátěže, detekce epileptických záchvatů, hodnocení fází spánku, detekce abnormálního EEG. Použití hlubokého učení v praxi pro výše zmíněné aplikace je ovšem opět omezeno zarušením EEG signálu různými druhy artefaktů, které je třeba pro správný chod hlubokého učení z dat odstranit, což ve většině současných studií probíhá ručně, a tedy i s velkými časovými nároky na specialisty. Další využívanou možností předzpracování dat pro hlubokému učení je odstranění artefaktů s využitím ICA, popř. filtrací ve frekvenční oblasti, které ovšem nejsou úplně optimálním řešením, protože odstraňují ze záznamu i jiná než artefaktová/abnormální data, která by při precizním ručním procházení záznamu z dat odstraněna nebyla. [37; 38]

Konvoluční neuronové sítě

Konvoluční neuronové sítě (CNN) jsou podmnožinou skupiny neuronových sítí. Jejich hlavními znaky je použití konvolučních a tzv. pooling vrstev. V poslední době dosahují obrovských pokroků zejména při zpracování dat velkých rozměrů, kterými mohou být např. obrázky. Jejich učení vychází z poznatků člověka, který pro ně vytvoří učící množinu dat a konvoluční neuronová síť po mnoha iteracích a změnách svých vnitřních parametrů optimalizuje své nastavení tak, aby bylo dosaženo minimalizace chyby predikce. V oblasti EEG se tento typ neuronových sítí uplatňuje zejména jako klasifikátor, do kterého vstupují spektrogramy. [37; 38]

3 Cíle práce

Cílem této práce je vytvoření metodiky automatizované klasifikace EEG záznamu založené na adaptivní segmentaci s využitím analýzy v časové a frekvenční oblasti. Dalším cílem je tuto metodiku realizovat v programovém prostředí MATLAB a tuto metodiku otestovat na simulovaných a reálných datech a kvantitativně vyhodnotit její efektivitu skrze měřitelné metriky. Dalším cílem je na základě výsledků a jejích dílčích procesů navrhnout příznaky a procesy pro řešení klasifikace EEG záznamu metodami založenými na kooperaci s umělou inteligencí využívající hluboké učení.

4 Metodika

4.1 Hypotetický základ a princip navrhované metody

Z dosavadního studia metod segmentací a klasifikací EEG záznamu vzešlo několik závěrů a zásad, které vedly ke stanovení několika hypotéz zmíněných v této kapitole, které jsou základními kameny navrhované metody.

4.1.1 Hypotéza 1: Komplementární efekt různě dlouhých oken

Použití různě dlouhých oken jednotlivě v kombinaci s dalšími parametry má nestejnorodou úspěšnost a v různých nastaveních přináší nepredikovatelnou míru úspěchu.

Úspěšnost segmentace dat s využitím oken je přímo závislá na velikosti daného okna (počet vzorků) a různé nastavení této délky přináší různorodé výsledky. Souvisí to zřejmě s různě dlouhými projevy událostí v EEG záznamu, kdy úspěšnost klesá u událostí výrazně delších než okno a zároveň klesá u událostí výrazně kratších než okno. Pak je vlastně největší výzvou volba optimální velikosti okna, což je prakticky nemožné kvůli individualitě každého EEG záznamu, která se projevuje jak v síle a intenzitě některých projevů, tak v přítomnosti/nepřítomnosti a délce různých artefaktů a jiných abnormálních událostí. Například v čase velice krátké a intenzivní elektrodové artefakty způsobené nedostatečným množstvím vodivého gelu/roztoku pod elektrodou zaplní signál artefakty, jejichž množství a charakter nutně zkreslí celkový záznam a hodnoty parametrů jsou pak ovlivněny jejich příspěvkem.

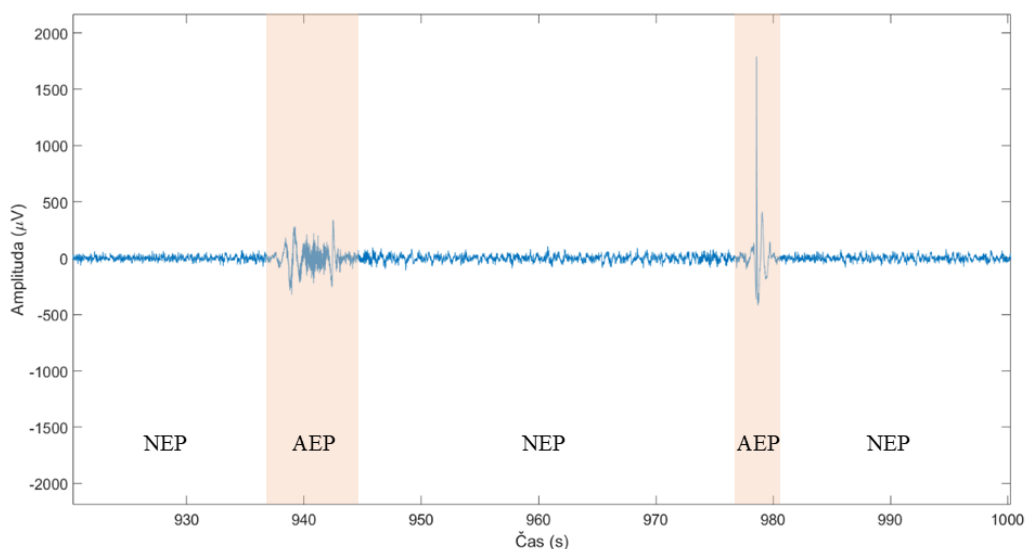
Řešením by mohlo být použití různě dlouhých oken, které by mohly komplementárně přispět k přesnějším výsledkům segmentace přírůstkovým způsobem, tzn. výsledky všech oken se agregují. V navržené metodě budou použity 3 různě dlouhá okna o délkách 256, 512 a 1024 vzorků a předpokládám, že kumulované výsledky za všechna okna budou úspěšnější než každé jednotlivé okno zvlášť.

4.1.2 Hypotéza 2: Abnormalitu lze definovat pouze normalitou

Když se základním pohledem trochu abstraktně podíváme na EEG signál, je vysvětlení podstaty artefaktů definováno jako nestandardní chování naměřeného signálu, nebo jako extrémní výkyv hodnot oběma směry s naprosto nesourodou a neočekávanou intenzitou a délkou v čase. To znamená, že vedle těchto výkyvů a extrémních událostí v záznamu musí existovat klidová (normální) část signálu, která je běžnou klidovou činností mozku. Takovéto normální činnosti by v záznamu měla být většina a měla by tedy být nejlépe popsitelná pomocí standardních parametrů a jejich prostřednictvím pak nejjednodušeji identifikovatelná stacionární část celého EEG záznamu.

Pro potřeby implikace této hypotézy do navrhované metody adaptivní segmentace EEG záznamu jsem definoval dvě tyto části jako NEP („Normal Event Part“ – běžná klidová činnost mozku) a AEP („Abnormal Event Part“ – artefakty a jiné grafoelementy).

Na rozdíl od ostatních metod používaných pro segmentaci dat, které se soustředí na nalezení nestacionarit a diferencí, tedy definici abnormalit, se mnou navržená metoda naopak soustředí na definici a nalezení NEP, jehož prostřednictvím pak jednoduše oddělíme AEP (viz obrázek 4.1). Teoreticky musí každý AEP ležet mezi dvěma NEP, což zároveň definuje jeho začátek i konec, protože konec jednoho NEP je začátkem následujícího AEP, a naopak začátek každého NEP je koncem předcházejícího AEP. Tento přístup by měl být realizovatelný pomocí jednoduchých dotazů a měl by výrazně snížit nároky na výpočetní čas, HW vybavení atd.



Obrázek 4.1. Ilustrativní obrázek NEP a AEP

4.1.3 Hypotéza 3: Hranice mezi NEP a AEP není bod, ale časový úsek

Při analyzování anotovaných EEG záznamů si lze všimnout, že pokud záznam anotuje více než jeden diagnostik, shodují se skoro stoprocentně v identifikaci jednotlivých artefaktů, ale naprosto se odlišují v identifikaci začátků a konců těchto artefaktů. Stejný jev můžeme pozorovat i v případě výsledků strojových analýz, které se v identifikaci začátků/konců také značně liší. Hypoteticky to v rámci teorie existence NEP a AEP znamená, že pokud budeme vnímat každý jednotlivý přechod mezi NEP a AEP jako náhodně dlouhý časový úsek, nikoli jako jeden konkrétní bod, může to přispět k lepší identifikaci AEP. Pokud budeme vnímat výsledky analýz oken různých délek jako všechny pravdivé, pak tento časový úsek musí ležet mezi dvěma krajními hodnotami na časové ose, proto je určení začátku a konce AEP podmíněno minimálně dvěma a více analýzami.

4.2 Technické řešení navrhované metody adaptivní segmentace

Nejprve je v kapitole 4.2.1 je uveden popis a matematický základ sledovaných parametrů (statistické, frekvenční, vycházející z tvaru), které byly využity při návrhu metody adaptivní segmentace

V kapitole 4.2.2 naleznete schéma principu navrhované metody adaptivní segmentace. Celý proces byl rozdělen do tří fází, které probíhaly samostatně a nezávisle pro jednotlivá okna. Jednotlivé fáze jsou detailněji rozebrány v kapitolách 4.2.3–4.2.5.

Po trojfázové analýze třemi různými okny byly jejich jednotlivé výsledky sloučeny do finálního výsledku segmentace dat. Tento proces je popsán v kapitole 4.2.6.

4.2.1 Sledované parametry

Sledované statistické parametry

Všechny statistické parametry zmíněné v této kapitole byly vypočítány z hodnot amplitudy signálu jednotlivých segmentů a její 1. a 2. derivace.

1. Maximum (MAX)

Maximum je nejvyšší hodnotou v datech. V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *max*.

2. Minimum (MIN)

Minimum je nejnižší hodnotou v datech. V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *min*.

3. Aritmetický průměr (MEAN)

Aritmetický průměr je odhadem střední hodnoty výběru, který se řadí mezi nejdůležitější a nejpoužívanější charakteristiky statistických souborů. V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *mean*. Matematický vzorec výpočtu aritmetického průměru:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad (4.1)$$

kde x představuje jednotlivé proměnné a N rozsah výběrového souboru. [39]

4. Výběrový rozptyl (VAR)

Výběrový rozptyl popisuje variabilitu výběrového souboru. Jedná se o průměrnou hodnotu kvadrátu odchylek jednotlivých hodnot výběru od střední hodnoty daného výběru (v našem případě aritmetického průměru). V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *var*. Matematický vzorec výpočtu výběrového rozptylu:

$$VAR = \frac{1}{N-1} \sum_{i=1}^N |x_i - \bar{x}|^2, \quad (4.2)$$

kde x představuje jednotlivé proměnné, N je rozsah výběrového souboru a \bar{x} je aritmetický průměr. [39]

5. Výběrová směrodatná odchylka (SMO)

Výběrová směrodatná odchylka je definována jako druhá odmocnina z výběrového rozptylu a jedná se vlastně o míru rozptylu hodnot sledovaného znaku kolem střední hodnoty výběru (v našem případě aritmetického průměru). V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *std*. Matematický vzorec výpočtu výběrové směrodatné odchylky:

$$SMO = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |x_i - \bar{x}|^2}, \quad (4.3)$$

kde x představuje jednotlivé proměnné, N je rozsah výběrového souboru a \bar{x} je aritmetický průměr. [39]

6. Modus (MOD)

Modus představuje nejčetnější (nejčastěji se vyskytující) kvantitativní hodnotu studovaného souboru. V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *mode*.

7. Medián (MED)

Medián je kvantitativní prvek souboru uspořádaného v neklesajícím pořadí, který tuto řadu rozděluje na poloviny, kdy jedna polovina znaků má menší hodnotu než medián a druhá polovina má hodnotu vyšší než medián. V případě lichého počtu prvků seřazeného souboru se medián nachází přesně uprostřed řady. V případě sudého počtu prvků je hodnota mediánu vypočítána jako aritmetický průměr sousedních kvantitativních hodnot prvků nacházející se okolo pomyslného středu řady. V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *median*.

8. Koeficient šikmosti (SKEW)

Koeficient šikmosti je definován jako aritmetický průměr třetích mocnin odchylek jednotlivých prvků souboru od jejich aritmetického průměru vydělený třetí mocninou směrodatné odchylky. Jedná se o charakteristiku náhodného rozdělení, která popisuje míru asymetrie dat kolem střední hodnoty souboru. Kladná hodnota koeficientu šikmosti znamená, že napravo od střední hodnoty se nachází odlehlejší hodnoty a naopak. V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *skewness*.

Matematický vzorec výpočtu koeficientu šikmosti:

$$SKEW = \frac{E(x - \bar{x})^3}{\sigma^3}, \quad (4.4)$$

kde x představuje jednotlivé proměnné, \bar{x} je aritmetický průměr a σ je směrodatná odchylka. [40]

9. Koeficient špičatosti (KURT)

Koeficient špičatosti je charakteristika, která porovnává rozdělení náhodné veličiny s normálním rozdělením pravděpodobnosti. Kladná hodnota koeficientu špičatosti znamená, že křivka hustoty pravděpodobnosti je špičatější než u rozdělení normálního, v případě záporné hodnoty koeficientu je křivka naopak plošší. V programovacím prostředí MATLAB je pro jeho výpočet implementována funkce *kurtosis*.

Matematický vzorec výpočtu koeficientu špičatosti:

$$KURT = \frac{E(x - \bar{x})^4}{\sigma^4}, \quad (4.5)$$

kde x představuje jednotlivé proměnné, \bar{x} je aritmetický průměr a σ je směrodatná odchylka. [40]

Frekvenční analýza

Biologické záznamy, včetně EEG, jsou považovány za časové řady. Frekvenční analýza takovýchto signálů je jedním z nejdůležitějších a nejvíce relevantních nástrojů, protože i samotný lékař při hodnocení EEG vyhodnocuje jednotlivé frekvenční složky, které jsou obsaženy v daném úseku signálu. Abnormality je možné na základě frekvenční analýzy rozpoznávat mnoha způsoby, např. porovnáním hodnoty prahu (ať už pevně nastaveného, nebo matematicky určeného) s hodnotami ve spektru. [1; 3; 9; 15; 41]

Diskrétní Fourierova transformace – DFT

Fourierova transformace je základním algoritmem pro převádění signálu z časové do frekvenční oblasti a lze ji použít jak pro spojitá, tak pro diskrétní data. Samotný signál EEG sice patří do kategorie spojitých signálů, ale z důvodu záznamového zařízení, ve kterém probíhá vzorkování a kvantování signálu je v důsledku záznam EEG signálu diskretizován. K převodu diskrétních signálů z časové oblasti do frekvenční slouží tzv. diskrétní Fourierova transformace (DFT). Výpočtem DFT dochází k vyjádření signálu pomocí součtu mnoha komplexních exponenciál, což s využitím Eulerova vztahu (viz rovnice 4.6) vede na součet sinových a cosinových signálů, ze kterých je možné zjistit obsažené frekvence v signálu, tedy frekvenční spektrum signálu. Tato transformace se používá na konečné posloupnosti vzorků v čase, přičemž po aplikaci převodu do frekvenční oblasti má výsledné frekvenční spektrum signálu stejný počet vzorků jako vstupující diskrétní časová posloupnost. Z toho plyne, že čím delší je konečná posloupnost vzorků vstupujících do Fourierovy transformace, tím získáme vyšší rozlišení ve frekvenčním spektru. DFT je definována podle rovnice 4.7 a její výpočetní náročnost (počet matematických operací) je určena podle rovnice 4.8. [1; 3; 9; 15; 41]

$$e^{-j\omega} = \cos(\omega) - j \cdot \sin(\omega), \quad (4.6)$$

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}nk}, \quad (4.7)$$

$$O_{DFT} = N^2, \quad (4.8)$$

kde $X(k)$ označuje Fourierovo spektrum, $x(n)$ je vstupní časová posloupnost, n je index pořadí vzorku, k je koeficient umístění na frekvenční ose, N je celkový počet vzorků signálu, O_{DFT} je počet matematických operací potřebných k převodu signálu z časové do frekvenční oblasti.

Rychlá Fourierova transformace (FFT)

Z důvodu velké výpočetní náročnosti výše zmíněné DFT byl vyvinut algoritmus rychlé Fourierovy transformace (FFT). Jedná se o rekurzivní postup výpočtu, který pracuje na principu postupného rozdělování vstupní posloupnosti dat, která zmenšila počet operací nutných pro zjištění výsledného frekvenčního spektra (viz rovnice 4.9). [1; 2; 3]

$$O_{FFT} = N \cdot \log_2 N, \quad (4.9)$$

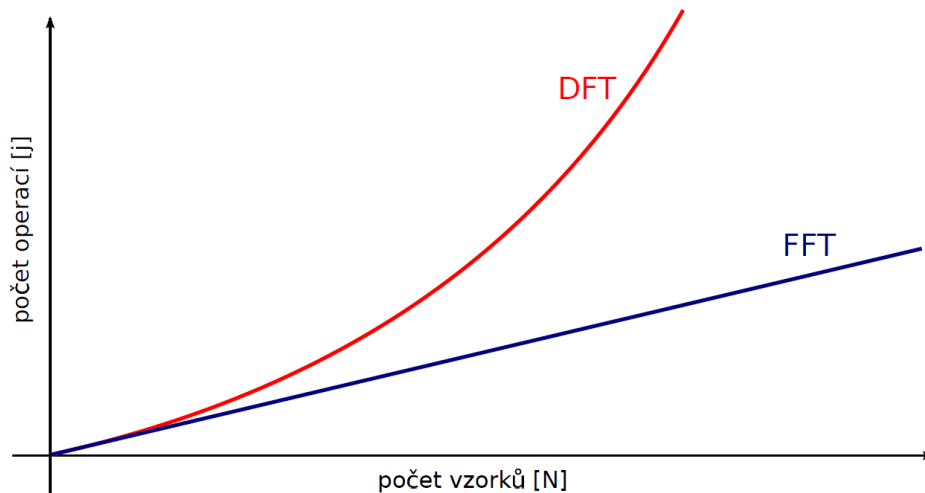
kde N je počet vzorků vstupujících do transformace.

Algoritmus rychlé Fourierovy transformace je v programovacím prostředí MATLAB implementován jako funkce *fft*. Jedinou podmínkou, kterou je nutné dodržet pro správné použití této funkce, je vstupní délka signálu, která musí být rovna mocnině čísla 2. V případě, že toto kritérium není dodrženo, funkce automaticky doplní vstupní datovou posloupnost nulami na první nejbližší hodnotu mocniny čísla 2, což v důsledku může vést ke zkreslení výstupu transformace.

Rozdíly mezi rychlostí výpočtu DFT a FFT (viz tabulka 4.1 a obrázek 4.2), tedy počtem nutných matematických operací pro provedení úkonu, stoupá s velikostí vstupního datového souboru.

Tabulka 4.1. Ukázka rozdílů ve výpočetní náročnosti DFT a FFT v závislosti na velikosti vstupního signálu

Velikost vstupního signálu N	$O_{DFT} = N^2$	$O_{FFT} = N \cdot \log_2 N$
128	16284	896
256	65536	2048
512	262144	4608
1024	1048576	10240



Obrázek 4.2. Porovnání algoritmů DFT a FFT. Obrázek převzat z [41]

Sledované frekvenční parametry

Frekvenční parametry využívané v této práci jsou založeny na výpočtu odhadu frekvenčního spektra signálu v daném okně s využitím funkce *fft* pro rychlou Fourierovu transformaci. Výsledný odhad frekvenčního spektra byl následně umocněn dvěma tak, abychom získali odhad výkonové spektrální hustoty a zároveň zesílili rozdíly mezi jednotlivými frekvenčními složkami. Následně bylo výkonové spektrum rozděleno do diagnosticky významných pásem EEG (delta, theta, alfa, beta, gama – viz kapitola 2.1.2) tak, že pro každé pásmo byly sečteny odpovídající složky odhadu výkonového spektra a převedeny na relativní hodnoty v procentech vydělením hodnoty pro dané pásmo hodnotou sumy všech spektrálních složek odhadu výkonového spektra s využitím funkce *Compute_RP* [42].

Frekvenčními parametry používanými v této práci tedy jsou:

- Relativní delta (1-4 Hz)
- Relativní theta (4-8 Hz)
- Relativní alfa (8-14 Hz)
- Relativní beta (14-30 Hz)
- Relativní gama (30-50 Hz)

Sledované parametry vycházející z tvaru signálu

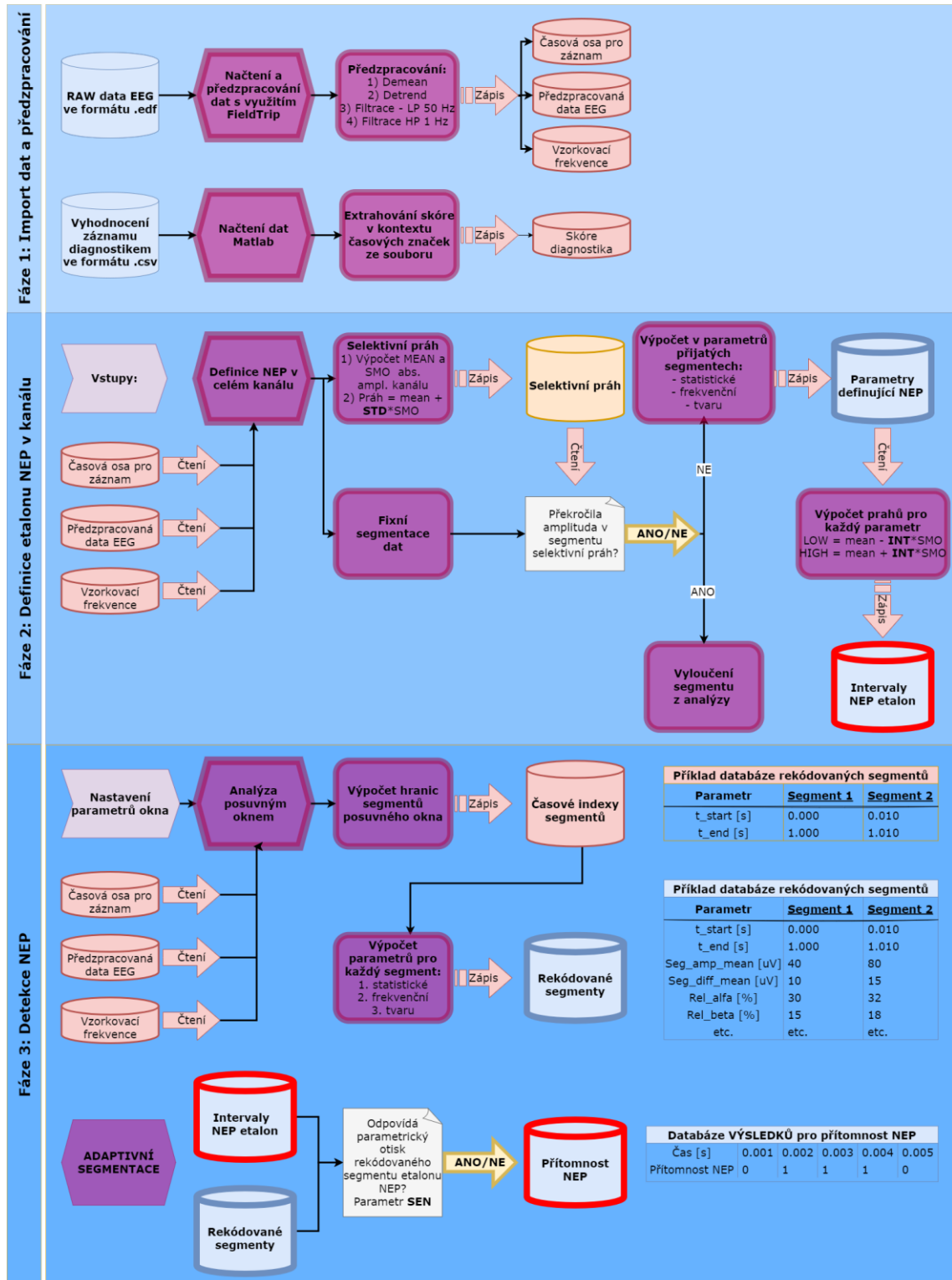
Mezi parametry vycházející z tvaru signálu byly zařazeny parametry získané pomocí identifikace lokálních extrémů (maxim a minim) v plovoucím okně s využitím implementovaných funkcí v programovacím prostředí MATLAB *islocalmin*, *islocalmax*. Lokální extrémy jsou podle definice body, ve kterých dochází ke změně znaménka první derivace. Počet lokálních maxim/minim byl získán součtem výskytu lokálních maxim/minim v segmentu. Relativní zastoupení lokálních maxim/minim bylo získáno jako počet jejich výskytů vydělený celkovým počtem vzorků segmentu. Parametry využívané v této práci tedy jsou:

- Počet lokálních maxim
- Počet lokálních minim
- Relativní zastoupení lokálních maxim
- Relativní zastoupení lokálních minim

Parametry vycházející z tvaru signálu jsou užitečné pro detekci jistých typů abnormálních událostí jako je epilepsie, kdy jednou z nejvýraznějších charakteristik epileptických grafoelementů je přítomnost hrotu (tedy lokálního extrému) v záznamu.

4.2.2 Schéma navržené metody adaptivní segmentace

Na obrázku 4.3 můžete vidět schéma navržené metody adaptivní segmentace. Fáze 2 a 3 probíhají v rámci jedné analýzy pro 3 různě dlouhá okna (256, 512, 1024 vzorků). Fáze 1, 2 a 3 jsou detailněji popsány v kapitolách 4.2.3–4.2.5.



Obrázek 4.3. Schéma navržené metody adaptivní segmentace

4.2.3 Fáze 1: Import dat a předzpracování

Soubory vstupující do první fáze řešení byly RAW data ve formátu *.edf* a k nim připojený soubor vyhodnocení ve formátu *.csv*.

K importu a předzpracování dat ve vstupním formátu *.edf* byl použit volně dostupný FieldTrip toolbox, který je dostupný z webové stránky [43] a je možné jej přidat jako knihovnu externích funkcí do programovacího prostředí MATLAB (pro potřeby této práce byla využita verze MATLAB 2022a). Předzpracování dat v kontinuálním formátu je dobrým krokem pro předcházení jistým typům artefaktů v záznamu, zejména s využitím filtrace. Samotná funkce *ft_preprocessing* z knihovny FieldTrip načítá data ze souboru jako jeden dlouhý souvislý segment kontinuálně naměřených dat bez další filtrace, a proto pro další kroky předzpracování dat bylo nutné doplnit FieldTripu konfigurační strukturu, která obsahovala další příkazy týkající se předzpracování dat. V rámci této diplomové práce bylo předzpracování dat provedeno s využitím funkcí: *demean*, *detrend*, *lpfilter*, *hpfilter*. Funkce *demean* slouží jako příkaz pro korekci základní linie a funkce *detrend* odstraňuje lineární trend z datové sady. Data byla dále filtrována s využitím funkce *hpfilter*, kdy mezní frekvence horní propusti byla nastavena na hodnotu 1 Hz. Posledním krokem předzpracování byla aplikace filtru typu dolní propust využitím funkce *lpfilter*, kdy byla mezní frekvence dolní propusti nastavena na hodnotu 50 Hz. Vzhledem k tomu, že použitá data z TUH (Temple University Hospital) databáze byla naměřena v USA, tedy obsahují síťový šum o frekvenci 60 Hz, nebylo nutné používat další filtr typu pásmová zádrž k odstranění tohoto rušení.

Vyhodnocení EEG záznamu diagnostikem ve vstupním formátu *.csv* bylo do programovacího prostředí MATLAB 2022a importováno pomocí standardní funkce *readtable*, a poté byl datový obsah (časové značky a označení daného úseku) extrahován pro pozdější využití v dalších fázích procesu.

Výstupy z první fáze technického řešení byly uloženy pro další práci do samostatných databází. Jednotlivé databáze obsahovaly: časovou osu pro pořizovaný záznam EEG, předzpracovaná data jednotlivých kanálů EEG, vzorkovací frekvenci, skóre diagnostika.

4.2.4 Fáze 2: Definice etalonu NEP v kanálu

Vstupy do fáze 2 byly časová osa pro pořízený záznam EEG, předzpracovaná data jednotlivých kanálů EEG a vzorkovací frekvence (v případě všech simulovaných dat 250 Hz, v případě reálných dat 400 Hz). Další vstupním parametrem bylo nastavení velikosti okna pro fixní segmentaci dat. S ohledem na pozdější analýzy, zejména na rychlou Fourierovu transformaci, je vhodné volit velikost okna ve formátu 2^N tak, aby funkce rychlé Fourierovy transformace (*fft*) nemusela automaticky doplňovat vzorky datové řady do výše zmíněné velikosti. Z tohoto důvodu byly velikosti okna fixní segmentace nastaveny na hodnoty: 256, 512, 1024 vzorků, což při vzorkovací frekvenci 250 Hz simulovaných dat odpovídá přibližně 1, 2 a 4 sekundám záznamu. Reálná data měla vzorkovací frekvenci 400 Hz, což odpovídá přibližně 0,6 s (okno 256 vzorků), 1,3 s (okno 512 vzorků), 2,6 s (okno 1024 vzorků).

Prvním krokem pro definici NEP v celém kanálu byl zjištěn selektivní práh absolutní amplitudy signálu v daném kanálu jako:

$$\text{Selektivní práh} = \text{mean} + \text{STD} * \text{SMO} , \quad (4.10)$$

kde *mean* je hodnota aritmetického průměru, *SMO* je směrodatná odchylka a *STD* je uživatelské nastavení násobku směrodatné odchylky.

Absolutní hodnota selektivního prahu byla poté použita s kladným i záporným znaménkem k vymezení oblasti, ze které se bude systém učit, jaký parametrický otisk má NEP.

Druhým krokem byla fixní segmentace kanálu. V každém z těchto segmentů bylo zjištěno, zda jeho amplituda alespoň jednou překročila počítanou hodnotu selektivního prahu. Pokud ano, pak byl tento segment vyloučen z dalších analýz pro definici NEP, protože výskyt amplitudy vyšší, než hodnota selektivního prahu je první indikací artefaktové, nebo jiné abnormální činnosti v záznamu, která by negativně ovlivnila přesnost definice etalonu NEP. V každém z fixních segmentů, které vyhověly kritériu selektivního prahu, byly následně vypočítány sledované parametry (statistické, frekvenční, vycházející z tvaru popsané v kapitole 4.2.1) popisující parametrický otisk daného časového úseku záznamu. Výsledky parametrických analýz jednotlivých fixních segmentů byly uloženy do samostatné databáze „Parametry definující NEP“ (viz obrázek 4.3).

Pro každý jednotlivý parametr přijatých segmentů bylo v posledním kroku fáze 2 určeno intervalové pásmo NEP tak, aby byly potlačeny extrémní hodnoty parametrů odpovídající abnormální činnosti v EEG záznamu. Povoleno interval pro každý sledovaný parametr NEP etalonu byl určen jako:

$$NEP\ etalon = mean \pm INT * SMO, \quad (4.11)$$

kde *mean* je hodnota aritmetického průměru, *SMO* je směrodatná odchylka a *INT* je uživatelské nastavení násobku směrodatné odchylky.

Výstupem fáze 2 navrhovaného řešení je databáze povolených intervalů pro parametry NEP etalonu v celém kanálu.

4.2.5 Fáze 3: Detekce NEP

Vstupy třetí fáze jsou: časová osa pořízeného záznamu EEG, předzpracovaná data kanálu EEG, vzorkovací frekvence (získané jako výstup z fáze 1) a databáze intervalů pro parametry etalonu NEP (získané jako výstup z fáze 2). Dále je na vstupu třetí fáze nutné nastavit parametry posuvného okna:

- velikost okna – počet vzorků okna (tato hodnota koresponduje s předchozím nastavením velikosti okna ve fázi dvě, aby byla zachována integrita porovnávání parametrických hodnot)
- posun okna – počet vzorků, o které bude okno posunuto po záznamu při další iteraci výpočtů (posun okna byl pro všechna testovaná nastavení roven 50)

Prvním krokem byla segmentace kanálu EEG posuvným oknem. Časové indexy takto vzniklých segmentů byly uloženy do samostatné databáze pro pozdější využití. V každém z těchto segmentů byly zjištěny hodnoty parametrů (statistické, frekvenční, vycházející z tvaru) popisujících parametrický otisk segmentu a tyto hodnoty spolu s časovými indexy daného segmentu byly přehledně uloženy do databáze „Rekódované segmenty“.

Druhým krokem bylo porovnání parametrického otisku segmentů s databází povolených intervalů NEP etalonu s uživatelským nastavením citlivosti metody (parametr SEN), která může nabývat hodnot od 0 do 1 (podle rovnice 4.12) – hodnota byla zaokrouhlována dolů.

$$Počet\ splněných\ parametrů\ pro\ NEP \doteq SEN * 36, \quad (4.12)$$

kde *SEN* je citlivost metody, která odráží, kolik % parametrů z celkových 36 sledovaných se musí nacházet v odpovídajícím povoleném intervalu NEP etalonu, aby byl segment binárně klasifikován jako NEP.

Výsledky této analýzy každého segmentu byly zapsány do samostatné databáze spojené s časovou osou pořízeného záznamu. Výstupem třetí fáze navrhovaného řešení je tedy databáze „Přítomnost NEP“ obsahující binární klasifikaci pro přítomnost NEP (1), resp. přítomnost AEP (0) v kontextu původní časové osy záznamu.

4.2.6 Sloučení výsledků jednotlivých oken – klasifikace NEP a AEP

Po trojfázové analýze signálu třemi různými okny byla pro každé jednotlivé okno vytvořena binární matice výsledků pro každý jednotlivý vzorek záznamu, kdy 1 značí přítomnost NEP a 0 značí přítomnost AEP. Tyto výsledky byly na závěr sloučeny podle pravidla, kdy klasifikace vzorku jako NEP měla přednost před klasifikací AEP. Tedy všechny vzorky označené jednotlivými okny jako NEP se dostaly do finálního sloučení s klasifikací NEP.

4.3 Statistické zpracování

Pro statistické zpracování výsledků všech testovaných nastavení metody byl pro jednotlivá nastavení proveden výpočet senzitivity, specificity, PPV (prediktivní hodnota pozitivního testu) a NPV (prediktivní hodnota negativního testu) v kontextu jednotlivých vzorků záznamu, přičemž hodnoty senzitivity, specificity, PPV a NPV se pohybují v rozsahu 0–100 %. Senzitivita reflektuje úspěšnost detekce AEP a specificita reflektuje úspěšnost detekce NEP. PPV je hodnota pravděpodobnosti, že vzorky označené metodou jako AEP, byly opravdu pozitivní na přítomnost AEP. NPV je hodnota pravděpodobnosti, že vzorky označené jako NEP, byly opravdu negativní na přítomnost AEP. Rovnice výpočtu senzitivity, specificity, PPV a NPV:

$$\text{Senzitivita} = \frac{TP}{TP+FN} \quad (4.13)$$

$$\text{Specificita} = \frac{TN}{TN+FP} \quad (4.14)$$

$$\text{PPV} = \frac{TP}{TP+FP} \quad (4.15)$$

$$\text{NPV} = \frac{TN}{TN+FN}, \quad (4.16)$$

kde TP (true positive) je počet skutečně pozitivních vzorků záznamu, TN (true negative) je počet skutečně negativních vzorků záznamu, FP je počet falešně negativních vzorků záznamu a FN je počet falešně negativních vzorků záznamu.

Dále bylo pro určení nejlepšího nastavení parametrů nutné sloučit výsledky metody do jedné porovnatelné hodnoty s využitím indexování výsledků [44], proto bylo vypočteno skóre úspěšnosti *Index/Score* jednotlivých nastavení podle rovnice 4.17, které při absolutní úspěšnosti metody může dosáhnout maximální hodnoty 400, podle rovnice:

$$\text{Index/Score} = \text{Senzitivita} + \text{Specificita} + \text{PPV} + \text{NPV} \quad (4.17)$$

Další statistickou metrikou použitou v této práci je přesnost, která byla vypočtena podle rovnice 4.18 [45]:

$$\text{Přesnost} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4.18)$$

Reálná data byla testována v 60 různých kombinacích nastavení parametrů (STD, INT, SEN) a výsledky byly zjištěny pro jednotlivá okna (256, 512, 1024 vzorků) a jejich sloučení, což vedlo k zisku obrovské databáze výsledků. Pro další analýzy těchto výsledků bylo využito prostředí Power BI společnosti Microsoft.

4.4 Použitá data

4.4.1 Simulovaný syntetický signál (SIM_SYNT)

Pro pilotní testování navrhované metodiky byl v rámci práce vytvořen simulovaný syntetický signál v programovacím prostředí MATLAB s využitím modulace sinusových signálů. Pozadí mozkové činnosti bylo simulováno jako složení 5 různých sinusových signálů s níže popsanými hodnotami amplitudy a frekvence (viz tabulka 4.2) v celé časové periodě syntetického signálu. Na tento signál byly v dalším kroku modulovány syntetické artefakty v podobě sinusových signálů v náhodně zvolených časech (viz tabulka 4.3). Simulovaný syntetický signál byl vytvořen složením těchto sinusových signálů a bylo získáno 13 specifických částí signálu (viz tabulka 4.4 a obrázek 4.4) včetně jejich časových značek tak, aby bylo možné kvantitativně validovat úspěšnost metody. Detaily jednotlivých částí a přechodů mezi událostmi najdete na obrázcích 4.5–4.10.

Tabulka 4.2. Přehled simulace pro bezartefaktový simulovaný syntetický signál

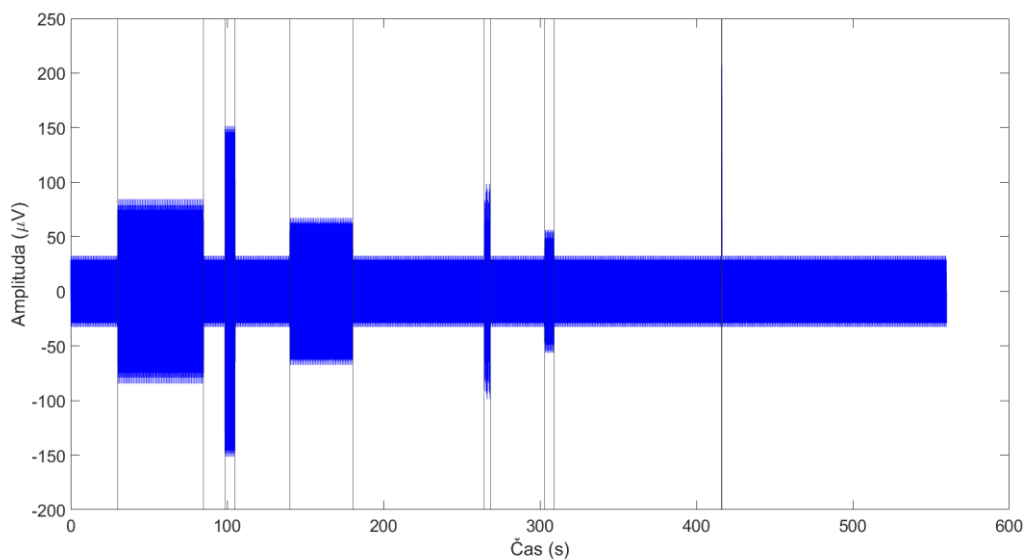
Rytmus	Amplituda (μV)	Frekvence (Hz)
Delta	10	3
Theta	5	7
Alfa	9	16
Beta	15	22
Gama	6	40

Tabulka 4.3. Přehled simulace syntetických artefaktů

Artefakt	Amplituda (μV)	Frekvence (Hz)
Sval	60	45
Oko	135	3
Malý sval	40	48
Malé oko	70	3,5
Síťový šum	30	50
Elektroda	200	1

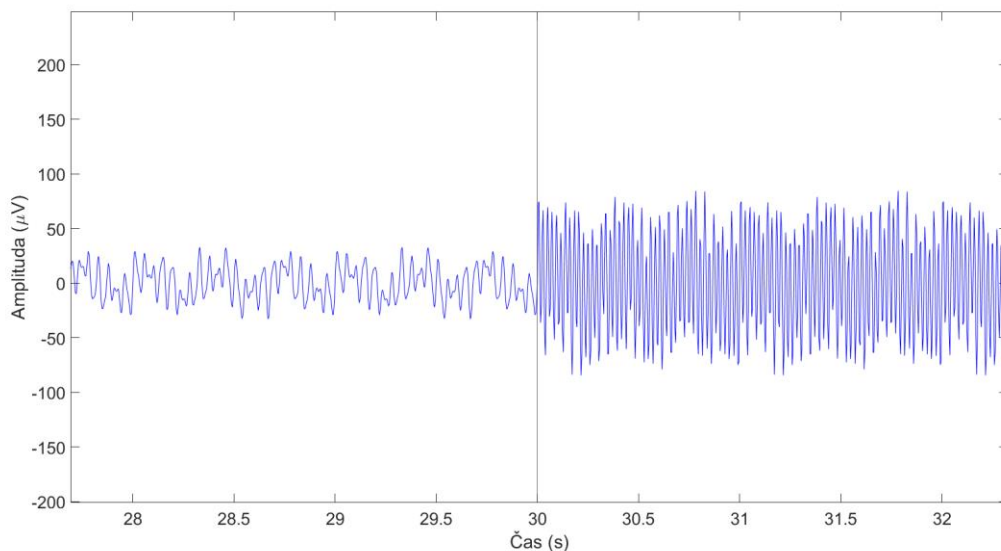
Tabulka 4.4. Přehled částí simulovaného syntetického signálu

Část signálu	Událost	Obrázek
1	Syntetické EEG	
2	Syntetické EEG + syntetický sval	Obrázek 4.5
3	Syntetické EEG	
4	Syntetické EEG + syntetické oko	Obrázek 4.6
5	Syntetické EEG	
6	Syntetické EEG + syntetický malý sval	Obrázek 4.7
7	Syntetické EEG	
8	Syntetické EEG + syntetické malé oko	Obrázek 4.8
9	Syntetické EEG	
10	Syntetické EEG + syntetický síťový šum	Obrázek 4.9
11	Syntetické EEG	
12	Syntetické EEG + syntetická elektroda	Obrázek 4.10
13	Syntetické EEG	

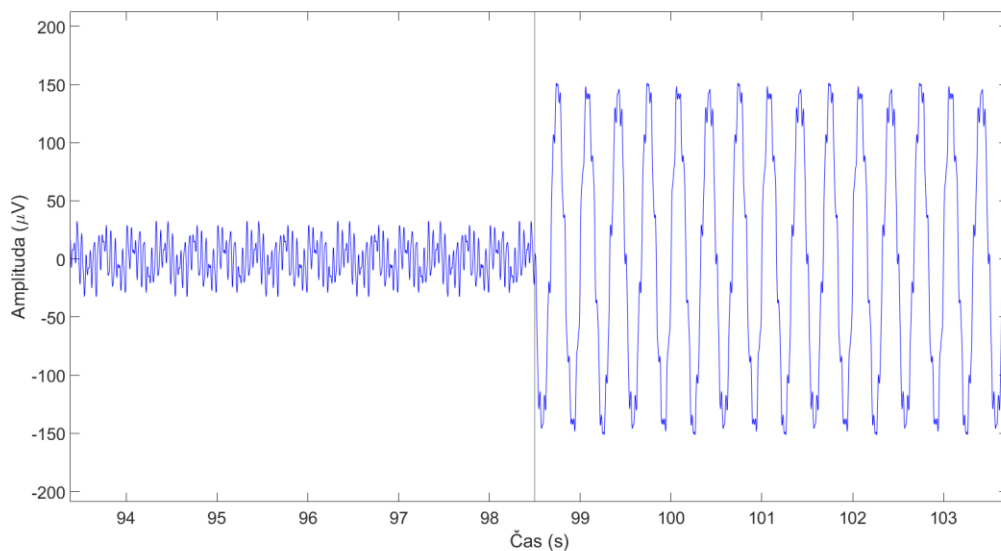


Obrázek 4.4. Ukázka celého SIM_SYNT signálu s vyznačenými hranicemi syntetických artefaktů.

Na obrázku 4.5 níže je vidět v detailu přechod mezi částmi 1 a 2 v simulovaném syntetickém signálu – v části 2 se nachází syntetický svalový artefakt. Na obrázku 4.6 níže je vidět v detailu přechod mezi částmi 2 a 3 v simulovaném syntetickém signálu – v části 3 je vidět syntetický okulární artefakt.

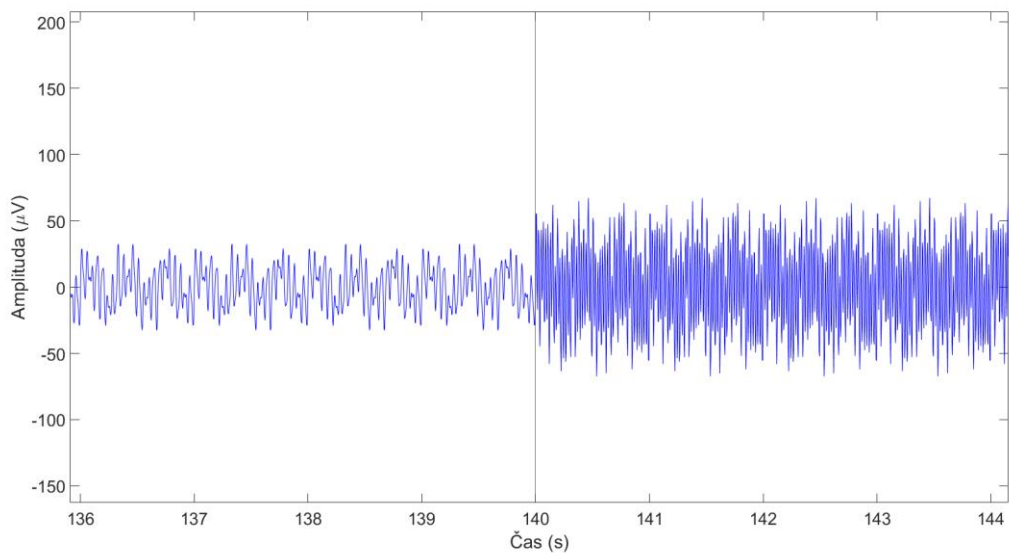


Obrázek 4.5. Syntetický signál. Detail přechodu mezi částmi 1 a 2.

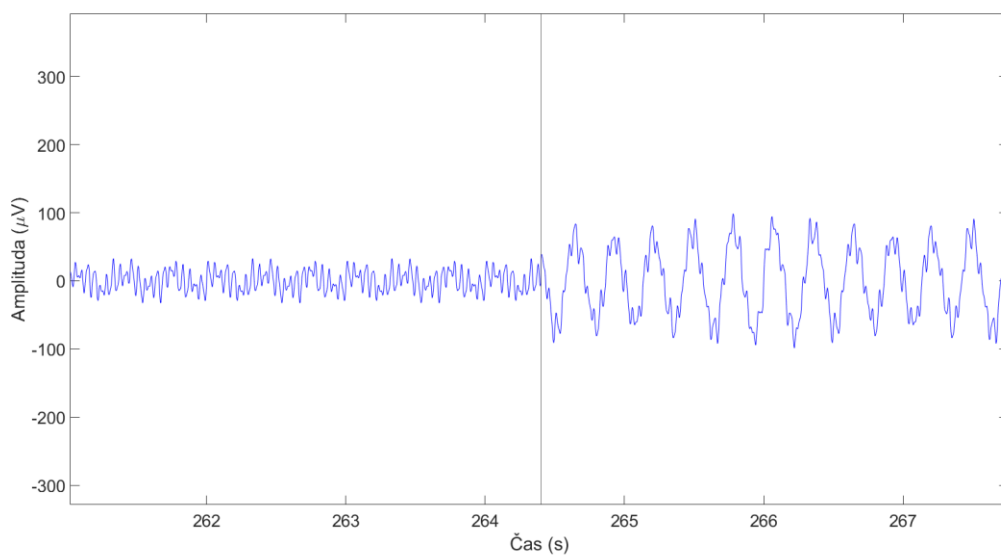


Obrázek 4.6. Syntetický signál. Detail přechodu mezi částmi 3 a 4.

Na obrázku 4.7 níže je vidět v detailu přechod mezi částmi 5 a 6 v simulovaném syntetickém signálu – v části 6 se nachází syntetický svalový artefakt s nízkou amplitudou. Na obrázku 4.8 níže je vidět v detailu přechod mezi částmi 7 a 8 v simulovaném syntetickém signálu – v části 8 je vidět syntetický okulární artefakt s nízkou amplitudou.

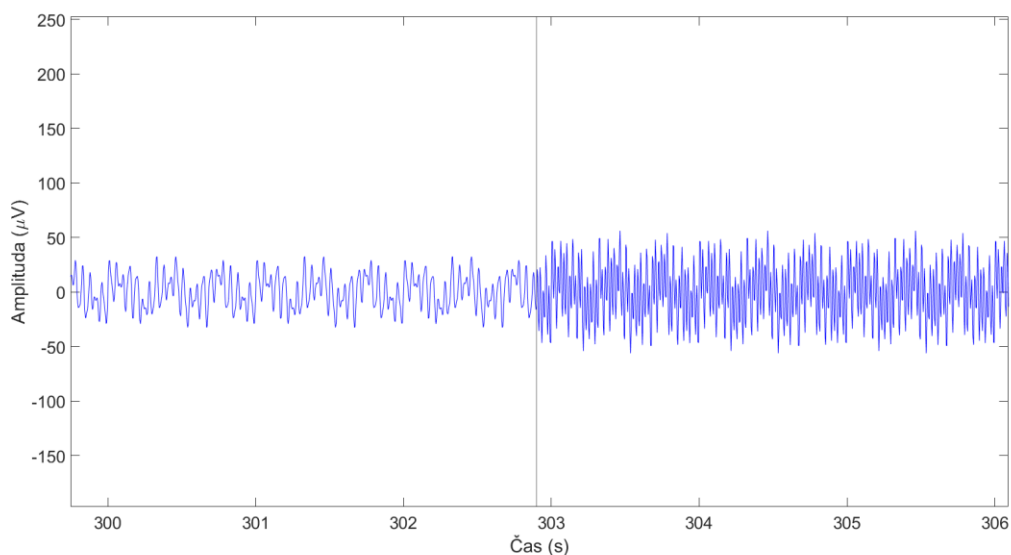


Obrázek 4.7. Syntetický signál. Detail přechodu mezi částmi 5 a 6.

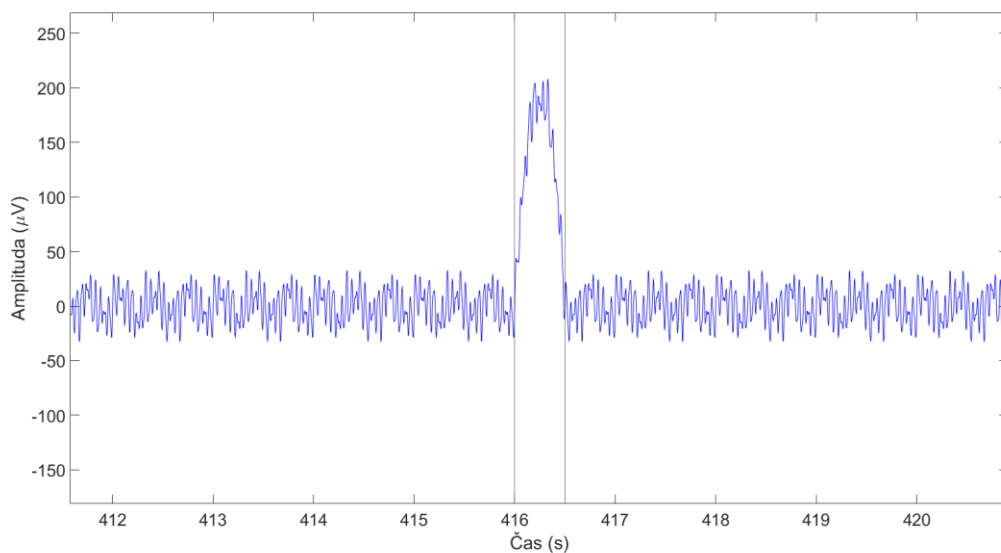


Obrázek 4.8. Syntetický signál. Detail přechodu mezi částmi 7 a 8.

Na obrázku 4.9 níže je vidět v detailu přechod mezi částmi 9 a 10 v simulovaném syntetickém signálu – v části 10 se nachází syntetický síťový šum. Na obrázku 4.10 níže je vidět v detailu přechod mezi částmi 11, 12 a 13 v simulovaném syntetickém signálu – v části 12 je vidět syntetický elektrodový artefakt.



Obrázek 4.9. Syntetický signál. Detail přechodu mezi částmi 9 a 10.



Obrázek 4.10. Syntetický signál. Detail přechodu mezi částmi 11, 12 a 13.

4.4.2 Simulovaný semisyntetický signál (SIM_SEMISYNT)

Při testování správnosti odstranění artefaktů z EEG signálu, např. pomocí odstranění komponent ICA, se používají tzv. semisyntetická data. Tato data jsou složena z EEG signálu, který je naprosto čistý (tzn. bez artefaktů a jiných abnormálních událostí) a na něj je modulován signál artefaktu. [46]

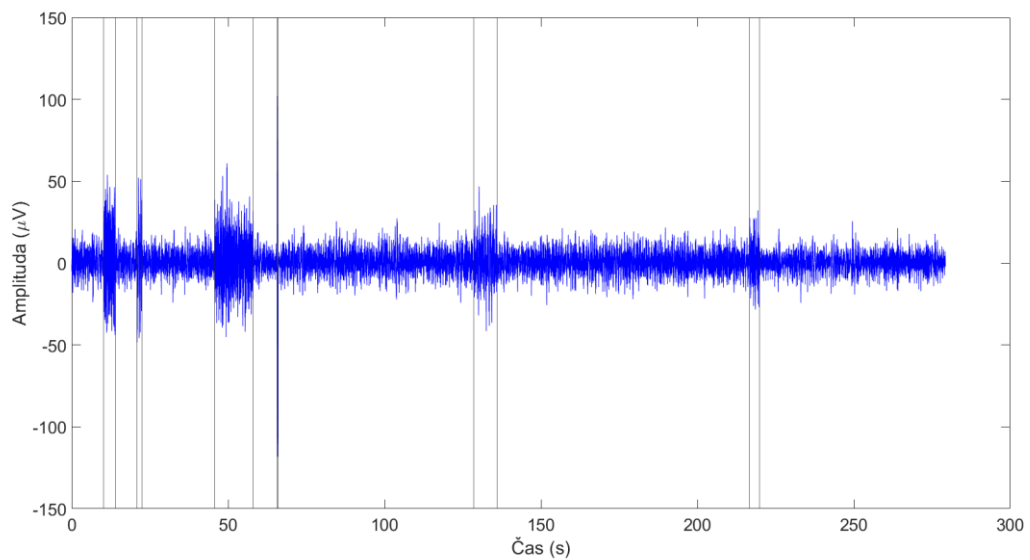
Tento přístup mě inspiroval k vytvoření vlastního simulovaného semisyntetického signálu pro testování mnou navržené metodiky. Klidový reálný EEG signál byl získán ručním procházením a vystřiháním odpovídajících úseků klidového EEG z reálného záznamu jednoho signálu z databáze TUH. Na tento signál byly poté modulovány syntetické artefakty (viz tabulka 4.5), aby bylo možné otestovat přesnost jejich detekce v kontextu přesných časových značek začátku a konce artefaktů. Přehled vytvořených částí simulovaného semisyntetického signálu najdete v tabulce 4.6 a celý vytvořený signál najdete na obrázku 4.11, detail jednotlivých částí semisyntetického signálu je na obrázcích 4.12–4.17.

Tabulka 4.5. Přehled simulovaných syntetických artefaktů

Artefakt	Amplituda (μV)	Frekvence (Hz)
Sval	20	45
Oko	30	3
Síťový šum	10	50
Elektroda	120	2
Malé oko	7	1,5
Malý sval	1	44.5

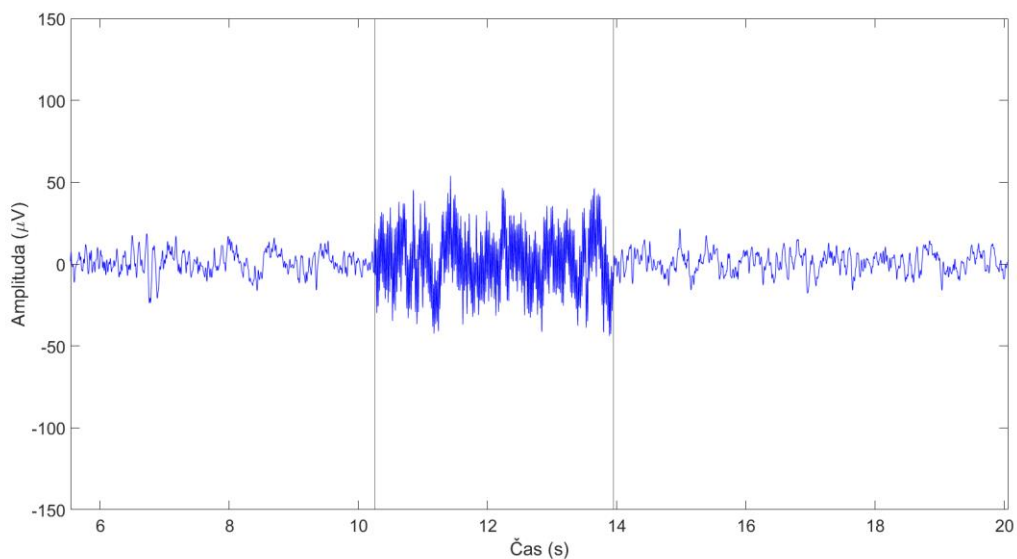
Tabulka 4.6. Přehled částí simulovaného semisyntetického signálu

Část signálu	Událost	Obrázek
1	Reálné EEG	
2	Reálné EEG + syntetický sval	Obrázek 4.12
3	Reálné EEG	
4	Reálné EEG + syntetické oko	Obrázek 4.13
5	Reálné EEG	
6	Reálné EEG + syntetický síťový šum	Obrázek 4.14
7	Reálné EEG	
8	Reálné EEG + syntetická elektroda	Obrázek 4.15
9	Reálné EEG	
10	Reálné EEG + syntetické malé oko	Obrázek 4.16
11	Reálné EEG	
12	Reálné EEG + syntetický malý sval	Obrázek 4.17
13	Reálné EEG	

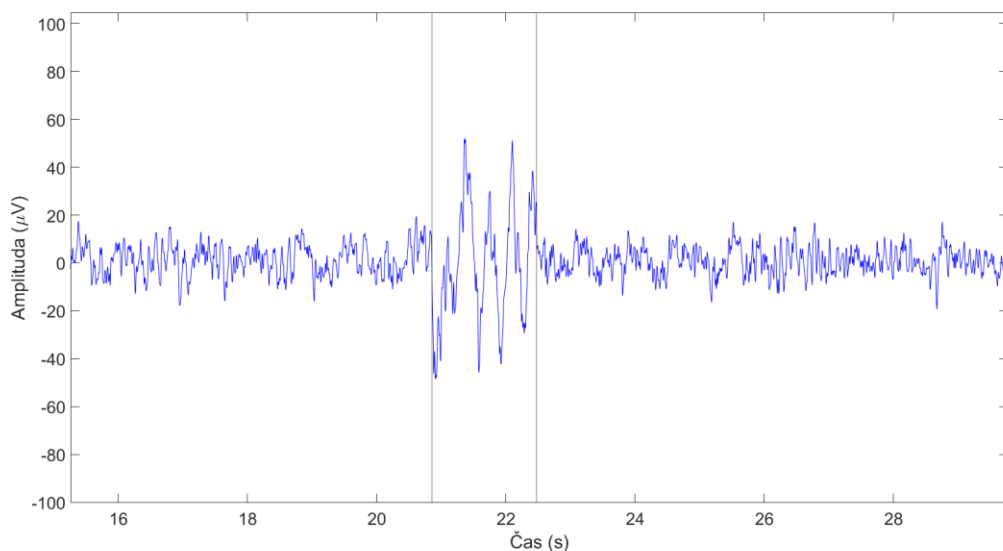


Obrázek 4.11. Semisyntetický signál s vyznačenými hranicemi syntetických artefaktů.

Na obrázku 4.12 níže je vidět v detailu přechod mezi částmi 1, 2 a 3 v simulovaném semisyntetickém signálu – v části 2 se nachází syntetický svalový artefakt. Na obrázku 4.13 níže je vidět v detailu přechod mezi částmi 3, 4 a 5 v simulovaném semisyntetickém signálu – v části 4 je vidět syntetický okulární artefakt.

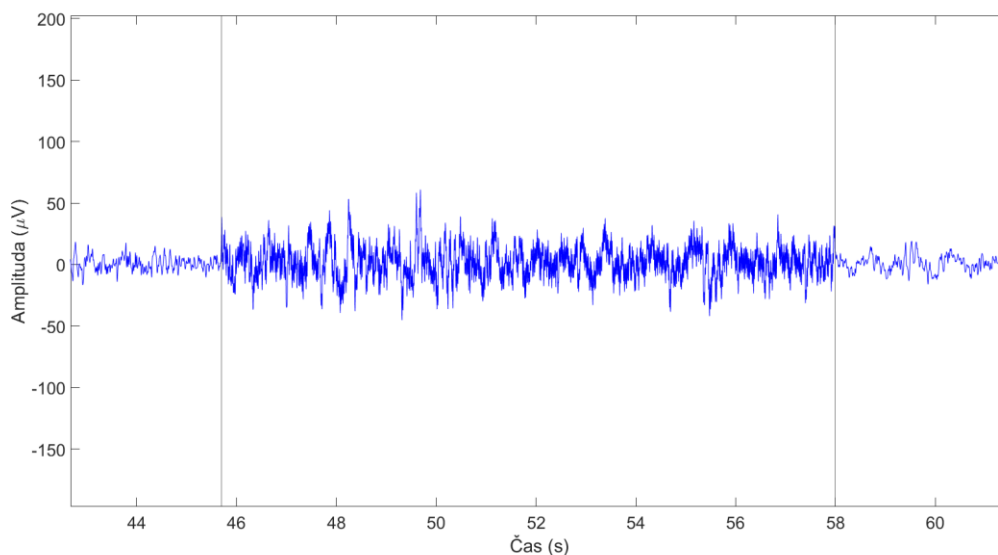


Obrázek 4.12. Semisyntetický signál. Detail přechodu mezi částmi 1, 2 a 3. V části 2 je syntetický svalový artefakt.

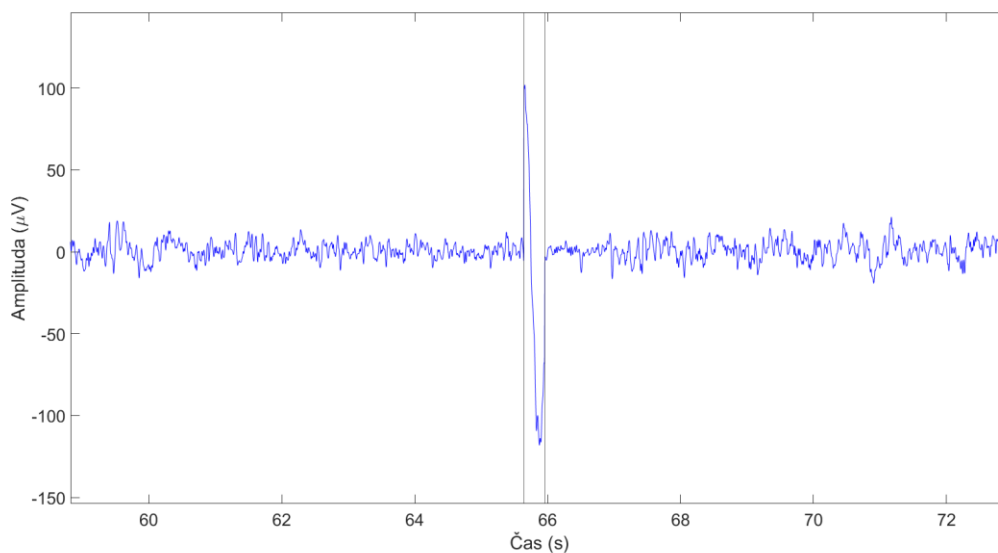


Obrázek 4.13. Semisyntetický signál. Detail přechodu mezi částmi 3, 4 a 5. V části 4 je syntetický okulární artefakt.

Na obrázku 4.14 níže je vidět v detailu přechod mezi částmi 5, 6 a 7 v simulovaném semisyntetickém signálu – v části 6 se nachází syntetický síťový šum. Na obrázku 4.15 níže je vidět v detailu přechod mezi částmi 7, 8 a 9 v simulovaném semisyntetickém signálu – v části 8 je vidět syntetický elektrodový artefakt.

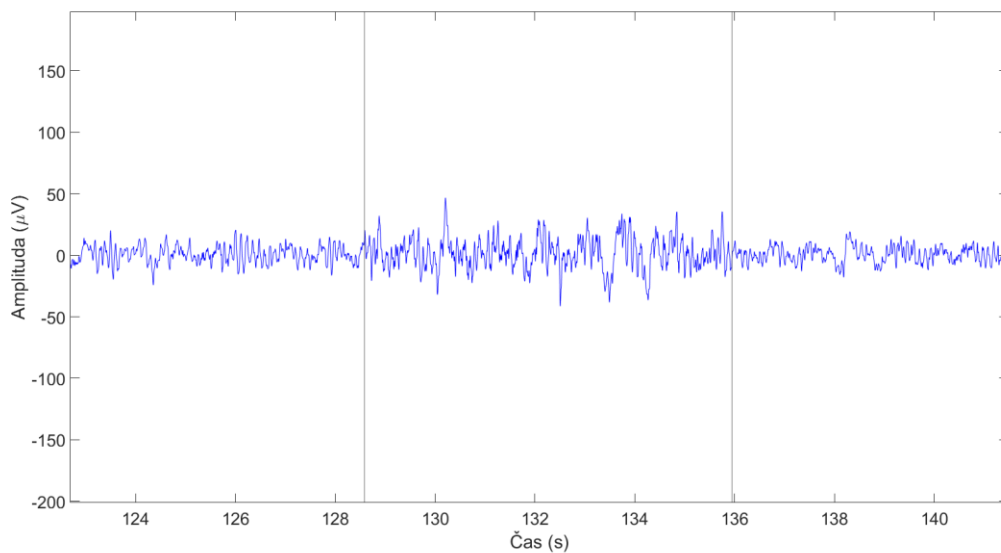


Obrázek 4.14. Semisyntetický signál. Detail přechodu mezi částmi 5, 6 a 7. V části 6 je syntetický síťový šum.

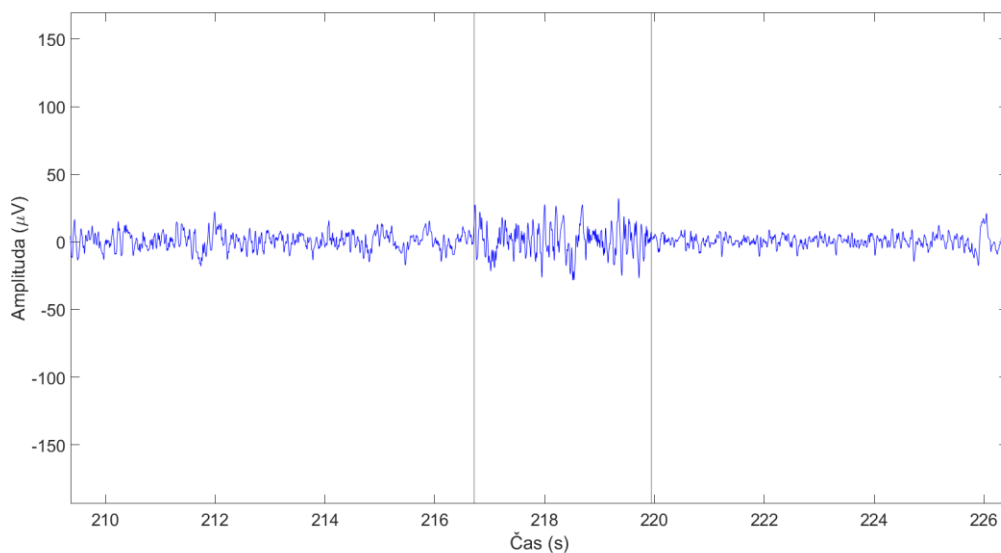


Obrázek 4.15. Semisyntetický signál. Detail přechodu mezi částmi 7, 8 a 9. V části 8 je syntetický elektrodový artefakt.

Na obrázku 4.16 níže je vidět v detailu přechod mezi částmi 9, 10 a 11 v simulovaném semisyntetickém signálu – v části 10 se nachází syntetický okulární artefakt s nízkou amplitudou. Na obrázku 4.17 níže je vidět v detailu přechod mezi částmi 11, 12 a 13 v simulovaném semisyntetickém signálu – v části 12 je vidět syntetický svalový artefakt s nízkou amplitudou.



Obrázek 4.16. Semisyntetický signál. Detail přechodu mezi částmi 9, 10 a 11. V části 10 je syntetický okulární artefakt s nízkou amplitudou.



Obrázek 4.17. Semisyntetický signál. Detail přechodu mezi částmi 11, 12 a 13. V části 12 je syntetický svalový artefakt s nízkou amplitudou.

4.4.3 Temple University Hospital Data Corpus (TUH)

Klinická data TUH databáze [8] byla shromážděna s informovaným souhlasem ze záznamů nemocnice v souladu s Helsinskou deklarácí [47]. Všechny záznamy byly poté důkladně anonymizovány a veškeré potenciální identifikátory pacientů (např. jméno, rodné číslo atd.) byly ze záznamů odstraněny a nahrazeny randomizovanými identifikátory záznamů.

Celý korpus TUH obsahuje cca 30 000 záznamů ve formátu *.edf* spolu s anotací diagnostického týmu univerzity ve formátu *.csv*. Vzorkovací frekvence záznamů korpusu se pohybuje od 250 do 512 Hz, přičemž valná většina (87 %) byla vzorkována frekvencí 250 Hz a spolu se záznamem EEG byly měřeny i některé další biologické signály (např. EKG, EOG). Celý tento korpus je uživatelům plně dostupný pro výzkumnou činnost po registraci, ovšem s podmínkou, kdy uživatel nesmí tyto záznamy poskytovat dalším subjektům bez registrace, proto RAW data použitá v této diplomové práci nemohou být přiložena v elektronických přílohách diplomové práce.

Temple University Artifact Corpus (TUAR)

Data použitá v této práci pochází z korpusu TUAR [48], který je částí celé široké TUH databáze EEG souborů. TUAR korpus se specializuje na identifikaci artefaktů (pohyb očí, žvýkání, chvění, elektrodové artefakty, svalové artefakty) a tato databáze může sloužit k testování SW pro identifikaci artefaktů, nebo jako trénovací množiny pro učící se systémy. Jednotlivá data v tomto korpusu jsou rozdělena do podsložek podle použitého zapojení elektrod:

- Průměrná referenční elektroda
- Referenční elektroda spojených uší

Proces anotace souborů

Anotaci souborů provádí tým diagnostiků, kteří se věnují speciálně interpretaci EEG. Jedná se o studenty Temple University a jejich školitele, kteří prošli několika měsíci školení. Během prvního kola anotace jeden člen týmu kontroluje záchvatové události, popř. další události, jako jsou artefakty. V případě, že není schopen definitivně rozhodnout o dané události, posouvá soubor ke kontrole dalšímu v hierarchii profesně staršímu členu týmu s vyšší odborností. V případě, že ani ten není schopen rozhodnout o daném úseku záznamu, je tato část signálu projednána a přezkoumána během týdenního zasedání, kdy je rozhodnuto na základě konsensu týmu. V posledním kole před umístěním dat do veřejného korpusu se provádí ještě revize anotovaných záznamů tak, aby byla zajištěna co nejvyšší přesnost anotace. [17; 19]

4.4.4 Simulovaná reálná data (SIM_REAL1 a SIM_REAL2)

Simulovaná reálná data byla vytvořena manuální kontrolou, selekcí a následnou segmentací klidových úseků EEG z reálného záznamu jednoho kanálu vzorkovaného 250 Hz z databáze TUAR. Dále byly tímto způsobem získány ze stejného kanálu záznamu i artefaktové úseky EEG záznamu. Tento postup byl zvolen proto, aby bylo jednoznačně známo, kdy artefakt začíná a končí v čase. Takto získané úseky z jednoho kanálu byly spojeny do jedné časové řady, aby byl přesně znám časový index vzorku, kdy abnormální událost začala a skončila. Kvůli maximální možné eliminaci bodů nespojitosti při spojování dat byly hraniční vzorky pro spojení úseků signálu vybrány v rozmezí amplitudy $-0,2-0,2 \mu\text{V}$.

Výše zmíněným způsobem byly vytvořeny dva simulované reálné signály (SIM_REAL1 a SIM_REAL2) od dvou různých subjektů z databáze TUAR, které byly před tím samozřejmě předzpracovány postupem pospaným v kapitole 4.2.3 a následně testovány navrženou metodou v programovém prostředí MATLAB.

4.4.5 Anotovaná reálná data (REAL_ANOT)

Anotovaná reálná data použitá v této práci získána v databázi TUH byla vzorkována frekvencí 400 Hz a jsou částí menšího korpusu TUAR (Temple University Artifact Corpus). V této práci byl testován jeden kanál (T3) datového souboru naměřeného v režimu zprůměrované reference s identifikátorem *aaaaadao_s001_t004*, ve kterém se vyskytovaly anotované artefakty z chvění pacienta, svalové artefakty a elektrodový artefakt.

5 Výsledky

5.1 Výsledky simulovaných dat

V kapitole 5.1.1 naleznete popis testovaných kombinací nastavení na simulovaných datech. V navazujících kapitolách jsou výsledky těchto nastavení pro simulovaná syntetická data (kapitola 5.1.2), simulovaná semisyntetická data (kapitola 5.1.3), simulovaná reálná data 1 (kapitola 5.1.4), simulovaná reálná data 2 (kapitola 5.1.5). Výsledky v jednotlivých kapitolách jsou prezentovány pro jednotlivá použitá okna a jejich sloučení.

Grafické ukázky výsledků nejlepší kandidátů metody na všech simulovaných datech jsou v kapitole 5.1.6.

5.1.1 Popis nastavení parametrů metody pro simulovaná data

V následujících kapitolách jsou popsána nastavení, která byla použita k testování navržené metody na simulovaných datech.

Nastavení 1

V nastavení 1 parametrů metody byly nastaveny výpočty pro selektivní práh absolutní amplitudy a interval NEP etalonu podle rovnic:

$$\text{Selektivní práh} = \text{mean} + 2 * SMO \quad (5.1)$$

$$\text{NEP etalon} = \text{mean} \pm 2 * SMO, \quad (5.2)$$

kde *mean* je hodnota aritmetického průměru a *SMO* je hodnota směrodatné odchylky.

Metoda byla testována pro 4 různé citlivosti metody, které nabývaly hodnot 70, 80, 90 a 100 % (viz rovnice 4.12).

Nastavení 2

V nastavení parametrů metody 2 byly nastaveny výpočty pro selektivní práh absolutní amplitudy, interval NEP etalonu podle rovnic:

$$\text{Selektivní práh} = \text{mean} + 3 * SMO \quad (5.3)$$

$$\text{NEP etalon} = \text{mean} \pm 2 * SMO, \quad (5.4)$$

kde *mean* je hodnota aritmetického průměru a *SMO* je hodnota směrodatné odchylky.

Metoda byla testována pro 4 různé citlivosti metody, které nabývaly hodnot 70, 80, 90 a 100 % (viz rovnice 4.12).

5.1.2 Výsledky SIM_SYNT

Celkem bylo testováno 8 kombinací nastavení parametrů metody STD, INT, SEN, přičemž výsledky byly získány pro analýzy signálu jednotlivými okny (256, 512, 1024 vzorků) a pro jejich sloučení – celkem bylo získáno 32 unikátních výsledků. V tabulce 5.1 jsou výsledky testování metody na simulovaných syntetických datech. Učení NEP (%) označuje, z kolika % záznamu systém určoval etalon NEP. NEP (%) označuje kolik % záznamu bylo nakonec označeno jako NEP. Výsledky metrik Index/Score a Přesnost vyšší než 90 % jsou označeny zelenou šipkou, výsledky v rozmezí 80-90 % žlutou šipkou, ostatní červenou.

Úspěšnosti přes 90 % v obou metrikách (Index/Score, Přesnost) dosáhlo 25 z 32 testovaných nastavení (78 % testovaných nastavení).

Tabulka 5.1. Výsledky metody na SIM_SYNT pro testovaná nastavení.

Okno	STD	INT	SEN	TP	TN	FP	FN	Senz.	Spec.	PPV	NPV	Index Score (%)	Přesnost (%)	Učení NEP (%)	NEP (%)
OKNO256	STD20	INT20	SEN70	28091	110766	1156	0	100	99	96	100	↑ 98,8	↑ 99,2	79	79
OKNO512	STD20	INT20	SEN70	28090	108919	3003	1	100	97	90	100	↑ 96,9	↑ 97,9	78	78
OKNO1024	STD20	INT20	SEN70	28091	104568	7354	0	100	93	79	100	↑ 93,2	↑ 94,7	76	75
SLOUČENÁ	STD20	INT20	SEN70	28090	111115	807	1	100	99	97	100	↑ 99,1	↑ 99,4	NaN	79
OKNO256	STD20	INT20	SEN80	28091	110610	1312	0	100	99	96	100	↑ 98,6	↑ 99,1	79	79
OKNO512	STD20	INT20	SEN80	28090	108919	3003	1	100	97	90	100	↑ 96,9	↑ 97,9	78	78
OKNO1024	STD20	INT20	SEN80	28091	104568	7354	0	100	93	79	100	↑ 93,2	↑ 94,7	76	75
SLOUČENÁ	STD20	INT20	SEN80	28090	111115	807	1	100	99	97	100	↑ 99,1	↑ 99,4	NaN	79
OKNO256	STD20	INT20	SEN90	28091	109206	2716	0	100	98	91	100	↑ 97,2	↑ 98,1	79	78
OKNO512	STD20	INT20	SEN90	28090	106809	5113	1	100	95	85	100	↑ 95,0	↑ 96,3	78	76
OKNO1024	STD20	INT20	SEN90	28091	99898	12024	0	100	89	70	100	⇒ 89,8	↑ 91,4	76	71
SLOUČENÁ	STD20	INT20	SEN90	28090	111115	807	1	100	99	97	100	↑ 99,1	↑ 99,4	NaN	79
OKNO256	STD20	INT20	SEN100	28091	65184	46738	0	100	58	38	100	↓ 73,9	↓ 66,6	79	47
OKNO512	STD20	INT20	SEN100	28091	49712	62210	0	100	44	31	100	↓ 68,9	↓ 55,6	78	36
OKNO1024	STD20	INT20	SEN100	28091	38212	73710	0	100	34	28	100	↓ 65,4	↓ 47,4	76	27
SLOUČENÁ	STD20	INT20	SEN100	28091	87756	24166	0	100	78	54	100	⇒ 83,0	⇒ 82,7	NaN	63
OKNO256	STD30	INT20	SEN70	26716	110877	1045	1375	95	99	96	99	↑ 97,3	↑ 98,3	87	80
OKNO512	STD30	INT20	SEN70	27783	110410	1512	308	99	99	95	100	↑ 98,0	↑ 98,7	87	79
OKNO1024	STD30	INT20	SEN70	26638	106961	4961	1453	95	96	84	99	↑ 93,3	↑ 95,4	86	77
SLOUČENÁ	STD30	INT20	SEN70	25633	111494	428	2458	91	100	98	98	↑ 96,8	↑ 97,9	NaN	81
OKNO256	STD30	INT20	SEN80	28091	110766	1156	0	100	99	96	100	↑ 98,8	↑ 99,2	87	79
OKNO512	STD30	INT20	SEN80	28073	109364	2558	18	100	98	92	100	↑ 97,3	↑ 98,2	87	78
OKNO1024	STD30	INT20	SEN80	28091	104568	7354	0	100	93	79	100	↑ 93,2	↑ 94,7	86	75
SLOUČENÁ	STD30	INT20	SEN80	28073	111244	678	18	100	99	98	100	↑ 99,2	↑ 99,5	NaN	79
OKNO256	STD30	INT20	SEN90	28091	110766	1156	0	100	99	96	100	↑ 98,8	↑ 99,2	87	79
OKNO512	STD30	INT20	SEN90	28090	108919	3003	1	100	97	90	100	↑ 96,9	↑ 97,9	87	78
OKNO1024	STD30	INT20	SEN90	28091	104568	7354	0	100	93	79	100	↑ 93,2	↑ 94,7	86	75
SLOUČENÁ	STD30	INT20	SEN90	28090	111115	807	1	100	99	97	100	↑ 99,1	↑ 99,4	NaN	79
OKNO256	STD30	INT20	SEN100	28091	103390	8532	0	100	92	77	100	↑ 92,3	↑ 93,9	87	74
OKNO512	STD30	INT20	SEN100	28090	87145	24777	1	100	78	53	100	⇒ 82,7	⇒ 82,3	87	62
OKNO1024	STD30	INT20	SEN100	28091	87736	24186	0	100	78	54	100	⇒ 83,0	⇒ 82,7	86	63
SLOUČENÁ	STD30	INT20	SEN100	28090	110671	1251	1	100	99	96	100	↑ 98,7	↑ 99,1	NaN	79

5.1.3 Výsledky SIM_SEMISYNT

Celkem bylo testováno 8 kombinací nastavení parametrů metody STD, INT, SEN, přičemž výsledky byly získány pro analýzy signálu jednotlivými okny (256, 512, 1024 vzorků) a pro jejich sloučení – celkem bylo získáno 32 unikátních výsledků. V tabulce 5.2 jsou výsledky testování metody na simulovaných semisyntetických datech. Učení NEP (%) označuje, z kolika % záznamu systém určoval etalon NEP. NEP (%) označuje kolik % záznamu bylo nakonec označeno jako NEP. Výsledky metrik Index/Score a Přesnost vyšší než 90 % jsou označeny zelenou šipkou, výsledky v rozmezí 80-90 % žlutou šipkou, ostatní červenou.

Úspěšnosti přes 90 % v obou metrikách (Index/Score, Přesnost) dosáhlo 8 z 32 testovaných nastavení (25 % testovaných nastavení).

Tabulka 5.2. Výsledky metody na SIM_SEMISYNT pro testovaná nastavení.

Okno	STD	INT	SEN	TP	TN	FP	FN	Senz.	Spec.	PPV	NPV	Index Score (%)	Přesnost (%)	Učení NEP (%)	NEP (%)
OKNO256	STD20	INT20	SEN70	7132	60340	2353	0	100	96	75	100	↑ 92,9	↑ 96,6	53	86
OKNO512	STD20	INT20	SEN70	7123	58241	4452	9	100	93	62	100	⇒ 88,6	↑ 93,6	32	83
OKNO1024	STD20	INT20	SEN70	7132	30794	31899	0	100	49	18	100	↓ 66,8	↓ 54,3	10	44
SLOUČENÁ	STD20	INT20	SEN70	7123	61159	1534	9	100	98	82	100	↑ 94,9	↑ 97,8	NaN	88
OKNO256	STD20	INT20	SEN80	7132	57182	5511	0	100	91	56	100	⇒ 86,9	↑ 92,1	53	82
OKNO512	STD20	INT20	SEN80	7123	51711	10982	9	100	82	39	100	⇒ 80,4	⇒ 84,3	32	74
OKNO1024	STD20	INT20	SEN80	7132	18956	43737	0	100	30	14	100	↓ 61,1	↓ 37,4	10	27
SLOUČENÁ	STD20	INT20	SEN80	7123	58259	4434	9	100	93	62	100	⇒ 88,6	↑ 93,6	NaN	83
OKNO256	STD20	INT20	SEN90	7132	48338	14355	0	100	77	33	100	↓ 77,6	↓ 79,4	53	69
OKNO512	STD20	INT20	SEN90	7123	35779	26914	9	100	57	21	100	↓ 69,5	↓ 61,4	32	51
OKNO1024	STD20	INT20	SEN90	7132	9116	53577	0	100	15	12	100	↓ 56,6	↓ 23,3	10	13
SLOUČENÁ	STD20	INT20	SEN90	7123	50751	11942	9	100	81	37	100	↓ 79,5	⇒ 82,9	NaN	73
OKNO256	STD20	INT20	SEN100	7132	16358	46335	0	100	26	13	100	↓ 59,9	↓ 33,6	53	23
OKNO512	STD20	INT20	SEN100	7132	11014	51679	0	100	18	12	100	↓ 57,4	↓ 26,0	32	16
OKNO1024	STD20	INT20	SEN100	7132	1024	61669	0	100	2	10	100	↓ 53,0	↓ 11,7	10	1
SLOUČENÁ	STD20	INT20	SEN100	7132	22042	40651	0	100	35	15	100	↓ 62,5	↓ 41,8	NaN	32
OKNO256	STD30	INT20	SEN70	7132	61376	1317	0	100	98	84	100	↑ 95,6	↑ 98,1	84	88
OKNO512	STD30	INT20	SEN70	7123	60401	2292	9	100	96	76	100	↑ 93,0	↑ 96,7	80	87
OKNO1024	STD30	INT20	SEN70	7087	57771	4922	45	99	92	59	100	⇒ 87,6	↑ 92,9	66	83
SLOUČENÁ	STD30	INT20	SEN70	7078	62074	619	54	99	99	92	100	↑ 97,5	↑ 99,0	NaN	89
OKNO256	STD30	INT20	SEN80	7132	60340	2353	0	100	96	75	100	↑ 92,9	↑ 96,6	84	86
OKNO512	STD30	INT20	SEN80	7123	59115	3578	9	100	94	67	100	↑ 90,2	↑ 94,9	80	85
OKNO1024	STD30	INT20	SEN80	7132	56842	5851	0	100	91	55	100	⇒ 86,4	↑ 91,6	66	81
SLOUČENÁ	STD30	INT20	SEN80	7123	61547	1146	9	100	98	86	100	↑ 96,0	↑ 98,3	NaN	88
OKNO256	STD30	INT20	SEN90	7132	56764	5929	0	100	91	55	100	⇒ 86,3	↑ 91,5	84	81
OKNO512	STD30	INT20	SEN90	7132	53006	9687	0	100	85	42	100	⇒ 81,7	⇒ 86,1	80	76
OKNO1024	STD30	INT20	SEN90	7132	50174	12519	0	100	80	36	100	↓ 79,1	⇒ 82,1	66	72
SLOUČENÁ	STD30	INT20	SEN90	7132	59552	3141	0	100	95	69	100	↑ 91,1	↑ 95,5	NaN	85
OKNO256	STD30	INT20	SEN100	7132	26740	35953	0	100	43	17	100	↓ 64,8	↓ 48,5	84	38
OKNO512	STD30	INT20	SEN100	7132	28758	33935	0	100	46	17	100	↓ 65,8	↓ 51,4	80	41
OKNO1024	STD30	INT20	SEN100	7132	23102	39591	0	100	37	15	100	↓ 63,0	↓ 43,3	66	33
SLOUČENÁ	STD30	INT20	SEN100	7132	44452	18241	0	100	71	28	100	↓ 74,8	↓ 73,9	NaN	64

5.1.4 Výsledky SIM_REAL1

Celkem bylo testováno 8 kombinací nastavení parametrů metody STD, INT, SEN, přičemž výsledky byly získány pro analýzy signálu jednotlivými okny (256, 512, 1024 vzorků) a pro jejich sloučení – celkem bylo získáno 32 unikátních výsledků. V tabulce 5.3 jsou výsledky testování metody na simulovaných reálných datech 1. Učení NEP (%) označuje, z kolika % záznamu systém určoval etalon NEP. NEP (%) označuje kolik % záznamu bylo nakonec označeno jako NEP. Výsledky metrik Index/Score a Přesnost vyšší než 90 % jsou označeny zelenou šipkou, výsledky v rozmezí 80-90 % žlutou šipkou, ostatní červenou.

Úspěšnosti přes 90 % v obou metrikách (Index/Score, Přesnost) dosáhlo 12 z 32 testovaných nastavení (37,5 % testovaných nastavení).

Tabulka 5.3. Výsledky metody na SIM_REAL1 pro testovaná nastavení.

Okno	STD	INT	SEN	TP	TN	FP	FN	Senz.	Spec.	PPV	NPV	Index Score (%)	Přesnost (%)	Učení NEP (%)	NEP (%)
OKNO256	STD20	INT20	SEN70	28052	277998	6363	1228	96	98	82	100	↑ 93,7	↑ 97,6	77	89
OKNO512	STD20	INT20	SEN70	28277	274163	10198	1003	97	96	73	100	↑ 91,5	↑ 96,4	66	88
OKNO1024	STD20	INT20	SEN70	28839	260667	23694	441	98	92	55	100	⇒ 86,2	↑ 92,3	50	83
SLOUČENÁ	STD20	INT20	SEN70	27954	281748	2613	1326	95	99	91	100	↑ 96,4	↑ 98,7	NaN	90
OKNO256	STD20	INT20	SEN80	28062	267274	17087	1218	96	94	62	100	⇒ 87,9	↑ 94,2	77	86
OKNO512	STD20	INT20	SEN80	28277	258993	25368	1003	97	91	53	100	⇒ 85,0	↑ 91,6	66	83
OKNO1024	STD20	INT20	SEN80	28839	233121	51240	441	98	82	36	100	↓ 79,1	⇒ 83,5	50	74
SLOUČENÁ	STD20	INT20	SEN80	27964	276678	7683	1316	96	97	78	100	↑ 92,7	↑ 97,1	NaN	89
OKNO256	STD20	INT20	SEN90	28288	233734	50627	992	97	82	36	100	↓ 78,6	⇒ 83,5	77	75
OKNO512	STD20	INT20	SEN90	28689	207651	76710	591	98	73	27	100	↓ 74,5	↓ 75,4	66	66
OKNO1024	STD20	INT20	SEN90	28839	173333	111028	441	98	61	21	100	↓ 70,0	↓ 64,5	50	55
SLOUČENÁ	STD20	INT20	SEN90	28036	251484	32877	1244	96	88	46	100	⇒ 82,4	⇒ 89,1	NaN	81
OKNO256	STD20	INT20	SEN100	29024	124740	159621	256	99	44	15	100	↓ 64,5	↓ 49,0	77	40
OKNO512	STD20	INT20	SEN100	29280	98612	185749	0	100	35	14	100	↓ 62,1	↓ 40,8	66	31
OKNO1024	STD20	INT20	SEN100	29280	83392	200969	0	100	29	13	100	↓ 60,5	↓ 35,9	50	27
SLOUČENÁ	STD20	INT20	SEN100	29024	174060	110301	256	99	61	21	100	↓ 70,3	↓ 64,8	NaN	56
OKNO256	STD30	INT20	SEN70	28044	282664	1697	1236	96	99	94	100	↑ 97,3	↑ 99,1	90	91
OKNO512	STD30	INT20	SEN70	28037	281127	3234	1243	96	99	90	100	↑ 96,0	↑ 98,6	87	90
OKNO1024	STD30	INT20	SEN70	28831	273795	10566	449	98	96	73	100	↑ 91,9	↑ 96,5	81	87
SLOUČENÁ	STD30	INT20	SEN70	27768	283142	1219	1512	95	100	96	99	↑ 97,4	↑ 99,1	NaN	91
OKNO256	STD30	INT20	SEN80	28044	280586	3775	1236	96	99	88	100	↑ 95,5	↑ 98,4	90	90
OKNO512	STD30	INT20	SEN80	28277	277609	6752	1003	97	98	81	100	↑ 93,6	↑ 97,5	87	89
OKNO1024	STD30	INT20	SEN80	28831	267277	17084	449	98	94	63	100	⇒ 88,8	↑ 94,4	81	85
SLOUČENÁ	STD30	INT20	SEN80	27938	282564	1797	1342	95	99	94	100	↑ 97,1	↑ 99,0	NaN	91
OKNO256	STD30	INT20	SEN90	28052	263912	20449	1228	96	93	58	100	⇒ 86,5	↑ 93,1	90	85
OKNO512	STD30	INT20	SEN90	28277	260991	23370	1003	97	92	55	100	⇒ 85,7	↑ 92,2	87	84
OKNO1024	STD30	INT20	SEN90	28831	245775	38586	449	98	86	43	100	⇒ 81,9	⇒ 87,6	81	79
SLOUČENÁ	STD30	INT20	SEN90	27946	278368	5993	1334	95	98	82	100	↑ 93,8	↑ 97,7	NaN	89
OKNO256	STD30	INT20	SEN100	28532	157890	126471	748	97	56	18	100	↓ 67,7	↓ 59,4	90	51
OKNO512	STD30	INT20	SEN100	29280	142016	142345	0	100	50	17	100	↓ 66,8	↓ 54,6	87	45
OKNO1024	STD30	INT20	SEN100	28839	132351	152010	441	98	47	16	100	↓ 65,2	↓ 51,4	81	42
SLOUČENÁ	STD30	INT20	SEN100	28347	223313	61048	933	97	79	32	100	↓ 76,7	⇒ 80,2	NaN	71

5.1.5 Výsledky SIM_REAL2

Celkem bylo testováno 8 kombinací nastavení parametrů metody STD, INT, SEN, přičemž výsledky byly získány pro analýzy signálu jednotlivými okny (256, 512, 1024 vzorků) a pro jejich sloučení – celkem bylo získáno 32 unikátních výsledků. V tabulce 5.4 jsou výsledky testování metody na simulovaných semisyntetických datech. Učení NEP (%) označuje, z kolika % záznamu systém určoval etalon NEP. NEP (%) označuje kolik % záznamu bylo nakonec označeno jako NEP. Výsledky metrik Index/Score a Přesnost vyšší než 90 % jsou označeny zelenou šipkou, výsledky v rozmezí 80-90 % žlutou šipkou, ostatní červenou.

Úspěšnosti přes 90 % v obou metrikách (Index/Score, Přesnost) dosáhlo 15 z 32 testovaných nastavení (47 % testovaných nastavení).

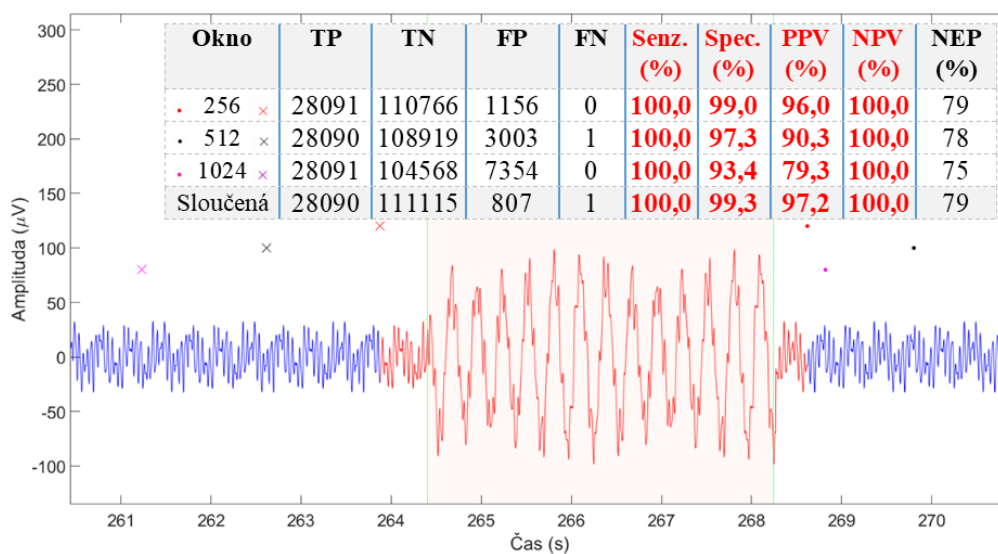
Tabulka 5.4. Výsledky metody na SIM_REAL2 pro testovaná nastavení.

Okno	STD	INT	SEN	TP	TN	FP	FN	Senz.	Spec.	PPV	NPV	Index Score (%)	Přesnost (%)	Učení NEP (%)	NEP (%)
OKNO256	STD20	INT20	SEN70	7534	52880	888	0	100	98	89	100	↑ 97,0	↑ 98,6	87	86
OKNO512	STD20	INT20	SEN70	7534	51120	2648	0	100	95	74	100	↑ 92,3	↑ 95,7	85	83
OKNO1024	STD20	INT20	SEN70	7534	48950	4818	0	100	91	61	100	⇒ 88,0	↑ 92,1	85	80
SLOUČENÁ	STD20	INT20	SEN70	7534	53208	560	0	100	99	93	100	↑ 98,0	↑ 99,1	NaN	87
OKNO256	STD20	INT20	SEN80	7534	52412	1356	0	100	97	85	100	↑ 95,6	↑ 97,8	87	85
OKNO512	STD20	INT20	SEN80	7534	50296	3472	0	100	94	68	100	↑ 90,5	↑ 94,3	85	82
OKNO1024	STD20	INT20	SEN80	7534	48026	5742	0	100	89	57	100	⇒ 86,5	↑ 90,6	85	78
SLOUČENÁ	STD20	INT20	SEN80	7534	53052	716	0	100	99	91	100	↑ 97,5	↑ 98,8	NaN	87
OKNO256	STD20	INT20	SEN90	7534	49566	4202	0	100	92	64	100	⇒ 89,1	↑ 93,1	87	81
OKNO512	STD20	INT20	SEN90	7534	46438	7330	0	100	86	51	100	⇒ 84,3	⇒ 88,0	85	76
OKNO1024	STD20	INT20	SEN90	7534	44280	9488	0	100	82	44	100	⇒ 81,7	⇒ 84,5	85	72
SLOUČENÁ	STD20	INT20	SEN90	7534	52092	1676	0	100	97	82	100	↑ 94,7	↑ 97,3	NaN	85
OKNO256	STD20	INT20	SEN100	7534	32670	21098	0	100	61	26	100	↓ 71,8	↓ 65,6	87	53
OKNO512	STD20	INT20	SEN100	7534	27734	26034	0	100	52	22	100	↓ 68,5	↓ 57,5	85	45
OKNO1024	STD20	INT20	SEN100	7534	24126	29642	0	100	45	20	100	↓ 66,3	↓ 51,6	85	39
SLOUČENÁ	STD20	INT20	SEN100	7534	43210	10558	0	100	80	42	100	⇒ 80,5	⇒ 82,8	NaN	70
OKNO256	STD30	INT20	SEN70	7534	52880	888	0	100	98	89	100	↑ 97,0	↑ 98,6	88	86
OKNO512	STD30	INT20	SEN70	7534	51532	2236	0	100	96	77	100	↑ 93,2	↑ 96,4	87	84
OKNO1024	STD30	INT20	SEN70	7534	48950	4818	0	100	91	61	100	⇒ 88,0	↑ 92,1	86	80
SLOUČENÁ	STD30	INT20	SEN70	7534	53208	560	0	100	99	93	100	↑ 98,0	↑ 99,1	NaN	87
OKNO256	STD30	INT20	SEN80	7534	52724	1044	0	100	98	88	100	↑ 96,5	↑ 98,3	88	86
OKNO512	STD30	INT20	SEN80	7534	51532	2236	0	100	96	77	100	↑ 93,2	↑ 96,4	87	84
OKNO1024	STD30	INT20	SEN80	7534	48026	5742	0	100	89	57	100	⇒ 86,5	↑ 90,6	86	78
SLOUČENÁ	STD30	INT20	SEN80	7534	53208	560	0	100	99	93	100	↑ 98,0	↑ 99,1	NaN	87
OKNO256	STD30	INT20	SEN90	7534	50758	3010	0	100	94	71	100	↑ 91,5	↑ 95,1	88	83
OKNO512	STD30	INT20	SEN90	7534	49834	3934	0	100	93	66	100	⇒ 89,6	↑ 93,6	87	81
OKNO1024	STD30	INT20	SEN90	7534	45204	8564	0	100	84	47	100	⇒ 82,7	⇒ 86,0	86	74
SLOUČENÁ	STD30	INT20	SEN90	7534	52608	1160	0	100	98	87	100	↑ 96,1	↑ 98,1	NaN	86
OKNO256	STD30	INT20	SEN100	7534	34418	19350	0	100	64	28	100	↓ 73,0	↓ 68,4	88	56
OKNO512	STD30	INT20	SEN100	7534	31330	22438	0	100	58	25	100	↓ 70,9	↓ 63,4	87	51
OKNO1024	STD30	INT20	SEN100	7534	25050	28718	0	100	47	21	100	↓ 66,8	↓ 53,2	86	41
SLOUČENÁ	STD30	INT20	SEN100	7534	44972	8796	0	100	84	46	100	⇒ 82,4	⇒ 85,7	NaN	73

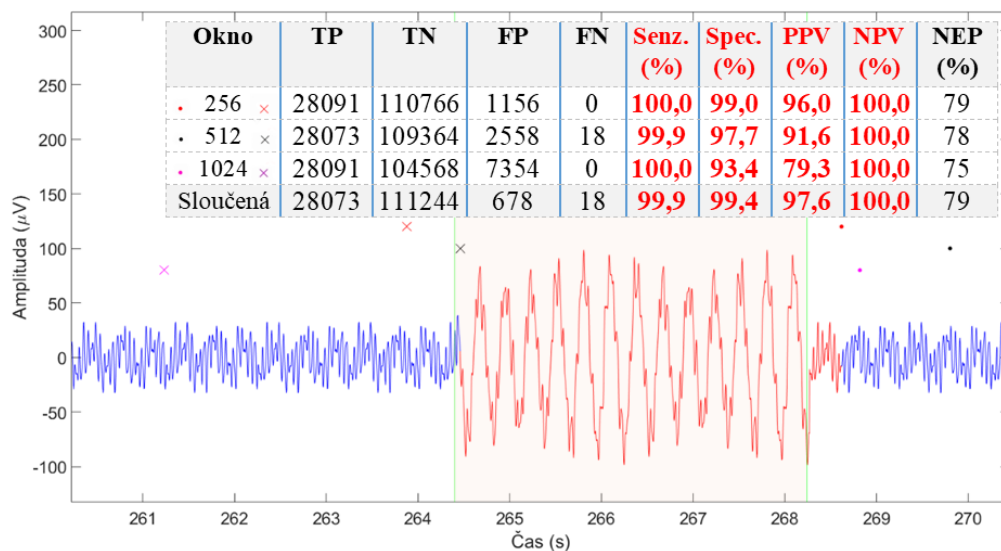
5.1.6 Grafická ukázka výsledků na simulacích pro nejlepších kandidátů nastavení

SIM_SYNT

Na obrázcích 5.1 (nastavení 1, SEN = 70 %) a 5.2 (nastavení 2, SEN = 70 %) naleznete grafické ukázky výsledků navržené metody adaptivní segmentace na simulovaných syntetický datech. Tečky v grafech označují začátek metodou klasifikovaného NEP úseku, křížky jeho konec pro jednotlivá použitá okna. Červená část v grafu reprezentuje výsledek sloučení analýz, tedy finální výsledek segmentace. Podbarvená oblast grafu je oblast, ve které se artefakt skutečně vyskytoval.



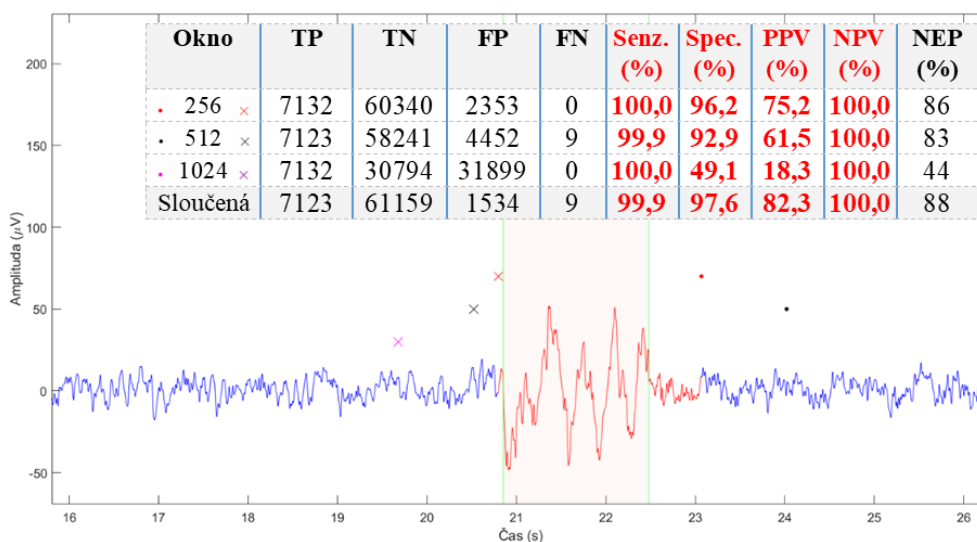
Obrázek 5.1. Grafické výsledky SIM_SYNT pro SEN = 70 % v kombinaci s Nastavením 1



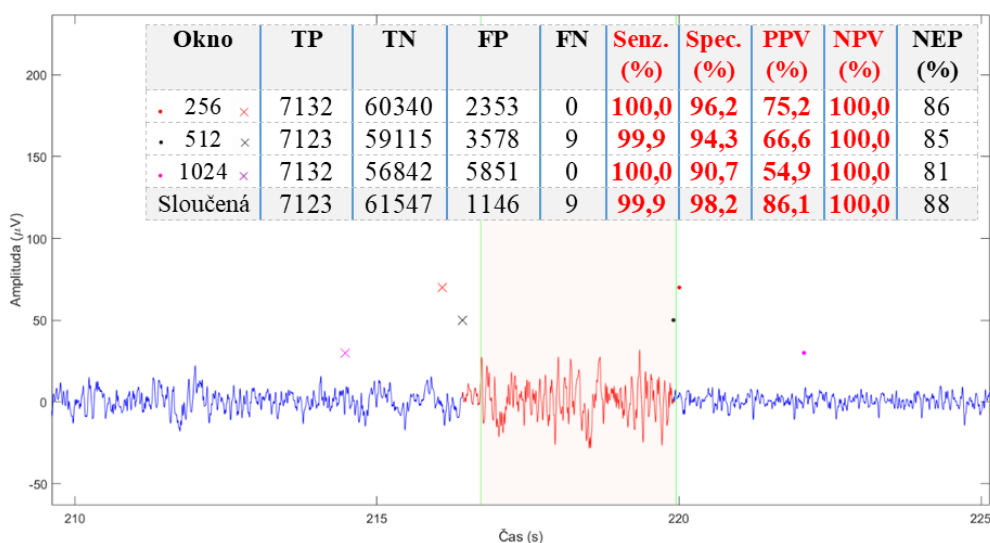
Obrázek 5.2. Grafické výsledky SIM_SYNT pro SEN = 80 % v kombinaci s Nastavením 2

SIM_SEMISYNT

Na obrázcích 5.3 (nastavení 1, SEN = 70 %) a 5.4 (nastavení 2, SEN = 80 %) naleznete grafické ukázky výsledků navržené metody adaptivní segmentace na simulovaných semisyntetický datech. Tečky v grafech označují začátek metodou klasifikovaného NEP úseku, křížky jeho konec pro jednotlivá použitá okna. Červená část v grafu reprezentuje výsledek sloučení analýz, tedy finální výsledek segmentace. Podbarvená oblast grafu je oblast, ve které se artefakt skutečně vyskytoval.



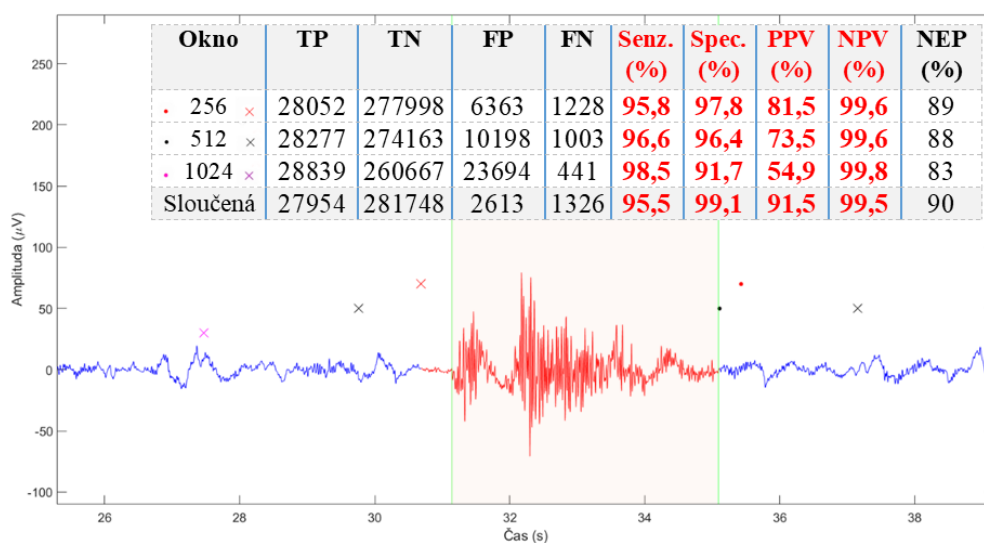
Obrázek 5.3. Grafické výsledky SIM_SEMISYNT pro SEN = 70 % v kombinaci s Nastavením 1



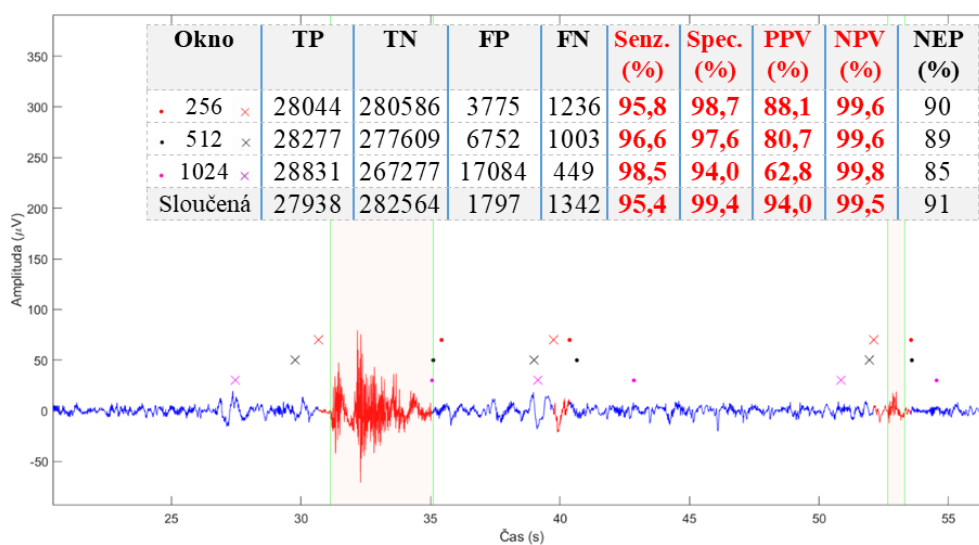
Obrázek 5.4. Grafické výsledky SIM_SYNT pro SEN = 80 % v kombinaci s Nastavením 2

SIM_REAL1

Na obrázcích 5.5 (nastavení 1, SEN = 70 %) a 5.6 (nastavení 2, SEN = 80 %) naleznete grafické ukázky výsledků navržené metody adaptivní segmentace na simulovaných reálných datech 1. Tečky v grafech označují začátek metodou klasifikovaného NEP úseku, křížky jeho konec pro jednotlivá použitá okna. Červená část v grafu reprezentuje výsledek sloučení analýz, tedy finální výsledek segmentace. Podbarvená oblast grafu je oblast, ve které se artefakt skutečně vyskytoval.



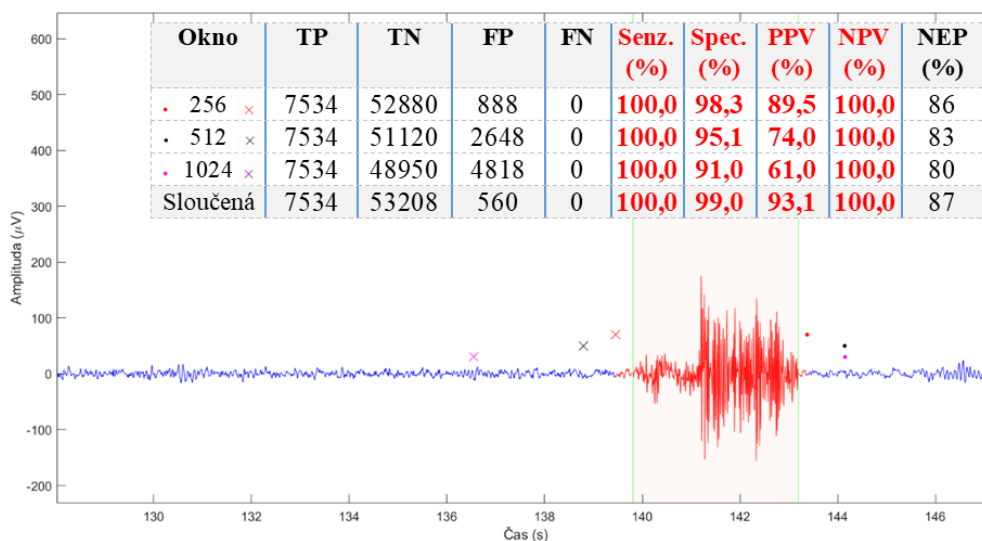
Obrázek 5.5. Grafické výsledky SIM_REAL1 pro SEN = 70 % v kombinaci s Nastavením 1



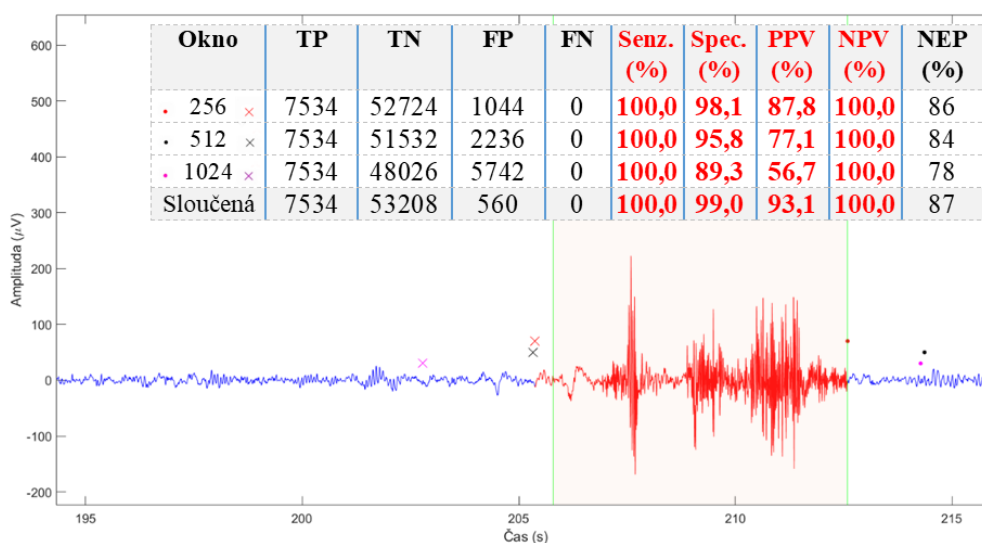
Obrázek 5.6. Grafické výsledky SIM_REAL1 pro SEN = 80 % v kombinaci s Nastavením 2

SIM_REAL2

Na obrázcích 5.7 (nastavení 1, SEN = 70 %) a 5.8 (nastavení 2, SEN = 80 %) naleznete grafické ukázky výsledků navržené metody adaptivní segmentace na simulovaných reálných datech 2. Tečky v grafech označují začátek metodou klasifikovaného NEP úseku, křížky jeho konec pro jednotlivá použitá okna. Červená část v grafu reprezentuje výsledek sloučení analýz, tedy finální výsledek segmentace. Podbarvená oblast grafu je oblast, ve které se artefakt skutečně vyskytoval.



Obrázek 5.7. Grafické výsledky SIM_REAL2 pro SEN = 70 % v kombinaci s Nastavením 1



Obrázek 5.8. Grafické výsledky SIM_REAL1 pro SEN = 80 % v kombinaci s Nastavením 2

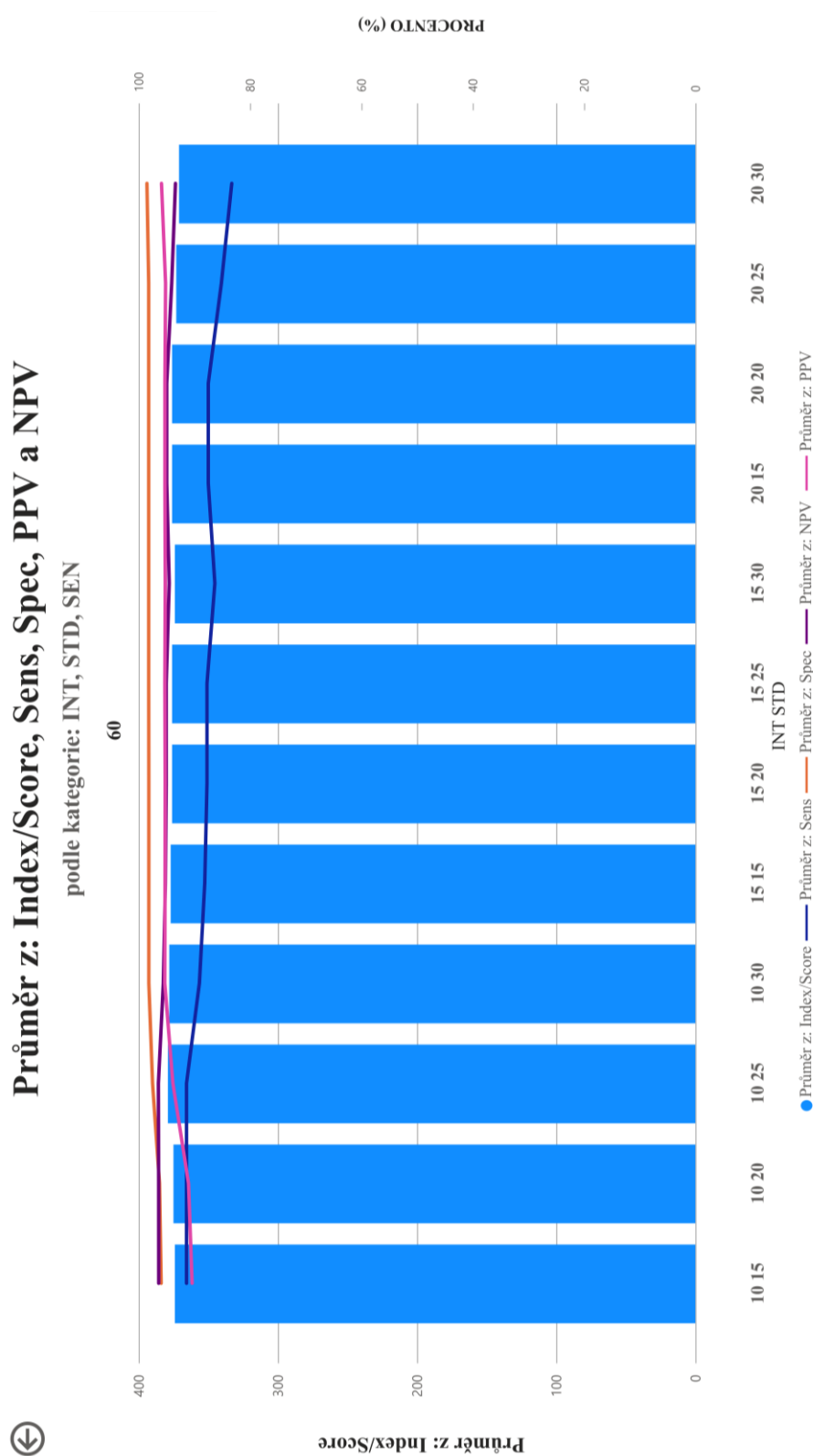
5.2 Výsledky anotovaných reálných dat

Celkem bylo testováno 60 kombinací nastavení parametrů metody STD, INT, SEN, přičemž výsledky byly získány pro analýzy signálu jednotlivými okny (256, 512, 1024 vzorků) a pro jejich sloučení – celkem bylo získáno 240 unikátních výsledků. 38 % (v absolutním čísle 92) z těchto výsledků dosahovalo úspěšnosti v obou statistických metrikách (Index/Score, Přesnost) přes 90 % úspěšnosti (označení zelenou šipkou). V tabulce 5.5 jsou výsledky pro 40 nejlepších kandidátů nastavení metody, kteří dosáhli úspěšnosti v obou metrikách (Index/Score, Přesnost) přes 94 %. V kapitole 5.2.1 jsou obrázky sumarizace výsledků metody pro různá nastavení citlivosti metody (SEN), v kapitole 5.2.2 jsou pak výsledky metody pro jednotlivá okna používaná v této práci a jejich finálního sloučení. Analýzu výsledků v prostředí Power BI společnosti Microsoft pro metriku Index/Score naleznete v kapitole 5.2.3. Grafickou ukázkou výsledků testování metody na anotovaných reálných datech naleznete v kapitole 5.2.4.

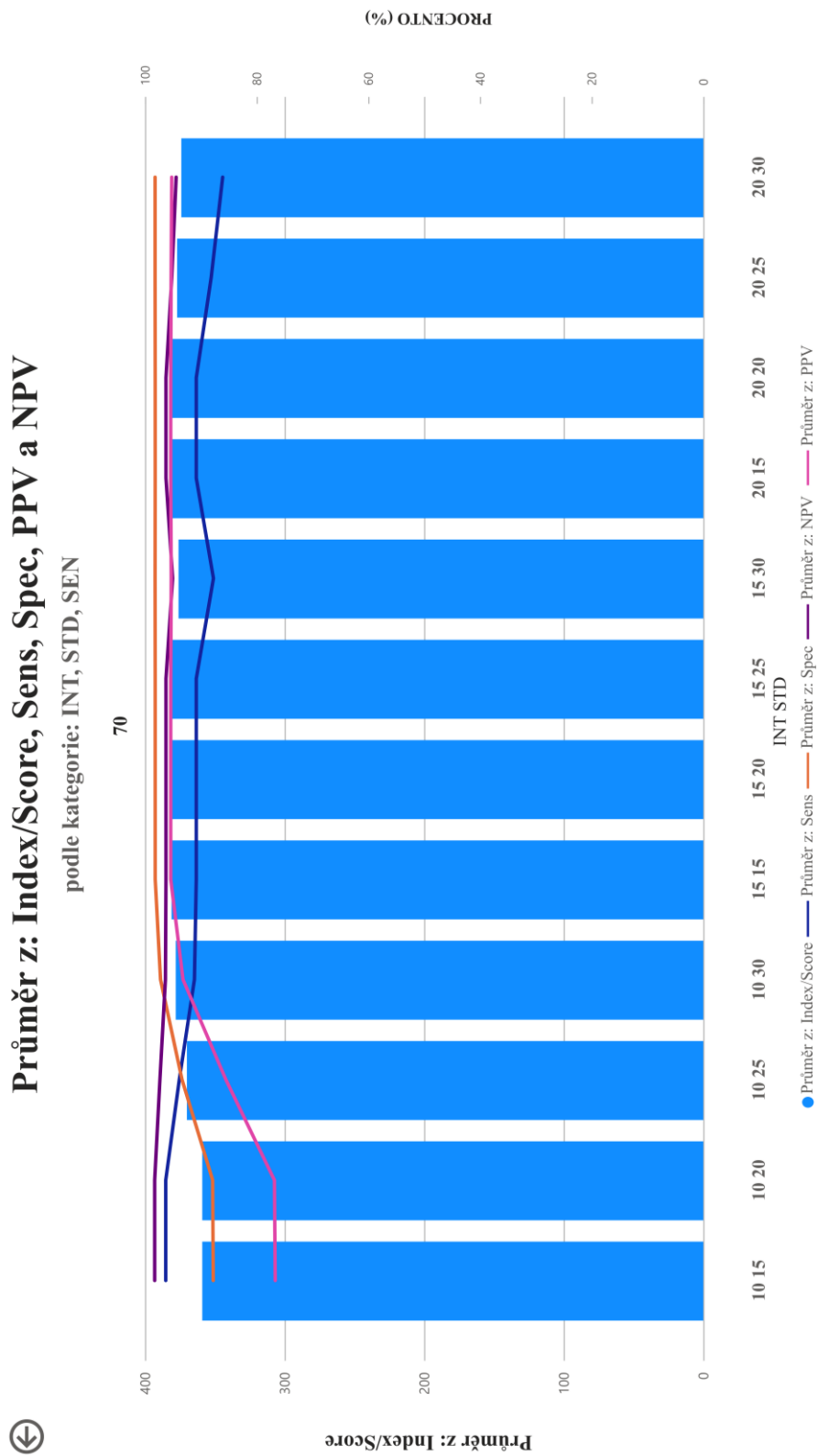
Tabulka 5.5. Výsledky 40 nejlepších nastavení na REAL_ANOT datech, která dosáhla v metrikách Index/Score i Přesnost více než 94 % úspěšnosti

Okno	STD	INT	SEN	TP	TN	FP	FN	Senz.	Spec.	PPV	NPV	Index Score (%)	Přesnost (%)	Učení NEP (%)	NEP (%)
OKNO256	STD25	INT20	SEN80	20976	50727	1498	799	96	97	93	98	↑ 96,3	↑ 96,9	72	70
SLOUČENÁ	STD15	INT20	SEN80	20443	51104	1121	1332	94	98	95	97	↑ 96,0	↑ 96,7	0	71
OKNO256	STD25	INT15	SEN70	20669	50832	1393	1106	95	97	94	98	↑ 96,0	↑ 96,6	72	70
SLOUČENÁ	STD20	INT20	SEN80	20367	51104	1121	1408	94	98	95	97	↑ 95,9	↑ 96,6	0	71
OKNO256	STD30	INT10	SEN60	20590	50803	1422	1185	95	97	94	98	↑ 95,8	↑ 96,5	73	70
OKNO256	STD15	INT20	SEN80	20976	50365	1860	799	96	96	92	98	↑ 95,8	↑ 96,4	72	69
OKNO256	STD20	INT20	SEN80	20976	50365	1860	799	96	96	92	98	↑ 95,8	↑ 96,4	72	69
OKNO256	STD25	INT15	SEN60	20190	51177	1048	1585	93	98	95	97	↑ 95,7	↑ 96,4	72	71
SLOUČENÁ	STD25	INT15	SEN80	20367	50898	1327	1408	94	97	94	97	↑ 95,5	↑ 96,3	0	71
OKNO256	STD30	INT15	SEN80	20396	50865	1360	1379	94	97	94	97	↑ 95,5	↑ 96,3	73	71
OKNO256	STD15	INT20	SEN70	20396	50815	1410	1379	94	97	94	97	↑ 95,5	↑ 96,2	72	71
OKNO256	STD20	INT20	SEN70	20396	50815	1410	1379	94	97	94	97	↑ 95,5	↑ 96,2	72	71
OKNO256	STD25	INT20	SEN70	19984	51177	1048	1791	92	98	95	97	↑ 95,4	↑ 96,2	72	72
OKNO256	STD30	INT15	SEN70	19984	51177	1048	1791	92	98	95	97	↑ 95,4	↑ 96,2	73	72
SLOUČENÁ	STD15	INT20	SEN90	20695	50412	1813	1080	95	97	92	98	↑ 95,4	↑ 96,1	0	70
SLOUČENÁ	STD20	INT20	SEN90	20695	50412	1813	1080	95	97	92	98	↑ 95,4	↑ 96,1	0	70
OKNO256	STD15	INT20	SEN60	19863	51262	963	1912	91	98	95	96	↑ 95,3	↑ 96,1	72	72
OKNO256	STD20	INT20	SEN60	19863	51262	963	1912	91	98	95	96	↑ 95,3	↑ 96,1	72	72
SLOUČENÁ	STD25	INT20	SEN80	19991	51116	1109	1784	92	98	95	97	↑ 95,3	↑ 96,1	0	71
SLOUČENÁ	STD15	INT15	SEN80	20695	50380	1845	1080	95	96	92	98	↑ 95,3	↑ 96,0	0	70
SLOUČENÁ	STD20	INT15	SEN80	20695	50380	1845	1080	95	96	92	98	↑ 95,3	↑ 96,0	0	70
SLOUČENÁ	STD15	INT15	SEN70	19776	51271	954	1999	91	98	95	96	↑ 95,2	↑ 96,0	0	72
SLOUČENÁ	STD15	INT20	SEN70	19776	51271	954	1999	91	98	95	96	↑ 95,2	↑ 96,0	0	72
SLOUČENÁ	STD20	INT15	SEN70	19776	51271	954	1999	91	98	95	96	↑ 95,2	↑ 96,0	0	72
SLOUČENÁ	STD20	INT20	SEN70	19776	51271	954	1999	91	98	95	96	↑ 95,2	↑ 96,0	0	72
SLOUČENÁ	STD25	INT15	SEN70	19776	51271	954	1999	91	98	95	96	↑ 95,2	↑ 96,0	0	72
OKNO256	STD15	INT15	SEN60	20190	50815	1410	1585	93	97	93	97	↑ 95,1	↑ 96,0	72	71
OKNO256	STD20	INT15	SEN60	20190	50815	1410	1585	93	97	93	97	↑ 95,1	↑ 96,0	72	71
OKNO256	STD30	INT20	SEN80	19828	51177	1048	1947	91	98	95	96	↑ 95,1	↑ 96,0	73	72
OKNO256	STD15	INT15	SEN70	20669	50208	2017	1106	95	96	91	98	↑ 95,0	↑ 95,8	72	69
OKNO256	STD20	INT15	SEN70	20669	50208	2017	1106	95	96	91	98	↑ 95,0	↑ 95,8	72	69
OKNO256	STD25	INT15	SEN80	20976	49847	2378	799	96	95	90	98	↑ 95,0	↑ 95,7	72	68
SLOUČENÁ	STD25	INT10	SEN60	19903	50922	1303	1872	91	98	94	96	↑ 94,8	↑ 95,7	0	71
SLOUČENÁ	STD25	INT20	SEN90	20012	50795	1430	1763	92	97	93	97	↑ 94,8	↑ 95,7	0	71
SLOUČENÁ	STD30	INT20	SEN90	19677	51098	1127	2098	90	98	95	96	↑ 94,7	↑ 95,6	0	72
OKNO256	STD30	INT15	SEN60	19451	51262	963	2324	89	98	95	96	↑ 94,6	↑ 95,6	73	72
OKNO256	STD30	INT20	SEN70	19451	51262	963	2324	89	98	95	96	↑ 94,6	↑ 95,6	73	72
SLOUČENÁ	STD30	INT15	SEN80	19400	51283	942	2375	89	98	95	96	↑ 94,6	↑ 95,5	0	73
SLOUČENÁ	STD30	INT10	SEN60	19400	51271	954	2375	89	98	95	96	↑ 94,5	↑ 95,5	0	72
SLOUČENÁ	STD30	INT10	SEN70	19848	50775	1450	1927	91	97	93	96	↑ 94,5	↑ 95,4	0	71

5.2.1 Sumarizace výsledků pro různá nastavení citlivosti metody (SEN)



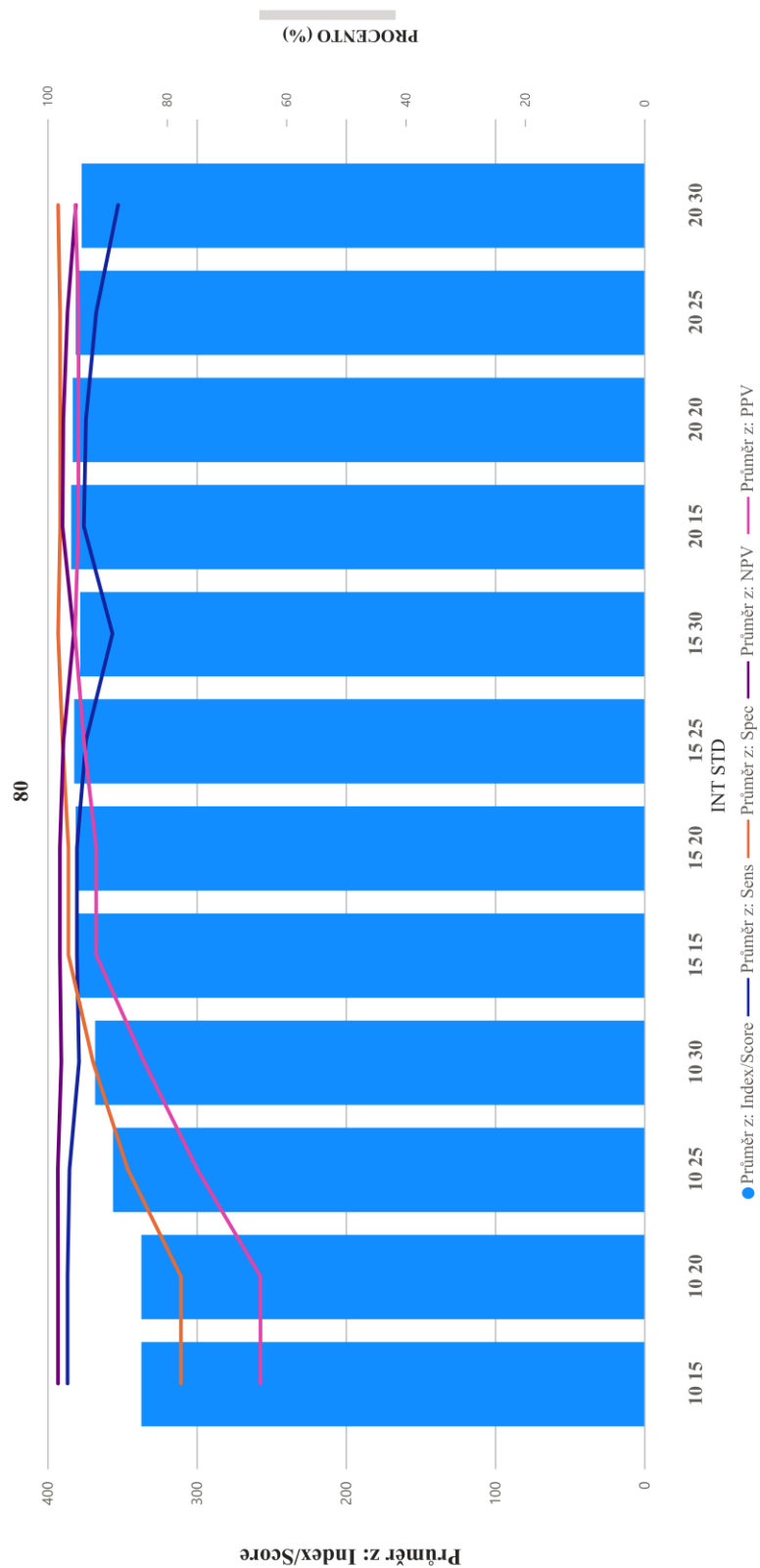
Obrázek 5.9. Sumarizace výsledků metody pro citlivost metody SEN = 60 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.



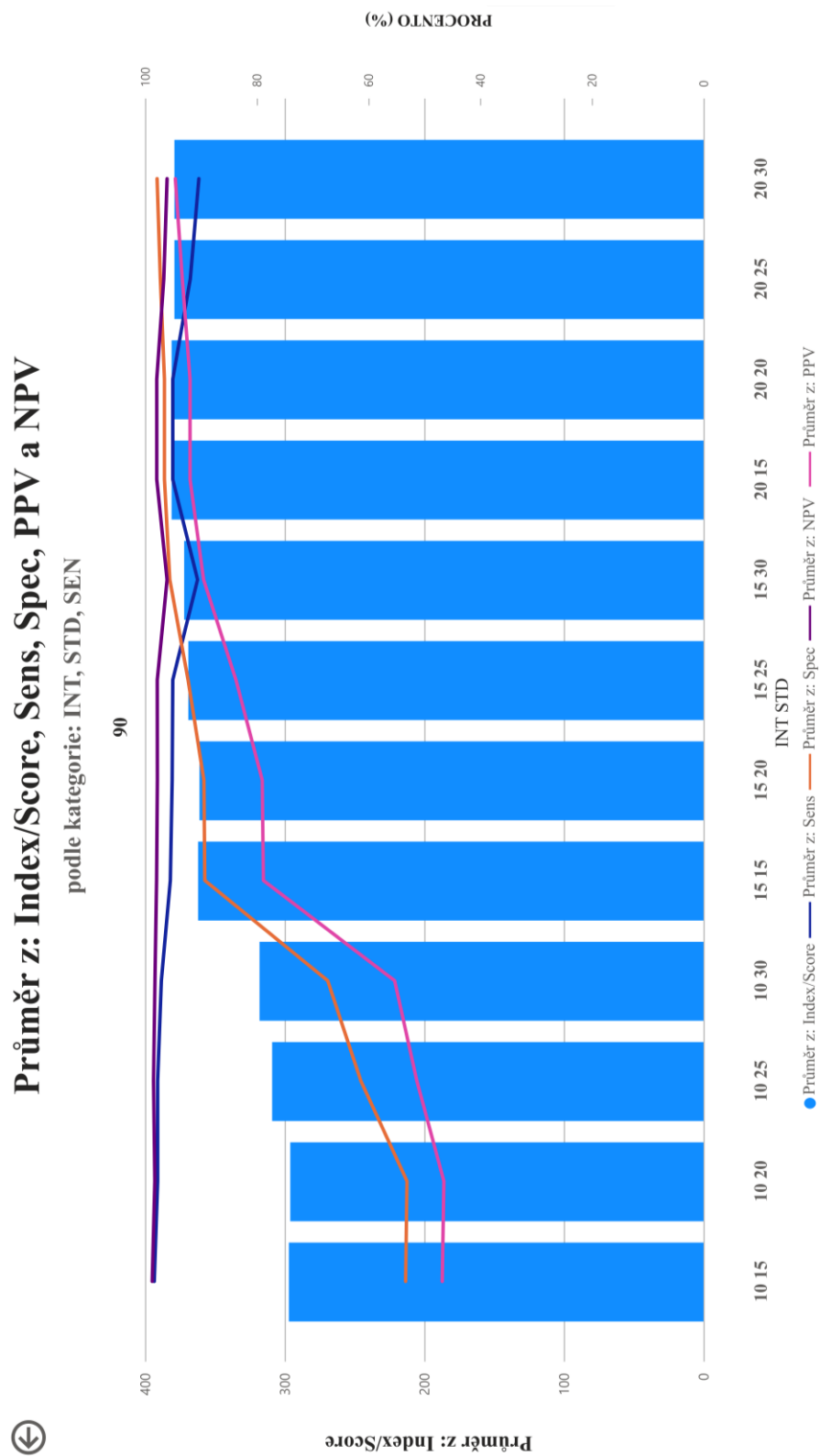
Obrázek 5.10. Sumarizace výsledků metody pro citlivost metody SEN = 70 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100 %) se vztahuje ke křivkám senzitivity, specifity, PPV a NPV.

Průměr z: Index/Score, Sens, Spec, PPV a NPV

podle kategorie: INT, STD, SEN



Obrázek 5.11. Sumarizace výsledků metody pro citlivost metody SEN = 80 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.

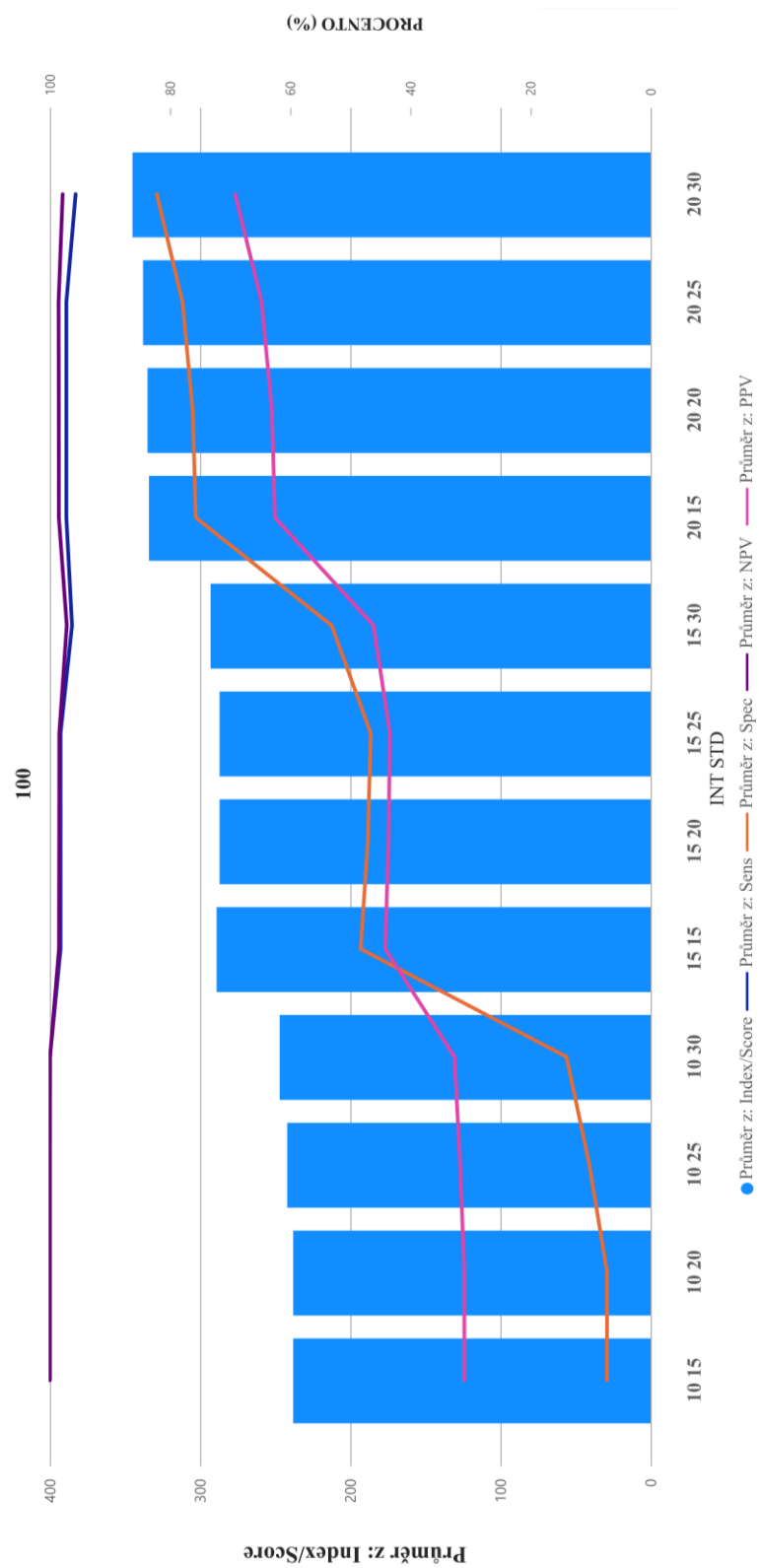


Obrázek 5.12. Sumarizace výsledků metody pro citlivost metody SEN = 90 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specifity, PPV a NPV.



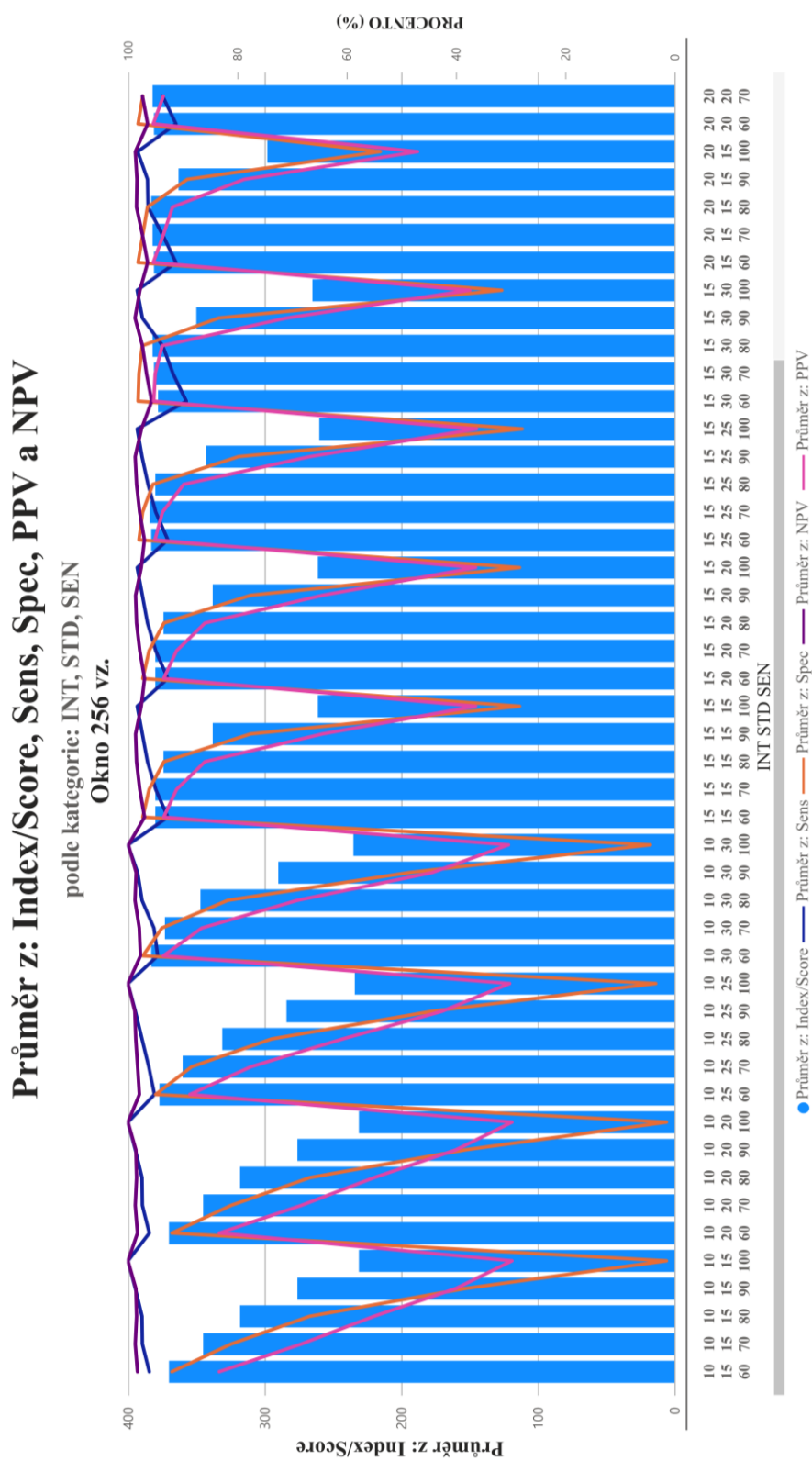
Průměr z: Index/Score, Sens, Spec, PPV a NPV

podle kategorie: INT, STD, SEN



Obrázek 5.13. Sumarizace výsledků metody pro citlivost metody SEN = 100 %. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.

5.2.2 Sumarizace výsledků pro jednotlivá okna a jejich sloučení

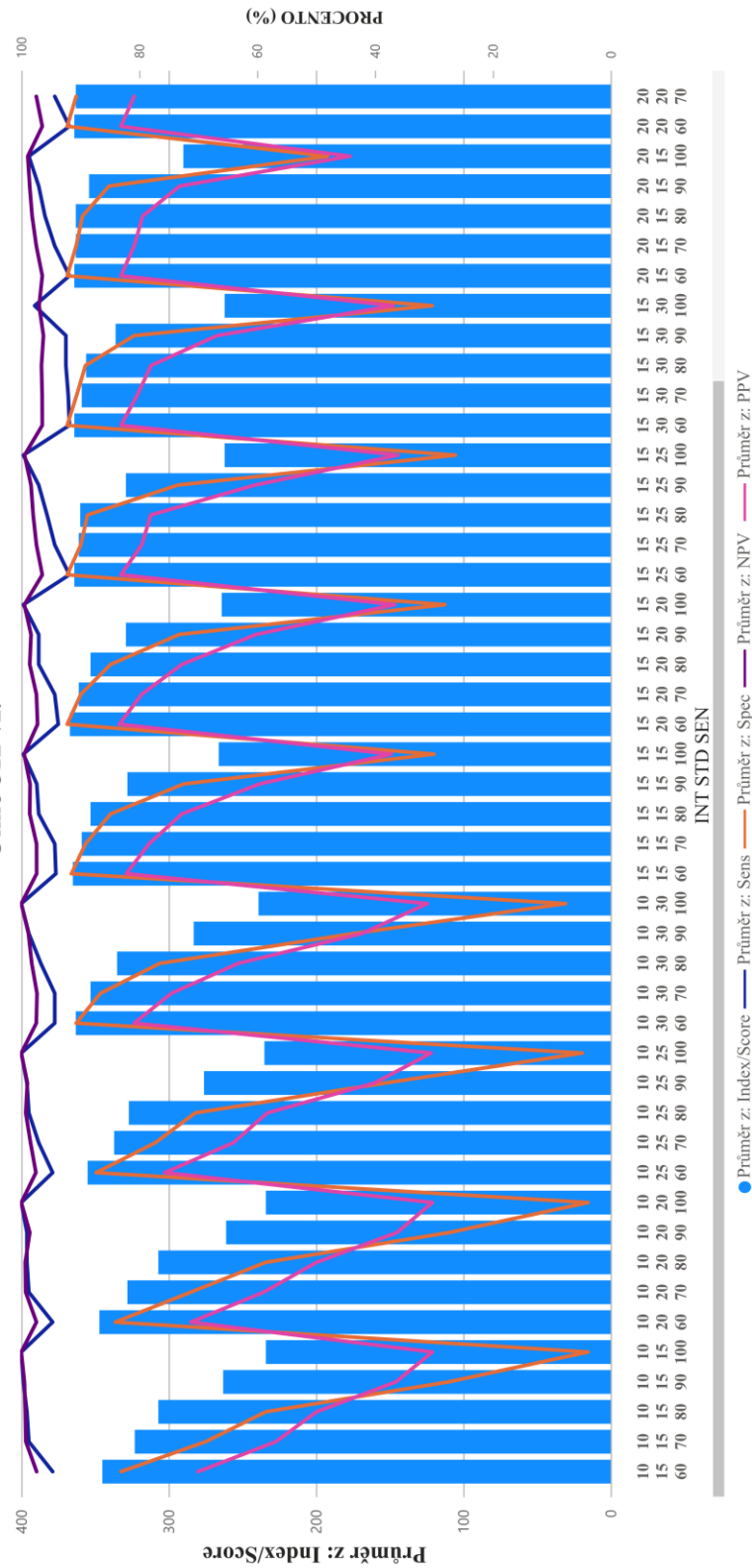


Obrázek 5.14. Sumarizace výsledků metody pro okno 256 vzorků. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.

Průměr z: Index/Score, Sens, Spec, PPV a NPV

podle kategorie: INT, STD, SEN

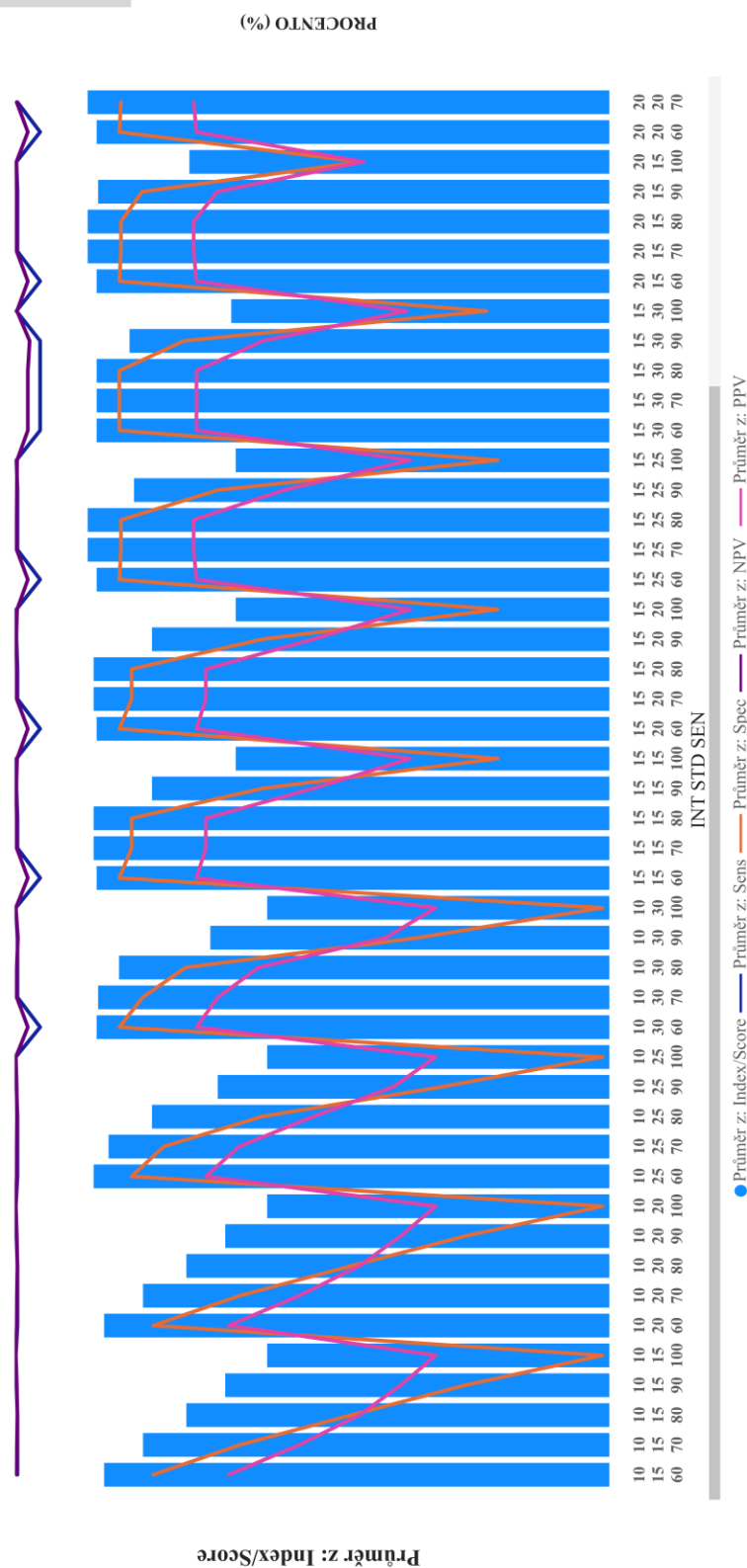
Okno 512 vz.



Obrázek 5.15. Sumarizace výsledků metody pro okno 512 vzorků. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specifity, PPV a NPV.

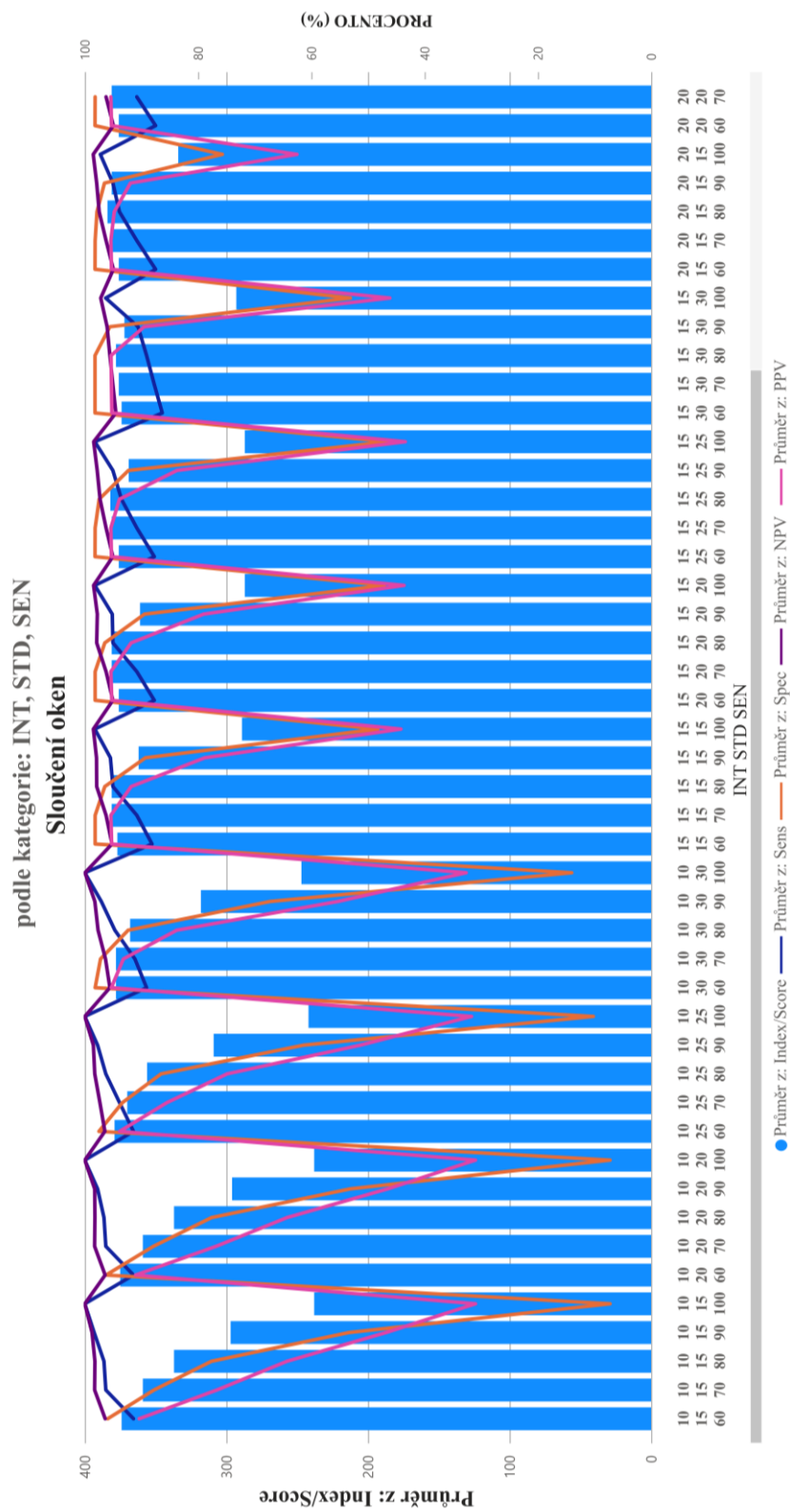
Průměr z: Index/Score, Sens, Spec, PPV a NPV

podle kategorie: INT, STD, SEN
Okno 1024 vz.



Obrázek 5.16. Sumarizace výsledků metody pro okno 1024 vzorků. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specificity, PPV a NPV.

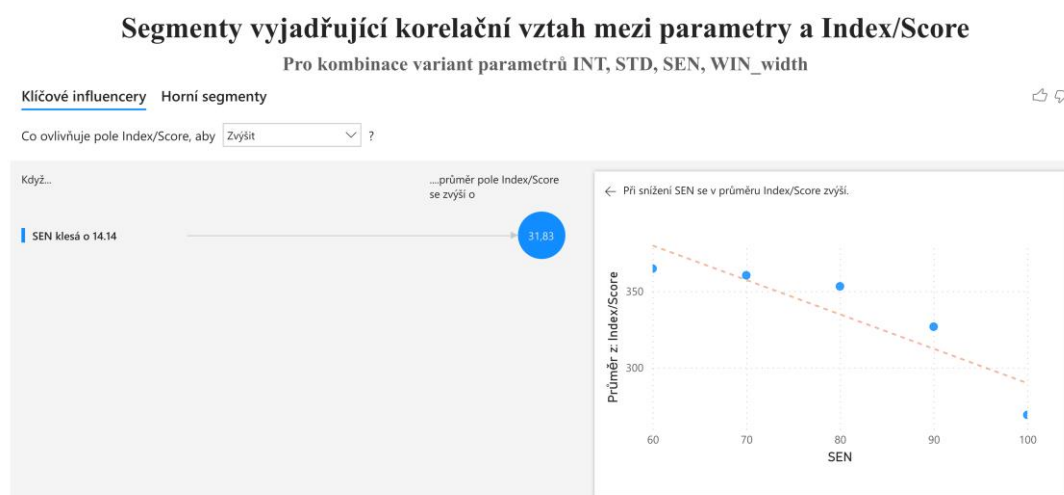
Průměr z: Index/Score, Sens, Spec, PPV a NPV



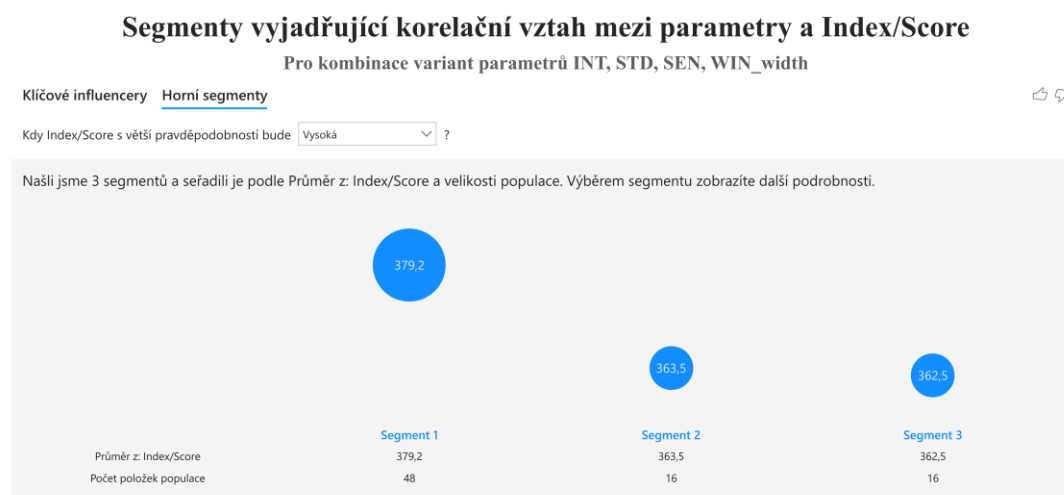
Obrázek 5.17. Sumarizace výsledků metody pro sloučení oken. Y-osa vlevo (rozsah 0-400) se vztahuje k metrice Index/Score. Y-osa vpravo (rozsah 0-100) % se vztahuje ke křivkám senzitivity, specifity, PPV a NPV.

5.2.3 Analýza v prostředí Power BI pro metriku Index/Score

Na obrázku 5.18 jsou výsledky analýzy v prostředí Power BI, která našla signifikantní vztah mezi parametrem SEN (klíčový influencer) a výsledky metriky Index/Score. Dále identifikovala v rámci všech získaných výsledků 3 ohraničené skupiny nastavení (segmenty) při kritériu vysoké pravděpodobnosti pro vysoké výsledky metriky Index/Score, jejichž detailní rozpis je na obrázku 5.19.



Parametr SEN nejvíce přispívá k růstu hodnoty Index/Score. V průměru se hodnota Index/Score zvyšuje o 8% - pokud snížíme nastavení hodnoty SEN o 15%.



Parametr SEN nejvíce přispívá k růstu hodnoty Index/Score. V průměru se hodnota Index/Score zvyšuje o 8% - pokud snížíme nastavení hodnoty SEN o 15%.

Obrázek 5.18. Výsledky analýzy v prostředí Power BI ve vztahu k metrice Index/Score. Klíčové influencery a identifikované segmenty různých nastavení.

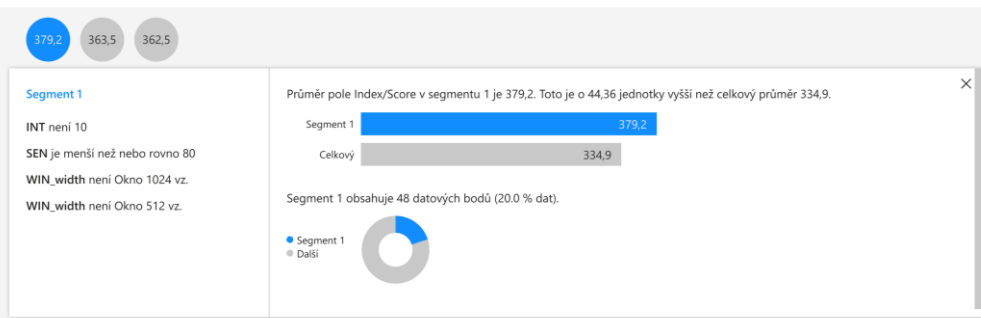
Segmenty vyjadřující korelační vztah mezi parametry a Index/Score

Pro kombinace variant parametrů INT, STD, SEN, WIN_width

Klíčové influencery Horní segmenty



Kdy Index/Score s větší pravděpodobností bude ?



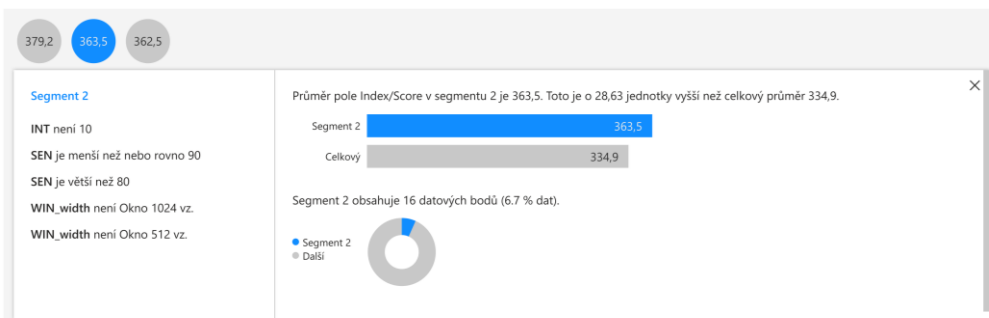
Segmenty vyjadřující korelační vztah mezi parametry a Index/Score

Pro kombinace variant parametrů INT, STD, SEN, WIN_width

Klíčové influencery Horní segmenty



Kdy Index/Score s větší pravděpodobností bude ?



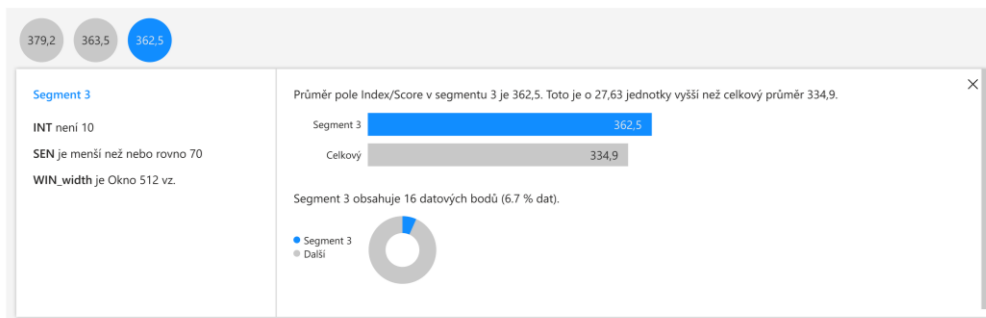
Segmenty vyjadřující korelační vztah mezi parametry a Index/Score

Pro kombinace variant parametrů INT, STD, SEN, WIN_width

Klíčové influencery Horní segmenty



Kdy Index/Score s větší pravděpodobností bude ?



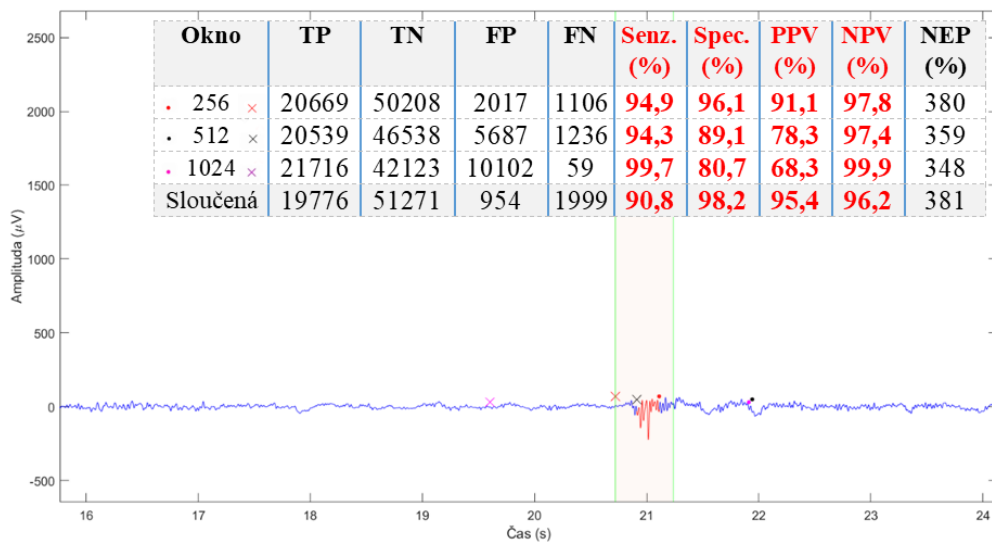
Obrázek 5.19. Detailní popis identifikovaných segmentů v prostředí Power BI

5.2.4 Grafická ukázka výsledků anotovaných reálných dat jednoho z nejlepších kandidátů nastavení

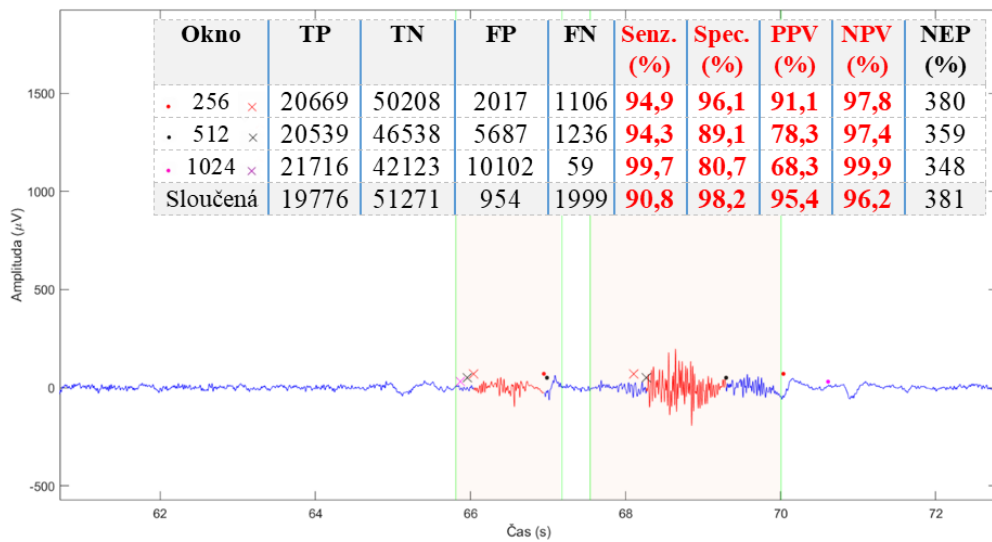
Na obrázcích 5.20-5.22 je pro ilustrační ukázkou zobrazeno hodnocení reálného záznamu jedním z nejúspěšnějších nastavení. Tečky v grafech označují začátek metodou klasifikovaného NEP úseku, křížky jeho konec pro jednotlivá použitá okna. Červená část v grafu reprezentuje výsledek sloučení analýz, tedy finální výsledek segmentace. Podbarvená oblast grafu je oblast, ve které se artefakt skutečně vyskytoval.

Toto nastavení dosáhlo hodnocení metrikou Index/Score na hodnotu 384 bodů úspěšnosti, což je v relativní hodnotě 96,0 % (vztažené k maximu možné úspěšnosti 400 bodů) a standartní metrika Přesnost dosáhla hodnoty 96,7 %. Jedná se o výsledek sloučení oken s nastavením:

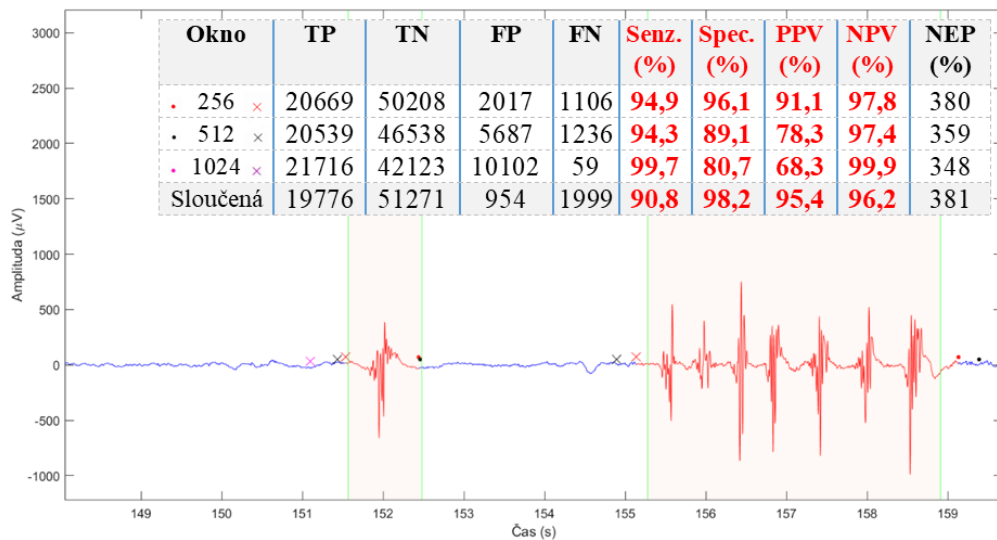
- STD = 1,5
- INT = 2,0
- SEN = 80 %



Obrázek 5.20. Grafická ukázka výsledků metody na anotovaných reálných datech.



Obrázek 5.21. Grafická ukázka výsledků metody na anotovaných reálných datech.



Obrázek 5.22. Grafická ukázka výsledků metody na anotovaných reálných datech.

5.3 Výpočetní čas na testovaných signálech

V tabulce 5.6 níže jsou uvedeny časy potřebné pro provedení analýzy navrženou metodou adaptivní segmentace. Všechny analýzy byly prováděny na notebooku značky Legion s parametry:

- Procesor: Intel(R) Core(TM) i5-10300H CPU @ 2.50GHz 2.50 GHz
- Paměť RAM: 16 GB
- Typ systému: 64bitový operační systém, procesor pro platformu x64

Tabulka 5.6. Výpočetní čas testovaných signálů

Signál	Počet vzorků signálu	Výpočetní čas (s)
SIM_SYNT	140013	2,1
SIM_SEMISYNT	69825	1
SIM_REAL1	313641	3,8
SIM_REAL2	61302	1
REAL_ANOT	74000	1

5.4 Návrh příznaků pro UI

Pro potřeby další optimalizace nastavení parametrů (STD, INT, SEN) navržené metody adaptivní segmentace navrhuji použít tyto parametry jako vstupy pro UI při dalším testování výsledků segmentace dat. Neuronové sítě by měly být schopny ještě více zpřesnit jejich nastavení než analýza v prostředí Power BI, i v kontextu jednotlivých specifických potřeb uživatelů, např. pro segmentaci jednotlivých artefaktů, epileptických záchvatů apod.

Neuronové sítě lze také využít v návaznosti na výsledky navržené metody adaptivní segmentace pro potřeby klasifikace segmentů označených metodou jako AEP do užších podskupin, např. okulární/svalový/elektrodový artefakt, epilepsie.

Pro využití neuronových sítí založených na jednotlivých příznacích, bych vzhledem k výsledkům navržené metody adaptivní segmentace, její sledované parametry taktéž navrhol pro použití při klasifikaci s využitím hlubokého učení. Tyto parametry je možné pro účely klasifikace ještě doplnit o parametry entropické oblasti, jako jsou např. Shannonova entropie, spektrální entropie a další [49]. Dalšími parametry vstupujícími do neuronové sítě by taktéž mohly být Hjortovy parametry.

Další možností využití neuronových sítí pro klasifikaci jsou konvoluční neuronové sítě. Pro tento typ neuronových sítí, které pracují s obrazovými daty, bych doporučil použití spektrogramu označených AEP úseků. Další možností konvolučních neuronových sítí je použití všech sledovaných parametrů v navržené metodě adaptivní segmentace rozšířených o další parametry (entropická oblast, Hjortovy parametry). V tomto případě by ovšem bylo nutné tyto parametry uspořádat do vyhovující matice, aby je byla CNN schopna zpracovat.

6 Diskuse

Hlavním cílem této práce bylo vytvoření metodiky automatizované klasifikace EEG záznamu s využitím adaptivní segmentace analyzující signál v časové a frekvenční oblasti. Tuto metodiku následně otestovat skrze měřitelné metriky a navrhnout příznaky a procesy, které by bylo možné využít pro klasifikaci založené na kooperaci s umělou inteligencí.

Navržený a popsáný postup (hypotetický základ a princip navrhované v kapitole 4.1 a technické řešení metody v kapitole 4.2), byl plně realizován. Navržený algoritmus z hlediska rychlosti výpočtů vykazoval hladký a rychlý průběh (viz kapitola 5.3). Tato výkonnost je pravděpodobně kombinací dvou faktorů. Prvním je, že větší množství menších dílčích analýz probíhá rychleji a s menšími nároky na systémové zdroje než jedna velká komplexní hluboká analýza, která se snaží v rámci jednoho setu mnoha dotazů naráz odpovědět na všechny otázky najednou. Druhým je způsob zápisu dílčích výsledků jednotlivých pomocných algoritmů, které se zapisují do samostatných databází s křížovým propojením na zdrojová data v původním předzpracovaném signálu. To umožňuje metodě, aby analyzovala jen data, která ke své dílčí analýze potřebuje, aniž by je musela vyhledávat v celém zdrojovém souboru a pro analýzu si dedikuje jen tu část dat, kterou k výpočtu potřebuje.

Na základě výsledků se dá předpokládat, že hypotéza 1 (viz kapitola 4.1.1) byla potvrzena. Dílčí výsledky jednotlivých délek plovoucích oken jsou menší než agregace výsledků ve sloučení oken (viz kapitoly 5.1 a 5.2.2). Zajímavým poznatkem je, že u všech typů testovaných signálů vykazovalo samostatně největší úspěšnost nejmenší okno o velikost 256 vzorků. Pokud další dvě okna hodnotíme jako doplňkový příspěvek k oknu 256 vzorků, pak je úspěšnějším kandidátem okno délky 512 vzorků. Tento efekt se v testovaných datech jeví jako pravidlo, ale uvědomuji si, že výsledky na simulovaných signálech nelze považovat za plně relevantní, jelikož charakter předělu mezi NEP a AEP je pro navrženou metodu jednoduše identifikovatelný kvůli skokové změně v charakteristice signálu.

Při vyhodnocování výsledků byly primárně testovány kombinace tří klíčových parametrů, které lze k výsledkům měřitelně přiřadit a zároveň jsou to proměnné, jejichž kombinace a nastavení jejich hladin jsou klíčové pro úspěšnost výsledků segmentace EEG signálu. STD (násobek směrodatné odchylky při stanovení hodnoty selektivního prahu) byl v metodě testován ve variantách násobků: 1,5; 2,0; 2,5; 3,0. INT (násobek směrodatné odchylky při stanovení etalonu NEP) byl testován ve variantách násobků: 1,0; 1,5; 2,0. Posledním parametrem je SEN (citlivost metody, tzn. kolik % sledovaných parametrů musí být splněno pro klasifikaci segmentu jako NEP). Tyto 3 parametry a jejich kombinace reprezentují 60 setů různých nastavení pro testování úspěšnosti, přičemž výsledky byly získány jak pro jednotlivá okna, tak pro jejich výsledné sloučení.

Úspěšnost byla vyhodnocována skrze standartní výkonové metriky (senzitivita, specificita, PPV, NPV) používané v medicínských studiích. Protože ale hodnoty výsledků těchto 4 výše zmíněných metrik vykazovaly vůči sobě navzájem těžko interpretovatelné výkyvy hodnot, které by korelovaly s mírou úspěšnosti v % nalezeného NEP, byly jako další hodnotící metriky dodány hodnoty Index/Score a Přesnost (viz kapitola 4.3). Index/Score je indexová hodnota vypočítaná součtem hodnot senzitivity, specificity, PPV a NPV. Protože všechny tyto 4 metriky (senzitivita, specificita, PPV a NPV) mají společnou jednotku (%) a vektor růstu úspěšnosti, kdy, čím je vyšší hodnota, tím je přesnější výsledek, považuji tuto metriku za relevantní. Součet jejich absolutních hodnot pak transparentně reflektuje výsledek bez ohledu na to, jestli jednotlivě rostou, nebo klesají, což lze v kontextu a ve vztahu k ostatním třem hodnotícím metrikám těžko vysvětlit. Takto kumulované skóre je ideální pro nastavení hodnotící škály, protože i u takto kumulované hodnoty skóre platí, že čím vyšší hodnota, tím lepší výsledek bez ohledu na dílčí pokles jedné z nich. Další vypočítanou hodnotou, skrze kterou lze hodnotit úspěšnost jednotlivých setů nastavení parametrů, je standartně používaná Přesnost (viz rovnice 4.18). Jako referenční hodnoty byla do kontextu datových výsledků ještě přidána hodnota % nalezeného NEP a informace o množství záznamu, které metoda ve druhé fázi vyhodnotila jako 100% nezpochybnitelný NEP, ze kterého určovala parametry pro definici NEP v celém záznamu. Tato hodnota sice o výkonu metody nic neříká, ale pro účely hlubší analýzy byla do výsledků vložena, aby se zvýšil počet informací o interních procesech analýz, které by mohly korelovat s mírou úspěchu daného setu nastavení, což se nakonec nepotvrdilo a procento nalezeného NEP v testovaných datech signifikantně nezvyšuje úspěšnost, nebo skóre setu.

Hodnotím-li výsledky skrze tyto dvě metriky, kdy hodnota Index/Score vyjadřuje spíše kvantitativní úspěšnost, pak zvyšující se hodnota Přesnosti hodnotí kvalitativní aspekt výsledku. Metrika Index/Score je navíc zasazena do kontextu senzitivity, specificity, PPV a NPV (viz kapitola 4.3) a lze hledat korelující vztahy mezi všemi těmito metrikami.

Výsledky takto nastaveného hodnocení pro mě byly naprostým překvapením. Ukazuje se, že metrika, která nejvíce stimuluje úspěšnost, je SEN (citlivost metody). Významnost této metriky byla předpokládána, ale ve smyslu, že větší procento splněných parametrů by mělo vykazovat úspěšnější identifikaci NEP. Překvapení spočívalo v tom, že je to přesně naopak. To znamená, že nejúspěšnější nastavení byla ta, kdy se hodnota SEN pohybovala na nižších hodnotách (v testu 60 a 70 %). Při ověřování tohoto trendu za pomoci kontroly statistické integrity dat jsem pochopil, že tento výsledek je v pořádku a vlastně dává naprosto smysl.

Klíčovým ukazatelem pro vysvětlení korelace rostoucího skóre úspěšnosti konkrétní kombinace ve vztahu k proměnné SEN je z mého pozorování vztah k hodnotě rozdílu mezi

procentem záznamu, ze kterého se kombinace učila a procentem skutečného NEP v záznamu. Čím „erudovanější“ etalon NEP nastavení získá, tím je výsledek lepší. Erudovanost etalonu NEP v této metodě se tedy zvyšuje s procentem skutečného NEP v záznamu, ze kterého se učí. Tento vztah je skoro lineární. V TOP40 nejúspěšnějších nastaveních neklesla hodnota rozdílu mezi % záznamu NEP pro učení a mezi skutečným % NEP v záznamu více než o 5%. Tím se vysvětluje i souvislost mezi hodnotou SEN a úspěšností kombinace nastavení parametrů. Kombinace popíše správněji parametry pro učení tím, že prostuduje více relevantního záznamu. Parametr SEN pak v logickém chápání databázových dotazů postupně a poměrně plynule mění logické parametry dotazu podle následujících situací:

1. Když se SEN blíží k hodnotě 100 %

Mezi hodnotami je vztah, kdy se rozhoduje více logickou souvislostí „a“. Parametry jsou pevně nastaveny do jedné univerzální formulace typického NEP, které vnímá prostřednictvím vyjádření „platí pouze, když platí všechny parametry společně“.

2. SEN se blíží k hodnotě 60 %

Mezi hodnotami je vztah, kdy se rozhoduje více logickou souvislostí „nebo“. Parametry nehledají typickou shodu, ale v rámci stejných podmínek musí formulaci typického parametru nastavit pro každé okno zvlášť, které vnímá prostřednictvím vyjádření „platí pouze, když platí alespoň nějaký počet parametrů“.

Protože "typický" univerzální etalon NEP je vlastně projekcí všech individuálních projevů NEP v záznamu, je vyšší úspěšnost nižších nastavení hodnoty SEN v kombinaci s lepším zdrojem učení logická. Čím nižší hodnota SEN - tím více může kombinace nastavení parametrů uplatnit to co se o NEP dozvěděla z etalonu NEP. S klesající hodnotou SEN tedy adekvátně roste adaptivita nastavení.

Tyto výsledky potvrzují systémovou adaptivitu metody, protože lepšího výsledku lze dosáhnout tím, že rozhodnutí o tom, jestli je tento moment NEP, necháme více na rozhodnutí metody než na splnění tvrdých parametrů prahu.

Během testování klasifikace EEG záznamů docházelo ke zkreslení výsledků senzitivity, specifity, PPV a NPV, což mělo původ v nejednoznačnosti označení začátku a konce artefaktu napříč diagnostiky. Z grafických výsledků v kapitolách 5.1.6 a 5.2.4 je patrné, že i analýzy jednotlivými okny reagují na změnu v různých časových bodech, které si ovšem bývají na časové rovině velice blízko, což opět naznačuje potvrzení hypotézy 3. Řešením tohoto předpokládaného problému, který byl předmětem hypotézy 3, je chápání a interpretace předělu mezi NEP a AEP nikoli jako jednoho časového řezu, ale časovým řezem by měl být celý tento časový úsek. Z anotovaných záznamů, kde jsou vizuálně oddělené ale agregované anotace více diagnostiků, lze pozorovat, jak se jednotlivá skóre mírně liší v interpretaci začátků

i konců. Z toho vyplývá, že i z hlediska lidského hodnocení je snaha o interpretaci předělu mezi NEP a AEP jako časového bodu v běžné praxi nereálná.

Metoda přistupuje k řešení v této části identifikace NEP velmi podobně. Každá jedna varianta nastavení parametrů vstupujících do výsledku simuluje jeden „diagnostický názor“, který je reprezentován jiným pohledem a zkušeností, ze které vzešel (procento a kvalita „učící“ části záznamu, ze kterých si systém určil etalon NEP). Výsledky sloučení více oken, které interpretují výsledek jako komplementární sumarizaci více názorů, tímto pohledem představují jakési „strojové konsilium diagnostiků“, kteří se sice naprosto shodují v identifikaci hledaného problému, ale s mírně odlišným názorem na míru jeho ohraničení, což v důsledku zpřesňuje výsledek minimálně tím, že významně eliminuje možnou chybu způsobenou např. atypickým zdrojem zkušenosti jediného rozhodovatele.

Pro doplnění výše uvedených výkonnostních metrik vysvětlujících úspěšnost kombinace nastavení parametrů STD, INT a SEN, byla pro zajímavost provedena analýza výsledků specializovaným programem pro vizualizaci, datamining a modelování: Power BI Desktop společnosti Microsoft pro nalezení dalších korelujících hodnot, pokud existují. Tato analýza skrze vícerozměrné datové pole našla i přes nízký počet ovlivnitelných vstupních parametrů několik souvisejících a ohraničených segmentů, které lze interpretovat skrze vstupní hodnoty parametrů nastavení (viz kapitola 5.2.3). V případě korelačního vztahu vstupních parametrů vztaženého k hodnotě Index/Score je identifikován shluk potvrzující větší efektivitu pro nízkou citlivost metody SEN (60, nebo 70 %) v kombinaci s oknem 256, nebo sloučením oken. Interpretace jiným pohledem na tuto analýzu by pak mohla být taková, že INT a STD neovlivňují výsledné skóre nijak významně odlišně v nastavených hodnotách, neznamená to ale, že by byly bezvýznamné. Znamená to, že i přes nastavení různých hodnot těchto nastavení, se vliv na skóre zásadním způsobem nezvyšuje ani nesnižuje.

Tyto korelace naznačené analýzou v prostředí Power BI Desktop by bylo potřeba doplnit o další proměnné a související hodnoty a skrze ně se pokusit o detailnější vysvětlení vztahů výkonnostních metrik a jejich vlivů na skóre, a to jsou možné příznaky, které lze definovat jako sérii souvisejících dotazů pro umělou inteligenci, která si další související informace dokáže extrahovat ze svého obrovského „big data“ zdroje učící množiny dat. Ukazuje se, že je reálná možnost, aby si UI dokonce vymyslela a definovala úplně nové metriky a škály hodnot, které efektivitu této metody ještě více zpřesní a automatizují. Další příležitostí pro UI je vizualizace samotného EEG záznamu včetně modelů a simulací obrovského množství variant a jejich skóre. Nejaktuálnější výzvou pro UI je určitě klasifikace nalezených AEP porovnáváním se vzorky z validovaných učících množin a databází způsobem, který může být řešením, jak kvantifikovat poznatky z osobních záznamů pacientů naprosto precizním, anonymním a etickým způsobem.

7 Závěr

V rámci práce byla navržena metodika automatizované klasifikace EEG záznamu založená na adaptivní segmentaci, která využívá analýzy v časové a frekvenční oblasti.

Navržená metoda založená na individuálním definování běžné činnosti mozku byla realizována v programovém prostředí MATLAB 2022a a testována skrze 4 typy simulovaných a anotovaných reálných dat a jeví se jako funkční a perspektivní nejen díky své úspěšnosti v identifikaci artefaktů, ale i svou nízkou náročností na výpočetní výkon.

Výsledky metodiky byly kvantitativně vyhodnoceny skrze metriky Index/Score a Přesnot. Při testování na simulovaných datech dosáhlo výsledků přes 90 % úspěšnosti pro metriku Index/Score a zároveň Přesnot v průměru 47 % (15 z 32) testovaných nastavení. Při testování na anotovaných reálných datech existovalo 38 % (92 z 240) testovaných nastavení, které dosahovaly úspěšnosti přes 90 % pro obě tyto metriky současně. Tyto výsledky potvrzují schopnost metody adaptovat se na signál a efektivně ho segmentovat.

Na základě výsledků navržené metody adaptivní segmentace byly navrženy příznaky STD, INT a SEN jako vstupní parametry pro UI, která by měla pomoci k optimalizaci jejich nastavení v závislosti na konkrétních požadavcích uživatele a zároveň v kontextu individuality záznamů jednotlivých subjektů, a tím výsledky navržené metody adaptivní segmentace ještě více zpřesnit. Dále byly navrženy příznaky využitelné pro klasifikaci vzniklých segmentů do užších podskupin.

V rámci navazující práce bylo by dobré otestovat metodu na dalších anotovaných reálných datech, která obsahují mimo jiné i události významné pro diagnostickou praxi. Další výzvou a logicky navazujícím krokem při testování této metody představuje další stupeň klasifikace nalezených artefaktů na klinicky významné/nevýznamné. Je možné, že i v tomto případě by metoda, která prvoplánově využívá hledání extrémních výkyvů pomocí definice neextrémních částí záznamu, mohla být užitečným nástrojem pro praxi.

8 Reference

- [1] COHEN, Michael. *Analyzing Neural Time Series Data: Theory and Practice*. 2014. ISBN 9780262319553.
- [2] SANEI, Saeid a Jonathon CHAMBERS. *EEG Signal Processing*. 1. Hoboken, 2007. ISBN 978-0-470-02581-9.
- [3] KRAJČA, Vladimír a Jitka MOHYLOVÁ. *Číslíkové zpracování neurofyziologických signálů*. 1. vyd. V Praze: České vysoké učení technické, 2011. ISBN 978-80-01-04721-7.
- [4] KHURANA, Vaishali, Monika GAHALAWAT, Pradeep KUMAR, Partha Pratim ROY, Debi Prosad DOGRA, Erik SCHEME a Mohammad SOLEYMANI. A Survey on Neuromarketing Using EEG Signals. *IEEE Transactions on Cognitive and Developmental Systems*. 2021, **13**(4), 732-749. ISSN 2379-8920. Dostupné z: doi:10.1109/TCDS.2021.3065200
- [5] O'CONNELL, Barbara, Steven WALDEN a Andrew POHLMANN. *Marketing and neuroscience what drives customer decisions.:* American Marketing Association. (2011).
- [6] PETTY, Silvia. *Neurofyziologické metody a techniky v marketingovém výskume v České republice*. Praha, 2016. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta.
- [7] ULLRICH, Petr. *Analýza a klasifikace dat ze snímače mozkové aktivity*. Brno, 2019. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií.
- [8] HARATI, A., S. LOPEZ, I. OBEID, J. PICONE, M. P. JACOBSON a S. TOBOCHNIK. The TUH EEG CORPUS: A big data resource for automated EEG interpretation. *2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2014, 1-5. ISBN 978-1-4799-8184-7. Dostupné z: doi:10.1109/SPMB.2014.7002953
- [9] CHÁBEROVÁ, Jana. *Automatická detekce artefaktů v EEG*. Praha, 2014. Bakalářská práce. České vysoké učení technické v Praze, Fakulta elektrotechnická.

- [10] KALA, David. *Metody a parametry adaptivní segmentace EEG záznamů*. Kladno, 2016. Diplomová práce. České vysoké učení technické v Praze, Fakulta biomedicínského inženýrství.
- [11] PIORECKÝ, Marek. *Analýza dat simultánního nahrávání EEG a fMRI*. Kladno, 2020. Disertační práce. České vysoké učení technické v Praze, Fakulta biomedicínského inženýrství.
- [12] ŠEBEK, Jakub. *Funkční konektivita zdrojů EEG pro časově vázané události*. Kladno, 2018. Diplomová práce. České vysoké učení technické v Praze, Fakulta biomedicínského inženýrství.
- [13] CHALOUPEK, Jaromír. *Detekce EMG artefaktu v reálném EEG záznamu za pomoci metody na bázi frekvenčního a časového prahu*. Kladno, 2020. Bakalářská práce. České vysoké učení technické v Praze, Fakulta biomedicínského inženýrství.
- [14] CHANG, B. S., S. C. SCHACHTER a D. L. SCHOMER. *Atlas of ambulatory EEG*. Amsterdam: Elsevier Academic Press. 2005. ISBN 9780126213454.
- [15] HODULÍKOVÁ, Tereza. *Analýza EEG během anestezie*. Brno, 2015. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií.
- [16] JANČOVÁ, Ivana. *Automatická detekce grafoelementů v signálu EEG*. Brno, 2013. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií.
- [17] *The Temple University Hospital EEG Corpus: Annotation Guidelines* [online]. 2020 [cit. 2023-05-16]. Dostupné z: www.isip.piconepress.com/publications/reports/2020/tuh_eeg/annotations/
- [18] TEPLAN, Michal. Fundamentals of EEG measurement. *Measurement science review*. 2002.
- [19] BUCKWALTER, G., S. CHHIN, S. RAHMAN, I. OBEID a J. PICONE. Recent Advances in the TUH EEG Corpus: Improving the Interrater Agreement for Artifacts and Epileptiform Events. *2021 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2021, 1-3. ISBN 978-1-6654-2897-2. Dostupné z: doi:10.1109/SPMB52430.2021.9672302
- [20] KLASS, Donald. The Continuing Challenge of Artifacts in the EEG. *American Journal of EEG Technology*. 1995, (35), 239-269. Dostupné z: doi:10.1080/00029238.1995.11080524

- [21] GARCES CORREA, M. Agustina a Eric LACIAR. Noise Removal from EEG Signals in Polisomnographic Records Applying Adaptive Filters in Cascade. *Adaptive Filtering Applications*. InTech, 2011. ISBN 978-953-307-306-4. Dostupné z: doi:10.5772/17219
- [22] KLUS, Michal. *Měření a analýza elektroencefalogramu*. Brno, 2011. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií.
- [23] STŘIHAVKOVÁ, Eva. *Snímání a hodnocení EEG*. Brno, 2012. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií.
- [24] STOYELL, Sally M., Janina WILMSKOETTER, Mary-Ann DOBROTA et al. High-Density EEG in Current Clinical Practice and Opportunities for the Future. *Journal of Clinical Neurophysiology*. 2021, **38**(2), 112-123. ISSN 0736-0258. Dostupné z: doi:10.1097/WNP.0000000000000807
- [25] POLANSKÝ, Štěpán. *ZPRACOVÁNÍ ELEKTROENCEFALOGRAFICKÝCH SIGNÁLŮ*. Brno, 2011. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií.
- [26] FERRELL, S., M. REFFORD, T. AHSAN, I. OBEID a J. PICONE. *The Temple University Hospital EEG Corpus: Electrode Location and Channel Labels* [online]. In: . [cit. 2023-05-16]. Dostupné z: https://www.isip.piconepress.com/publications/reports/2020/tuh_eeg/electrodes
- [27] GORGOLEWSKI, Krzysztof J., Tibor AUER, Vince D. CALHOUN et al. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*. 2016, **3**(1). ISSN 2052-4463. Dostupné z: doi:10.1038/sdata.2016.44
- [28] KEMP, Bob a Jesus OLIVAN. European data format 'plus' (EDF), an EDF alike standard format for the exchange of physiological data. *Clinical Neurophysiology*. 2003, **114**(9), 1755-1761. ISSN 13882457. Dostupné z: doi:10.1016/S1388-2457(03)00123-8
- [29] KRUPKA, Ondřej. *AUTOMATICKÁ SEGMENTACE DAT EEG*. Brno, 2013. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií.
- [30] PRAETORIUS, H.M, G BODENSTEIN a O.D CREUTZFELDT. Adaptive segmentation of EEG records: a new approach to automatic EEG analysis.

Electroencephalography and Clinical Neurophysiology. 1977, **42**(1), 84-94. ISSN 00134694. Dostupné z: doi:10.1016/0013-4694(77)90153-5

- [31] BODENSTEIN, G. a H.M. PRAETORIUS. Feature extraction from the electroencephalogram by adaptive segmentation. *Proceedings of the IEEE*. 1977, **65**(5), 642-652. ISSN 0018-9219. Dostupné z: doi:10.1109/PROC.1977.10543
- [32] AGARWAL, R. a J. GOTMAN. Adaptive segmentation of electroencephalographic data using a nonlinear energy operator. *ISCAS'99. Proceedings of the 1999 IEEE International Symposium on Circuits and Systems VLSI (Cat. No.99CH36349)*. IEEE, 1999, 199-202. ISBN 0-7803-5471-0. Dostupné z: doi:10.1109/ISCAS.1999.779976
- [33] KRAJČA, Vladimír, Svojmil PETRÁNEK, Ivana PATÁKOVÁ a Alpo VÄRRI. Automatic identification of significant graphoelements in multichannel EEG recordings by adaptive segmentation and fuzzy clustering. *International Journal of Bio-Medical Computing* [online]. 1991, 71-89 [cit. 2023-05-06]. ISSN 0020-7101. Dostupné z: <https://www.sciencedirect.com/science/article/pii/002071019190028D>
- [34] HASSANPOUR, H. a M. SHAHIRI. Adaptive Segmentation Using Wavelet Transform. *2007 International Conference on Electrical Engineering*. IEEE, 2007, 1-5. ISBN 1-4244-0892-X. Dostupné z: doi:10.1109/ICEE.2007.4287348
- [35] HORNIÁK, Robin. *Analýza EEG signálu*. Praha, 2010. Diplomová práce. České vysoké učení technické v Praze, Fakulta elektrotechnická.
- [36] PIORECKÝ, Marek. *Automatická klasifikace EEG segmentů metodou DBSCAN*. Kladno, 2016. Diplomová práce. České vysoké učení technické v Praze, Fakulta biomedicínského inženýrství.
- [37] CRAIK, Alexander, Yongtian HE a Jose L CONTRERAS-VIDAL. Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of Neural Engineering*. 2019, **16**(3). ISSN 1741-2560. Dostupné z: doi:10.1088/1741-2552/ab0ab5
- [38] ZACHA, Jiří. *Konvoluční neuronové síte pro klasifikaci objektu z LiDARových dat*. Praha, 2019. Bakalářská práce. České vysoké učení technické v Praze, Fakulta elektrotechnická.
- [39] KLADIVO, Petr. *Základy statistiky* [online]. Olomouc, 2013 [cit. 2023-05-16]. ISBN 978-80-244-3842-9.

- [40] KAISER, Tomáš. *Analýza epileptických EEG signálů*. Praha, 2012. Bakalářská práce. České vysoké učení technické v Praze, Fakulta elektrotechnická.
- [41] MIKLENDÁ, Petr. *SPEKTRÁLNÍ ANALÝZA EEG SIGNÁLU*. Brno, 2014. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií.
- [42] MONGE-ÁLVAREZ, Jesús. Relative Power (RP) of a signal. In: *MathWorks* [online]. [cit. 2023-05-16]. Dostupné z: <https://www.mathworks.com/matlabcentral/fileexchange/50294-relative-power-rp-of-a-signal>
- [43] 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
- [44] FISCHER, Jakub, Věra JEŘÁBKOVÁ, Ludmila PETKOVÁ, Veronika PTÁČKOVÁ a Petra ŠVARCOVÁ. *Základní metody statistického srovnávání*. Praha: Nakladatelství VŠE, 2019. ISBN 978-80-245-2342-2.
- [45] SHAJIHAN, Nimmisha. *Classification of stages of Diabetic Retinopathy using Deep Learning*. Dostupné z: doi:10.13140/RG.2.2.10503.62883
- [46] KLADOS, Manousos A. a Panagiotis D. BAMIDIS. A semi-simulated EEG/EOG dataset for the comparison of EOG artifact rejection techniques. *Data in Brief*. 2016, **8**, 1004-1006. ISSN 23523409. Dostupné z: doi:10.1016/j.dib.2016.06.032
- [47] World Medical Association Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects. In: *JAMA* [online]. 2013 [cit. 2023-05-16]. Dostupné z: <https://doi.org/10.1001/jama.2013.281053>
- [48] HAMID, A., K. GAGLIANO, S. RAHMAN, N. TULIN, V. TCHIONG, I. OBEID a J. PICONE. The Temple University Artifact Corpus: An Annotated Corpus of EEG Artifacts. *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*. IEEE, 2020, 1-4. ISBN 978-1-7281-8820-1. Dostupné z: doi:10.1109/SPMB50085.2020.9353647
- [49] MALEKZADEH, Anis, Assef ZARE, Mahdi YAGHOUBI, Hamid-Reza KOBRAVI a Roohallah ALIZADEHSANI. Epileptic Seizures Detection in EEG

Signals Using Fusion Handcrafted and Deep Learning Features. *Sensors*. 2021, 21(22). ISSN 1424-8220. Dostupné z: doi:10.3390/s21227710

- [50] EVANS, O. G. Electroencephalography. In: *Simply Psychology* [online]. 2023 [cit. 2023-05-16]. Dostupné z: <https://www.simplypsychology.org/what-is-an-eeeg-test-and-what-is-it-used-for.html>
- [51] ABU-FARAJ, Ziad. *Is the Term Paradoxical Sleep a Misnomer?*. 2020.
- [52] ŠČURKO, Lukáš. *Standardized data formats for electrophysiological data*. Plzeň, 2021. Diplomová práce. Západočeská univerzita v Plzni, Fakulta aplikovaných věd.

Příloha A: Obsah příložených souborů

Všechny vytvořené skripty a funkce jsou popsány v jejich hlavičkách a umístěny ve složce „*Skripty a Funkce*“.

Příložené skripty:

- Skript_SIM_SYNT – skript používaný pro testování metody na simulovaných syntetických datech
- Skript_SIM_SEMISYNT – skript používaný pro testování metody na simulovaných semisyntetických datech
- Skript_SIM_REAL – skript používaný pro testování metody na simulovaných reálných datech
- Skript_REAL_ANOT – skript používaný pro testování metody na anotovaných reálných datech z databáze TUAR

Příložené funkce:

- Funkce *selektivni_prah_vypocet* – funkce slouží pro výpočet selektivního prahu
- Funkce *segmentace* – funkce slouží k segmentaci dat
- Funkce *selekce_segmentu* – funkce slouží k výběru segmentů, které nepřekročily selektivní práh
- Funkce *vypocet_parametru_segmentu* – funkce slouží k výpočtu sledovaných parametrů v segmentu
- Funkce *etalon_NEP* – funkce slouží k výpočtu NEP etalonu
- Funkce *NEP_rozhodnuti* – funkce slouží k porovnání parametrického otisku analyzovaného segmentu s NEP etalonem
- Funkce *NEP_casova_osa_vypocet* – funkce slouží ke zjištění časových indexů segmentů označených jako NEP a vytvoření binární klasifikace každého vzorku záznamu
- Funkce *statistika* – funkce slouží ke statistickému hodnocení s využitím TP, TN, FP, FN, senzitivity, specificity, PPV a NPV na reálných anotovaných datech
- Funkce *statistika_simulace* – funkce slouží ke statistickému hodnocení s využitím TP, TN, FP, FN, senzitivity, specificity, PPV a NPV na simulovaných datech
- Funkce *Compute_RP* – funkce slouží k výpočtu relativního výkonu ve frekvenčních pásmech [42]

Příloha B: Obsah přiloženého CD/DVD

- Elektronická verze diplomové práce
- Zadání práce
- Abstrakt česky
- Abstrakt anglicky
- Klíčová slova česky
- Klíčová slova anglicky
- Skripty a funkce vytvořené v programovém prostředí Matlab 2022a