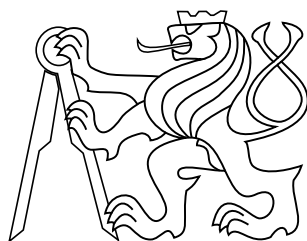


diplomová práce

Analýza EEG dětských pacientů se srdečními rhabdomyomy

Petr Štěpánek



Květen 2015

Ing. Petr Ježdík Ph.D.

České vysoké učení technické v Praze
Fakulta elektrotechnická, Katedra řídicí techniky

České vysoké učení technické v Praze
Fakulta elektrotechnická

katedra řídicí techniky

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: **Bc. Petr Štěpánek**

Studijní program: Kybernetika a robotika
Obor: Systémy a řízení

Název tématu: **Analýza EEG dětských pacientů se srdečními rhabdomyomy**

Pokyny pro vypracování:

1. Na předložených záznamech neinvazivních EEG pacientů proveďte analýzu četnosti epileptoformních výbojů vzhledem k času.
2. Optimalizujte předložený klastrovací algoritmus pro použití na tato data.
3. Proveďte analýzu prostorového výskytu výbojů vzhledem k polohám tuberózních skleróz dle vyšetření pacientů magnetickou rezonancí.
4. Formulujte a otestujte hypotézy o použitelnosti užitých výpočetních metod kvantitativního hodnocení EEG na těchto datech.

Seznam odborné literatury:

- [1] Brázdil M, Marusic P. a kol. Epilepsie temporálního laloku, TRITON 2006
- [2] Faber J. QEEG, GALÉN UK v Praze, 2005
- [3] Polanský Š, Zpracování Elektroencefalografických Signálů, Diplomová práce, VUT Brno, 2001

Vedoucí: Ing. Petr Ježdík, Ph.D.

Platnost zadání: do konce letního semestru 2015/2016



V Praze dne 12. 12. 2014

Poděkování

Chtěl bych touto formou poděkovat svému vedoucímu práce Ing. Petru Ježdíkovi, Ph.D za jeho čas a mnohé rady, které mi během posledního roku věnoval. Dále bych chtěl poděkovat 2. lékařské fakultě Univerzity Karlovy a Fakultní nemocnici v Motole za poskytnutá data, bez kterých by práce nebyla možná. Nakonec bych chtěl poděkovat své rodině za jejich podporu během celých pěti let mého studia na Českém vysokém učení technickém v Praze.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval samostatně, a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze, dne10.5.2015.....

..........

Podpis

Abstrakt

Studie se zabývá pacienty, kteří byli již v prenatálním stádiu diagnostikováni se srdečnými rhabdomyomy a později se u nich rozvinula epilepsie. EEG pacientů trpících epilepsií je charakteristické přítomností patologických artefaktů zvaných interiktální epileptiformní výboje (IED). Detekce IED v invazivním EEG je možné využít při lokalizaci epileptických zón, potřebných při plánování chirurgických zákroků. Cílem diplomové práce je prozkoumat využitelnost analýzy neinvazivního EEG při lokalizaci epileptických zón. Poskytnutá data obsahovala EEG záznamy pacientů a ohodnocená vyšetření magnetickou rezonancí (MRI). Polohy tuberózních skleróz určené vyšetřením MRI jsou za pomoci navrženého modelu transformovány na MRI profil pacienta. EEG data jsou jednak zpracována navrženým automatickým postupem, a také dodaným shlukovým algoritmem. Výsledky analýzy EEG jsou porovnány s referencí definovanou MRI profilem. Kvalita navrženého postupu je popsána senzitivitou, specificitou, prediktivní hodnotou pozitivního testu a přesností.

Klíčová slova

EEG, srdeční rhabdomyomy, epilepsie, epileptiformní výboj.

Abstract

Study focuses on epileptic patients diagnosed with cardiac rhabdomyoma during prenatal phase. EEG of patients with epilepsy is characterized by pathologic artifacts known as interictal epileptiform discharges (IEDs). It is possible to utilize detections of IEDs in intracranial EEG in localization and planning of surgical resection. Aim of this master thesis is to show whether it is feasible to use noninvasive EEG instead. Available data consists of EEG recordings of patients suffering from epilepsy and their magnetic resonance imaging (MRI) examinations. Localizations of tuberous sclerosis are translated into MRI profile defined by proposed model. EEG recordings of patients are then analyzed by proposed algorithm as well as by given clustering algorithm. Results of EEG analysis are compared with MRI reference. Quality of proposed algorithm is described by sensitivity, specificity, positive predictive value and accuracy.

Keywords

EEG, cardiac rhabdomyoma, epilepsy, epileptiform discharge.

Obsah

1 Úvod	1
2 Data pacientů	3
2.1 EEG data pacientů	3
2.1.1 Mezinárodní standard 10/20	4
2.1.2 Rhabdomyomy a jejich vztah k epilepsii	5
2.2 Data z magnetické rezonance	5
2.2.1 Definice kódu pro označení poloh tuberózních skleróz	6
3 Metodika zpracování dat	8
3.1 Konstrukce modelu vycházejícího z magnetických rezonancí	8
3.1.1 Zobrazení dat magnetických rezonancí na skupiny elektrod	10
3.1.2 Zavedení souřadného systému pro používané elektrody	10
3.1.3 Vícerozměrné normální rozdělení	11
3.1.4 Ohodnocení elektrod na základě vstupních dat	12
3.2 Četnostní analýza epileptoformních výbojů	12
3.2.1 Epileptoformní výboj	12
3.3 Analýza EEG dat	13
3.3.1 Analýza hlavních komponent	14
3.3.2 Určení chybovosti detektoru z dat zdravých novorozenců	14
3.3.3 Zpracování hlavní komponenty a určení váženého EEG profilu pacienta	15
3.4 Klasifikace dat vůči kvadrantům a lalokům	16
3.5 Statistické zpracování získaných dat	17
3.5.1 Hodnocení diagnostických testů	17
3.5.2 ROC křivky	18
4 Výsledky	19
4.1 Chybovost detektoru	19
4.1.1 Váhy určené z dat zdravých novorozenců w_j^I	19
4.2 Porovnání modelu s manuální klasifikací	20
4.2.1 První vyšetření magnetickou rezonancí	20
4.2.2 Druhé vyšetření magnetickou rezonancí	21
4.3 Váhy w_j^{II} EEG dat pacientů	22
4.3.1 Pacient 1	22
4.3.2 Pacient 2	22
4.3.3 Pacient 3	22

4.3.4	Pacient 4	22
4.3.5	Pacient 5	23
4.3.6	Pacient 6	23
4.3.7	Pacient 7	23
4.3.8	Pacient 8	24
4.3.9	Pacient 9	24
4.4	ROC křivka jako funkce parametru k	26
4.5	Kvalita klasifikace dat získaných shlukovou analýzou	28
4.5.1	První vyšetření magnetickou rezonancí	28
4.5.2	Druhé vyšetření magnetickou rezonancí	29
5	Diskuze	32
	Přílohy	
A	Appendix	35
	Literatura	36

Zkratky

Seznam použitých zkratek:

EEG	Elektroencefalogram
MRI	Magnetická rezonance
PCA	Analýza hlavních komponent
ISARG	Intracranial Signal Analysis Reasearch Group
IEDs	Interiktální epileptoformní výboje
SEN	Senzitivita
SPEC	Specificita
PPV	Prediktivní hodnota pozitivního testu
ACC	Přesnost
ROC	Reciever Operator Characteristic
FPR	Míra falešně pozitivních rozhodnutí

1 Úvod

Epilepsii je možné léčit, mimo jiné, pomocí léků. V případě farmakorezistivní epilepsie je však nutné hledat alternativní způsoby léčby. Tím může být například chirurgický zákrok, který se snaží odstranit epileptická ložiska. Lokalizaci ložisek je možné napomoci analýzou invazivního EEG, ve kterém je z prostorového výskytu interiktálních epileptiformních výbojů možné usoudit o poloze ložisek. Interiktální epileptiformní výboj je patologický artefakt EEG vyskytující se mezi záchvaty.

Cílem práce je ověřit zda je možné stejným způsobem využít neinvazivní EEG. Invazivní EEG je z hlediska kvality signálu nesrovnatelně lepší než neinvazivní EEG. Na druhou stranu měření neinvazivního EEG je jednodušší a nevyžaduje chirurgický zákrok. K analýze neinvazivního EEG byly poskytnuty detektor a shlukový algoritmus výzkumné skupiny ISARG. Tyto nástroje byly navrženy a optimalizovány pro použití na invazivním EEG.

Druhá kapitola se zabývá podrobným popisem dostupných dat. Data je možné rozdělit do tří skupin. První skupinu tvoří EEG záznamy zkoumaných pacientů, kterých bylo analyzováno celkem 9. Do druhé skupiny spadají EEG záznamy zdravých novorozenců. Zdravých novorozenců bylo 7. Poslední skupinu dat tvoří ohodnocená měření magnetickou rezonancí. Každý pacient podstoupil jedno až dvě MRI vyšetření.

Třetí kapitola popisuje navržený postup zpracování dat a způsob porovnání s referenčním modelem. Je zde popsán návrh modelu převádějící kódy vyšetření magnetickou rezonancí na ohodnocení jednotlivých elektrod zapojení. Obě skupiny EEG dat jsou zpracovány postupem popsáním v kapitole 3.3. EEG data pacientů jsou navíc analyzována shlukovým algoritmem. V závěru kapitoly je osvětleno jakým způsobem jsou výsledky analýzy EEG porovnávány s referencí z MRI.

Čtvrtá kapitola uvádí přehled získaných výsledků. Jako první jsou zkoumána EEG data zdravých novorozenců. Tato data slouží k určení chybovosti detektoru a k určení obecných vlastností jednotlivých kanálů neinvazivního EEG novorozenců. Následně jsou ověřeny kvalitativní vlastnosti navrženého modelu MRI porovnáním s manuální klasifikací. Po rozdělení dat pacientů do skupin příslušících některému vyšetření MRI, jsou pro tyto skupiny určeny váhy. Kombinací vah určených z dat zdravých novorozenců a vah vytvořených skupin, vzniká váha $w = w(k)$. Kvalita navrženého postupu je otestována pro několik hodnot parametru k . Na závěr je porovnána klasifikace na základě shlukové analýzy s referenčním modelem. Výsledky kapitoly 4.2 a 4.5 jsou uvedeny ve

1 Úvod

formátu (*Střední_hodnota* \pm *Směrodatná_odchylka*).

Poslední kapitola se zabývá analýzou a shrnutím výsledků. Z ROC křivek jsou určeny optimální hodnoty parametru k nastavujícího poměr zastoupení složek v celkové váze. Pro optimální hodnoty parametru k jsou určeny senzitivita a specificita klasifikace automatického zpracování dat. Senzitivita a specificita automatického zpracování jsou porovnány s odpovídajícími statistikami shlukové analýzy. Závěr kapitoly je věnován návrhům na vylepšení postupu.

2 Data pacientů

Veškerá data použitá v rámci diplomové práce byla poskytnuta 2. lékařskou fakultou Univerzity Karlovy a Fakultní nemocnicí v Motole. Anonymizovaná data zahrnují EEG záznamy pacientů trpících epilepsií a pro ně určené polohy tuberózních skleróz vyšetřením magnetickou rezonancí. Dále byly k dispozici EEG záznamy zdravých novorozenců.

Zpracováno bylo celkem devět pacientů. Tři z nich byly dívky a zbytek chlapci. Každý z nich měl alespoň pět naměřených EEG záznamů a jedno vyšetření magnetickou rezonancí. Stáří pacientů v průběhu měření EEG se pohybovalo v rozmezí od čtyř dnů po narození až do sedmi let věku. První vyšetření magnetickou rezonancí byla typicky provedena před dosažením druhého roku věku. Druhé vyšetření ve věku od dvou do sedmi let.

Zdravých novorozenců bylo sedm. Jeden z nich měl dva naměřené EEG záznamy, ostatní pouze jeden. Informace o pohlaví zdravých novorozenců nebyla k dispozici a tak jako v případě pacientů trpících epilepsií i tato data byla anonymizována.

2.1 EEG data pacientů

Záznamy pacientů i zdravých novorozenců byly měřeny pomocí povrchových elektrod. Měřících elektrod bylo 23 a signál byl vzorkován na frekvenci 200Hz . V některých případech nebylo zapojeno všech 23 elektrod. Z celkového počtu elektrod byly 4 elektrody pomocné:

A1, A2 Referenční elektrody na ušních boltcích. Elektroda A1 na levém uchu a elektroda A2 na pravém.

TIMESYNC Časová synchronizace elektrod.

Poly1 Kanál pro měření dodatečných signálů nemajících původ v EEG.

Rozmístění a pojmenování zbylých měřících elektrod se řídilo podle mezinárodního standardu 10/20. Délka EEG záznamů se v průměru pohybovala okolo dvaceti minut.

Počet EEG záznamů a základní údaje o pacientech jsou uvedeny v Tabulce 1. V případě *Pacienta 7* byl vyřazen jeden záznam z důvodu zhoršené kvality. Tabulka 2 obsahuje přehled informací pro zdravé novorozence. Záznamy *Novorozence 3* a *Novorozence 4* byly vyřazeny ze zpracování z důvodu zhoršené kvality.

	Sex	Věk	Total	Před	Po
Pacient 1	F	25	15	2	13
Pacient 2	F	29	18	4	14
Pacient 3	M	17	11	4	7
Pacient 4	M	31	8	3	5
Pacient 5	M	6	7	1	6
Pacient 6	M	11	9	2	7
Pacient 7	M	24	14	2	12
Pacient 8	M	78	8	2	6
Pacient 9	F	108	5	0	5

Tabulka 1 Přehled základních informací o pacientech. *Věk* pacientů při rozvoji epilepsie je uveden v týdnech. Sloupec *Total* zachycuje celkový počet EEG záznamů, sloupce *Před* a *Po* pak udávají počet EEG záznamů před a po rozvoji epilepsie.

	Věk	Total
Novorozenec 1	55	1
Novorozenec 2	29, 48	2
Novorozenec 3	27	1
Novorozenec 4	27	1
Novorozenec 5	40	1
Novorozenec 6	45	1
Novorozenec 7	19	1

Tabulka 2 Přehled základních informací o zdravých novorozencích. *Věk* novorozenců v době měření je uveden v týdnech. Sloupec *Total* udává počet EEG záznamů příslušných novorozenců.

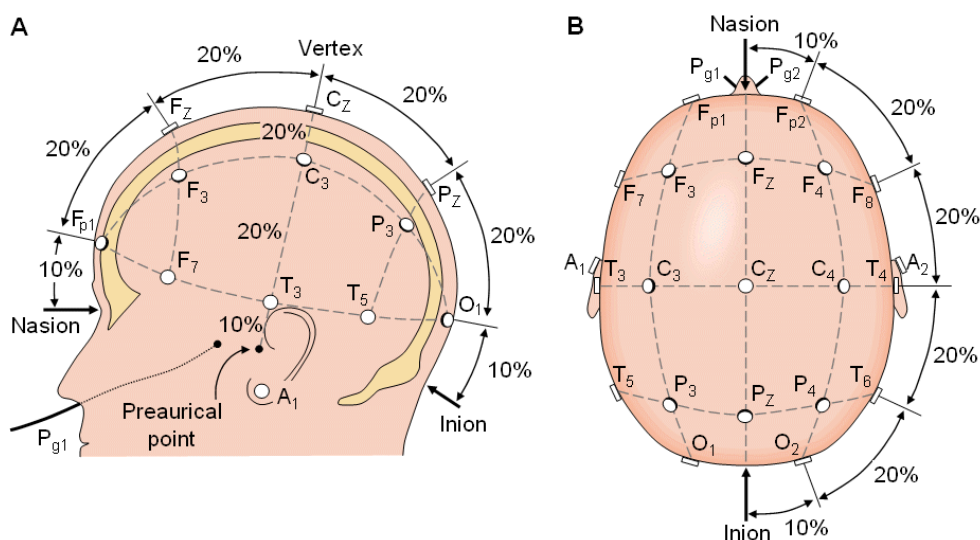
2.1.1 Mezinárodní standard 10/20

Mezinárodní standard 10/20 popisuje umístění a pojmenování elektrod při měření skalpového EEG. Čísla 10 a 20 symbolizují vzdálenost mezi dvěma sousedními elektrodami vyjádřenou v procentech celkové vzdálenosti zleva doprava nebo zepředu dozadu. Jako referenční body pro umístění elektrod se používají:

Inion Nejnižší bod na zadní straně lebky indikován výstupkem.

Nasion Bod v prohlubni mezi očima nad nosním můstkem.

Elektrody nacházející se nad levou hemisférou jsou indexovány lichými čísly, elektrody vpravo pak čísly sudými. Znaky používané pro označení elektrod mají následující význam:



Obrázek 1 Rozmístění elektrod na lebce pro standardní zapojení 10/20.

F Frontální lalok.

Fp Frontopolární oblast.

C Centrální oblast.

T Temporální lalok.

P Parietální lalok.

O Okcipitopolární lalok.

Rozmístění měřících elektrod podle mezinárodního standardu 10/20 je uvedeno na Obrázku 1.

2.1.2 Rhabdomyomy a jejich vztah k epilepsii

Rhabdomyom je nezhoubný nádor, který se může vytvořit na srdečním svalu, jeho výskyt ale není omezen pouze na srdce. Jeho přítomnost je možné diagnostikovat již ve 22. týdnu těhotenství. Studie ukazují [1], že až v 60% případů, kdy je plod diagnostikován se srdečním rhabdomyomem, dojde u těchto novorozenců k výskytu tuberózních skleróz. Jením z nejčastějších projevů tuberózních skleróz je epilepsie, a to u 80 – 90% pacientů [2].

Pacienti, kteří jsou diagnostikováni se srdečními rhabdomyomy, jsou již od narození pravidelně monitorováni. Data analyzovaná touto prací jsou daty právě těchto pacientů.

2.2 Data z magnetické rezonance

Každý pacient prošel alespoň jedním vyšetřením magnetickou rezonancí. První vyšetření bylo ve většině případů provedeno před dosažením druhého roku věku. Druhé vyšetření ve věku od dvou do šesti let. Data z vyšetření magnetickou rezonancí byla ohod-

	MRI 1		MRI 2	
Pacient 1	6,3	FPR, PLR	70	FPR, PLR
Pacient 2	5,6	FBR	23,2	FML, FPL
Pacient 3	3,6	FLR	29,4	FPR
Pacient 4	24,2	FBL	—	—
Pacient 5	1,6	FCR, FCoL	12,4	Chybí
Pacient 6	2,2	TLR	—	—
Pacient 7	3	PMR	32,6	OLL
Pacient 8	27,7	N/A	66,4	FCoR
Pacient 9	6,3	FCoL	82,2	FCoL

Tabulka 3 Přehled základních informací o datech z magnetických rezonancí. První sloupce vždy vyjadřují stáří pacienta v průběhu měření v měsících. Druhé sloupce obsahují polohy tuberózních skleróz.

nocena expertem v oboru a výsledek byl vyjádřen ve formě kódu definujícího polohu nalezených tuberózních skleróz.

Informace o datech z magnetických rezonancí jsou shrnuty v Tabulce 3. *Pacient 4* a *6* neabsolvovali druhé vyšetření magnetickou rezonancí. Druhé měření magnetickou rezonancí u *Pacienta 5* nebylo vyhodnoceno. V prvním měření *Pacienta 8* nenalezl expert žádné tuberózní sklerózy.

2.2.1 Definice kódu pro označení poloh tuberózních skleróz

První znak kódu určuje polohu tuberózních skleróz vůči mozkovým lalokům. Druhý znak přesněji specifikuje polohu uvnitř daného laloku a třetí znak označuje hemisféru, ve které se nález nachází.

První znak kódu:

- F** Frontal (Frontální lalok).
- T** Temporal (Temporální lalok).
- P** Parietal (Parietální lalok).
- O** Occipital (Okcipitopolární lalok).

Druhý znak kódu:

- C** Central.
- M** Mesial.
- L** Lateral.
- Co** Convex.

P Polar.

B Base.

Třetí znak kódu:

L Left (Levá hemisféra).

R Right (Pravá hemisféra).

Nakonec ještě poznamenejme, že za znakem frontálního laloku mohou následovat znaky **C**, **M**, **Co**, **P** a **B**. U ostatních laloků to jsou pouze znaky **M** a **L**.

3 Metodika zpracování dat

Cílem práce je ověřit zda je možné vhodnou analýzou EEG dat dospět ke stejným výsledkům jako v případě vyšetření magnetickou rezonancí. Jelikož byla data z magnetické rezonance ve formě kódu, musel být navržen model, který tento kód převede do formy vhodné pro porovnání s výsledky získanými analýzou EEG dat.

Soubory EEG záznamů přiřazené k příslušnému vyšetření magnetickou rezonancí jsou analyzovány za pomoci ISARG detektoru [3]. Výsledné četnosti epileptiformních výbojů na kanál jsou vyšetřeny analýzou hlavních komponent. Hlavní komponenta získaná tímto postupem je považována za EEG profil pacienta.

Data zdravých novorozenců jsou využívána k určení chybovosti detektoru. Na základě určené chybovosti detektoru jsou váženy výsledky analýzy EEG dat pacientů. Zpracování souboru EEG záznamů zdravých novorozenců je shodné jako v případě souborů epileptických pacientů. Zde existuje pouze jeden soubor EEG dat, který definuje chybovost detektoru pro všechna EEG data.

Extrahovaná hlavní komponenta EEG dat pacientů je modifikována a vážena získanou chybovostí detektoru. Tyto vážené profily a MRI profily získané z modelu jsou klasifikovány vůči kvadrantům nebo mozkovým lalokům. Přesnost klasifikace je nakonec statisticky zpracována.

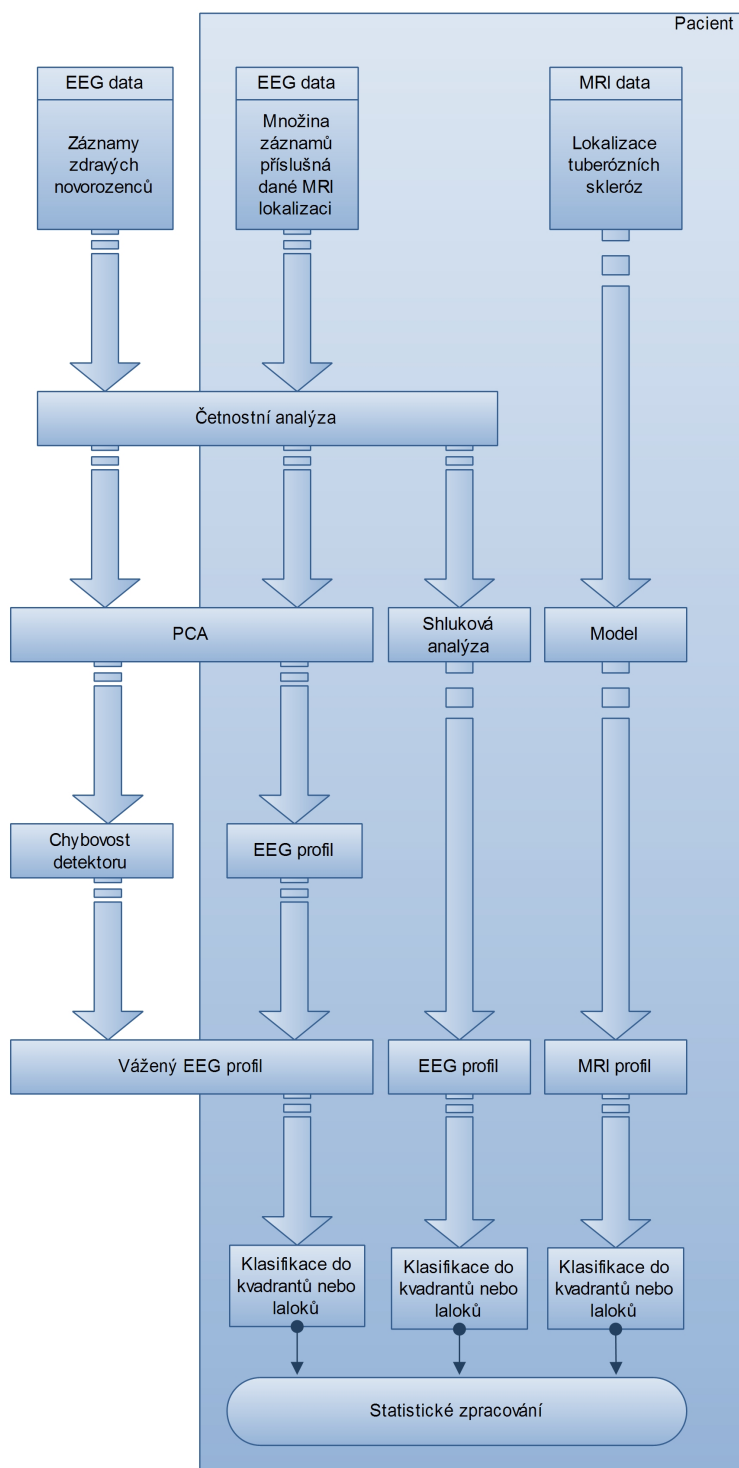
Blokové schéma navrženého postupu je uvedeno na Obrázku 2.

3.1 Konstrukce modelu vycházejícího z magnetických rezonancí

Pokud pacient absolvoval měření a měřením byla určena poloha tuberózních skleróz, výsledek bude reprezentován ve formě jednoho nebo více kódů. Definice a význam těchto kódů je uveden v kapitole 2.2.1.

Model se skládá ze souřadného systému definující polohu jednotlivých elektrod v dvojrozměrném prostoru a z přiřazení, které každému kódu přiřadí skupinu elektrod. Pro tuto skupinu elektrod je vypočítán její hmotný střed a do něj je umístěna dvourozměrná Gaussova funkce. Profil pacienta je dán jako součet příspěvků jednotlivých Gaussových funkcí, pokud má pacient danou více než jednu polohu tuberózních skleróz, v bodech definovaných jako souřadnice elektrod.

3.1 Konstrukce modelu vycházejícího z magnetických rezonancí



Obrázek 2 Blokový diagram navrženého postupu zpracování a porovnání dat.

LEFT	C	Co	P	B	M	L
F	Fp1, F3, C3, F7, Fz, Cz					—
T	—				C3	T3, T5
P	—				P3, Cz, Pz	C3, P3
O	—				O1	O1

Tabulka 4 Zobrazení přiřazující skupinu elektrod kódům pro levou hemisféru. Řádky odpovídají lalokům a sloupce jejich jemnějšímu dělení.

RIGHT	C	Co	P	B	M	L
F	Fp2, F4, C4, F8, Fz, Cz					—
T	—				C4	T4, T6
P	—				P4, Cz, Pz	C4, P4
O	—				O2	O2

Tabulka 5 Zobrazení přiřazující skupinu elektrod kódům pro pravou hemisféru. Řádky odpovídají lalokům a sloupce jejich jemnějšímu dělení.

3.1.1 Zobrazení dat magnetických rezonancí na skupiny elektrod

Výstupem z měření magnetickou rezonancí může být jeden nebo více kódů. Každý z nich odpovídá poloze tuberózních skleróz v mozku. Zobrazení z kódů na skupiny elektrod se snaží popsat, které elektrody v okolí ložiska epileptiformních výbojů, by měly vykazovat nejvyšší aktivitu.

Zobrazení pro levou hemisféru jsou definována v Tabulce 4. Zobrazení pro pravou hemisféru jsou definována v Tabulce 5.

3.1.2 Zavedení souřadného systému pro používané elektrody

Model byl parametrizován pomocí dvou parametrů. První z nich je vzdálenost mezi *Inion* a *Nasion*, definice viz. kapitola 2.1.1. Druhým parametrem je kovariance dvou-rozměrné Gaussovy funkce. Model pro zjednodušení předpokládá, že obraz lebky s rozmístěnými elektrodami v rovině je ve tvaru kružnice, jejíž poloměr je roven polovině vzdálenosti mezi *Inion* a *Nasion*. V textu je tato vzdálenost označena jako r_1 :

$$r_1 = \frac{1}{2} \| \text{Inion} - \text{Nasion} \| .$$

Počátek souřadné soustavy byl umístěn do středu výše definované kružnice. Kladná poloosa x jde ve směru ze středu lebky k pravému spánku. Kladná poloosa y jde ze středu lebky ve směru k nosu. Rozmístění elektrod v takto definované souřadné soustavě vychází z návodu na rozmístění elektrod pro případ mezinárodního zapojení 10/20 na

povrch hlavy [4].

Devadesát procent z poloměru r_1 je označeno jako r_2 a na základě této vzdálenosti jsou definovány souřadnice jednotlivých elektrod:

$$r_2 = 0,9r_1.$$

Rozmístění elektrod do souřadné soustavy:

$$\begin{aligned} \mathbf{Fp1} \quad [x_1, y_1] &= [r_2 \cos(0, 6\pi), r_2 \sin(0, 6\pi)] \\ \mathbf{Fp2} \quad [x_2, y_2] &= [r_2 \cos(0, 4\pi), r_2 \sin(0, 4\pi)] \\ \mathbf{F3} \quad [x_3, y_3] &= \left[-\frac{r_2}{4} + \frac{r_2}{2} \cos(0, 6\pi), \frac{r_2}{2} \sin(0, 8\pi) + \frac{r_2}{4}\right] \\ \mathbf{F4} \quad [x_4, y_4] &= \left[\frac{r_2}{4} - \frac{r_2}{2} \cos(0, 6\pi), \frac{r_2}{2} \sin(0, 8\pi) + \frac{r_2}{4}\right] \\ \mathbf{C3} \quad [x_5, y_5] &= \left[-\frac{r_2}{2}, 0\right] \\ \mathbf{C4} \quad [x_6, y_6] &= \left[\frac{r_2}{2}, 0\right] \\ \mathbf{P3} \quad [x_7, y_7] &= \left[-\frac{r_2}{4} + \frac{r_2}{2} \cos(0, 6\pi), -\frac{r_2}{2} \sin(0, 8\pi) - \frac{r_2}{4}\right] \\ \mathbf{P4} \quad [x_8, y_8] &= \left[\frac{r_2}{4} - \frac{r_2}{2} \cos(0, 6\pi), -\frac{r_2}{2} \sin(0, 8\pi) - \frac{r_2}{4}\right] \\ \mathbf{O1} \quad [x_9, y_9] &= [r_2 \cos(1, 4\pi), r_2 \sin(1, 4\pi)] \\ \mathbf{O2} \quad [x_{10}, y_{10}] &= [r_2 \cos(1, 6\pi), r_2 \sin(1, 6\pi)] \\ \mathbf{F7} \quad [x_{11}, y_{11}] &= [r_2 \cos(0, 8\pi), r_2 \sin(0, 8\pi)] \\ \mathbf{F8} \quad [x_{12}, y_{12}] &= [r_2 \cos(0, 2\pi), r_2 \sin(0, 2\pi)] \\ \mathbf{T3} \quad [x_{13}, y_{13}] &= [-r_2, 0] \\ \mathbf{T4} \quad [x_{14}, y_{14}] &= [r_2, 0] \\ \mathbf{T5} \quad [x_{15}, y_{15}] &= [r_2 \cos(1, 2\pi), r_2 \sin(1, 2\pi)] \\ \mathbf{T6} \quad [x_{16}, y_{16}] &= [r_2 \cos(1, 8\pi), r_2 \sin(1, 8\pi)] \\ \mathbf{Fz} \quad [x_{17}, y_{17}] &= \left[0, \frac{r_2}{2}\right] \\ \mathbf{Cz} \quad [x_{18}, y_{18}] &= [0, 0] \\ \mathbf{Pz} \quad [x_{19}, y_{19}] &= \left[0, -\frac{r_2}{2}\right] \end{aligned}$$

3.1.3 Vícerozměrné normální rozdělení

Definice: n -rozměrné normální rozdělení $N_n(\mu, V)$, kde μ je sloupcový vektor z \mathbb{R}^n a V je pozitivně definitní matice řádu n , je rozdělení s hustotou

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det V}} \exp\left(-\frac{(x - \mu)^T V^{-1} (x - \mu)}{2}\right),$$

$$\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n.$$

Parametr μ je střední hodnotou a V varianční maticí tohoto rozdělení [5].

Navržený model je konstruován pomocí dvourozměrného normálního rozdělení. Střední hodnota rozdělení je těžištěm skupiny elektrod a varianční matice je definována na základě parametrizace modelu.

3.1.4 Ohodnocení elektrod na základě vstupních dat

Vstupními daty modelu jsou kódy definující polohy tuberózních skleróz a parametrizace modelu. Postupem popsaným v kapitole 3.1.2 jsou každé elektrodě přiřazeny souřadnice v rovině. Kódy jsou převedeny na skupiny elektrod:

$$\text{kód}_i \longrightarrow \mathcal{M}_i,$$

kde index i jde přes kódy příslušné danému vyšetření magnetickou rezonancí. Pro každou množinu je určen hmotný střed:

$$\mathcal{M}_i \longrightarrow T_i.$$

Do každého hmotného bodu je umístěno dvourozměrné normální rozdělení:

$$f_i(x) = \mathcal{N}(T_i, V) = \frac{1}{2\pi\sqrt{\det V}} \exp\left(-\frac{(x - T_i)^T V^{-1} (x - T_i)}{2}\right).$$

Ohodnocení elektrod je dáno součtem příspěvků funkcí $f_i(x)$ v bodech do kterých byly elektrody umístěny:

$$w_j = \sum_{i=1}^n f_i(x_j),$$

kde x_j jsou souřadnice elektrod a index $j = 1, 2, \dots, 19$. Výstupem modelu je vektor ohodnocených elektrod, který je porovnáván s výsledky zpracování EEG záznamů jež byly přiřazeny k vyšetření magnetickou rezonancí:

$$W = (w_1, w_2, \dots, w_{19}).$$

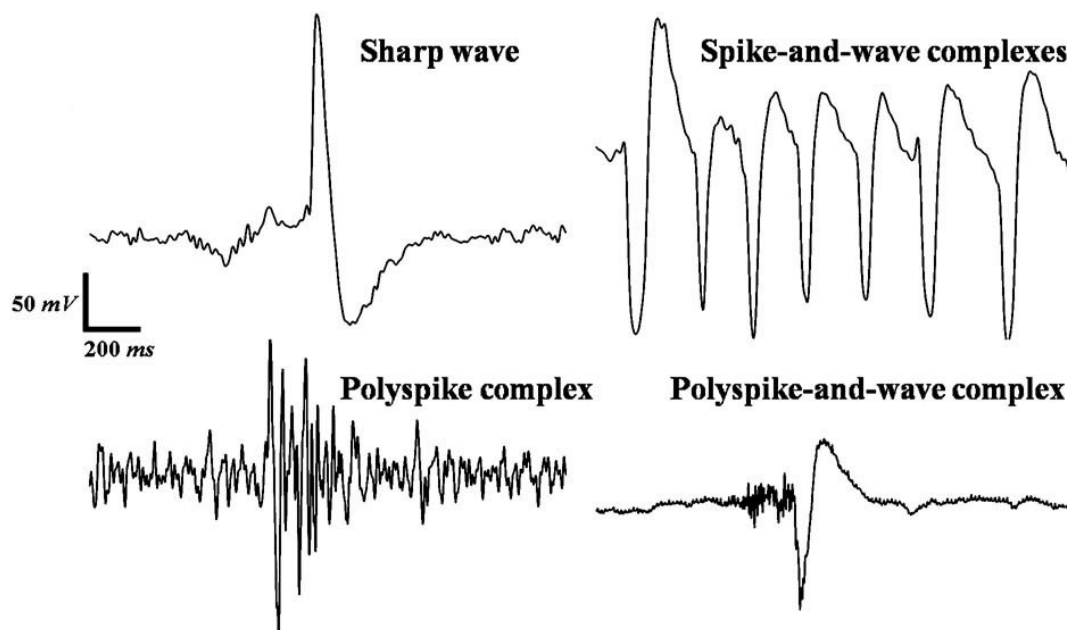
3.2 Četnostní analýza epileptiformních výbojů

Analýza četností epileptiformních výbojů vzhledem k času je prováděna za pomoci detektoru hrotů výzkumné skupiny ISARG. Poskytnutý detektor provádí jednonálovou analýzu signálu. Výstupem detektoru jsou, mimo jiné, určené indexy vzorků, časy označených epileptiformních výbojů, pro jednotlivé kanály záznamu. Ve výsledku jsou kanály ohodnoceny pomocí počtu epileptiformních výbojů za minutu.

Četnostní analýza je prováděna pro soubor EEG dat zdravých novorozenců a soubory EEG dat pacientů. Obě skupiny dat byly analyzovány pro shodné nastavení detektoru.

3.2.1 Epileptiformní výboj

U pacientů trpících epilepsií se mezi záchvaty mohou vyskytovat patologické artefakty nazývané interiktální epileptiformní výboje (IEDs). S první definicí IEDs přišel Gloor v roce 1975 [6]. Interiktální výboje lze rozdělit do čtyř skupin na základě jejich morfologie [7, 8]. Definice těchto skupin používaných v současné době:



Obrázek 3 Typické vzory epileptoformních výbojů [9].

Ostrá vlna je přechod jednoznačně odlišitelný od pozadí s ostrou špičkou a dobou trvání 70 – 200ms.

Hrot je podobný ostré vlně. Hrot trvá 20 – 70ms.

Komplex hrot-vlna je vzor skládající se z hrotu následovaného pomalou vlnou.

Mnohočetné hroty jsou komplexy hrot-vlna následující v těsné blízkosti za sebou.

Příklady interiktálních výbojů jsou uvedeny na Obrázku 3.

3.3 Analýza EEG dat

Výsledky četnostní analýzy za pomoci ISARG detektoru jsou dále zpracovávány dvěma různými způsoby.

1. Detekce jsou dále předány do shlukového algoritmu navrženého výzkumnou skupinou ISARG. Ze shluků jsou vybrány ty, které se nejvíce podobají epileptoformním výbojům a na základě nich jsou elektrody ohodnoceny.
2. Na detekce je aplikován algoritmus analýzy hlavních komponent a z dat je extrahována hlavní komponenta. Tato komponenta je vážena chybovostí detektoru získanou analýzou dat zdravých novorozenců.

Data zdravých novorozenců jsou analyzována pouze detektorem ISARG, shluková analýza není prováděna.

3.3.1 Analýza hlavních komponent

Analýza hlavních komponent (PCA) je jednoduchá neparametrická metoda extrahující relevantní data z množiny dat. Dokáže nalézt hlavní směry ve kterých se data mění a umožňuje redukci složitých množin dat na data s nižší dimenzí. Analýzou hlavních komponent je možné nalézt jednodušší skrytou dynamiku nacházející se v datech [10].

Kovarianční matice dat s nulovou střední hodnotou je matice $n \times n$ definovaná vztahem:

$$\Sigma = \frac{1}{n-1} X X^T,$$

kde matice X má v řádcích datové vektory a ve sloupcích všechna měření jednotlivých proměnných. Dále jsou definovány čtvercové matice Φ a Λ řádu $n \times n$. Matice Φ je ortogonální a její sloupce jsou vlastní vektory matice Σ . Matice Λ je diagonální s vlastními čísly matice Σ na diagonále. Pro takto definované matice platí vztah:

$$\Phi^T \Sigma \Phi = \Lambda.$$

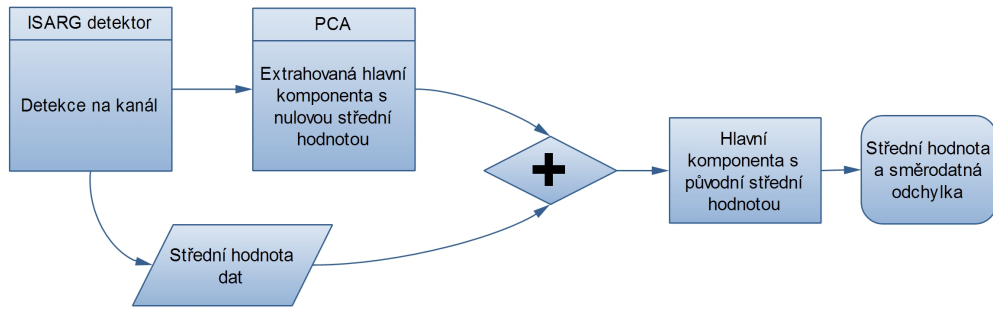
Matice Φ má roli lineární transformace, která transformuje data do nového souřadného systému v němž jsou data nekorelovaná [11].

3.3.2 Určení chybovosti detektoru z dat zdravých novorozenců

V ideálním případě by v těchto záznamech neměly být detekovány žádné epileptiformní výboje, u novorozenců může být problém s tím, jakým způsobem je epileptiformní výboj definován. Předpokládejme tedy, že veškeré detekce jsou falešně pozitivní. Detekce jsou předány ve formě matice typu $n \times 19$, kde n je počet realizací náhodného vektoru s délkou 19.

Detekce jsou následně zpracovány analýzou hlavních komponent. Při identifikaci hlavních směrů je vybráno k komponent tak, aby v nich bylo obsaženo alespoň 70% energie signálu. Jelikož funkce `pca()` v programu Matlab nejprve od vstupních dat odečte jejich střední hodnotu, je na závěr nutné střední hodnotu opět přičíst k výslednému signálu.

Extrahovaná hlavní komponenta je použita k výpočtu střední chybovosti detektoru na kanál. Směrodatná odchylka hlavní komponenty hraje v dalším zpracování roli věrohodnosti kanálu. Některé kanály mají tendenci obsahovat větší množství artefaktů, které pak způsobují falešné detekce. Kanálům s malou směrodatnou odchylkou je přiřkládána větší váha při hodnocení detekcí EEG záznamů pacientů. Popsaný postup je schematicky znázorněn na Obrázku 4.



Obrázek 4 Blokové schéma určení chybovosti detektoru.

3.3.3 Zpracování hlavní komponenty a určení váženého EEG profilu pacienta

Při zpracování získané hlavní komponenty a při konstrukci vah elektrod jsou využívány střední hodnoty a směrodatné odchylky kanálů. V navrženém postupu jsou využity obě skupiny dat, hlavní komponenta zdravých novorozenců i hlavní komponenta EEG dat pacientů. Označme střední hodnotu a směrodatnou odchylku j -tého kanálu hlavní komponenty zdravých novorozenců jako μ_j^I a σ_j^I . Obdobně pro hlavní komponentu EEG dat pacientů μ_j^{II} a σ_j^{II} .

Prvním krokem zpracování je odečtení střední hodnoty dat zdravých novorozenců od EEG dat pacientů. Pokud je rozdíl záporný, je prvek matice nahrazen nulou. Označíme-li zmíněný rozdíl jako $x_{ij} = d_{ij} - \mu_j^I$ a výslednou matici jako A pak:

$$a_{ij} = \begin{cases} x_{ij}, & \text{pokud } x_{ij} > 0. \\ 0, & \text{jinak.} \end{cases}$$

V druhém kroku jsou určeny váhy, kterými jsou ohodnoceny jednotlivé kanály. Výsledná váha je dána jako lineární kombinace vah zkonstruovaných z jednotlivých momentů skupin dat. Váha j -tého kanálu daná daty zdravých novorozenců je definovaná vztahem:

$$w_j^I = 1 - \frac{\sigma_j^I}{\mu_j^I}. \quad (1)$$

Váha určená EEG daty pacientů:

$$w_j^{II} = 1 - \frac{\sigma_j^{II}}{\mu_j^{II}}. \quad (2)$$

Výsledné váhy kanálů jsou definovány podle:

$$w_j = kw_j^I + (1 - k)w_j^{II}, \quad (3)$$

kde konstanta k je číslo z intervalu $\langle 0, 1 \rangle$. Obě dílčí váhy jsou vždy kladné, a proto je pro $k \in \langle 0, 1 \rangle$ i výsledná váha vždy kladná. Navíc pro kanál s malou směrodatnou odchylkou platí, že jeho váha bude větší, než v případě kanálu s vysokou směrodatnou odchylkou. Řádkový vektor vah je označen jako:

$$W = (w_1, w_2, \dots, w_{19}).$$

Posledním krokem navrženého postupu je vynásobení řádků výsledné matice A vektorem vah prvek po prvku. Vážená hlavní komponenta EEG dat pacientů je nyní připravena ke klasifikaci vůči kvadrantům a lalokům.

3.4 Klasifikace dat vůči kvadrantům a lalokům

Klasifikace dat se provádí u MRI profilu získaného z modelu, váženého EEG profilu získaného ze zpracování dat z detektoru a EEG profilu pacienta získaného ze shlukové analýzy. Data jsou klasifikována vůči kvadrantům a lalokům. Klasifikace MRI profilu se bere jako referenční a je porovnávána s klasifikací váženého EEG profilu a EEG profilu ze shlukové analýzy.

Definované kvadranty pro klasifikaci odpovídají polohou kvadrantům souřadného systému definovaného v kapitole 3.1.2, první kvadrant je vpravo nahoře, druhý vlevo nahoře, třetí vlevo dole a čtvrtý vpravo dole. K těmto kvadrantům jsou definovány množiny elektrod:

Kvadrant 1 = $\{Fp2, F4, C4, F8, T4, Fz\}$,

Kvadrant 2 = $\{Fp1, F3, C3, F7, T3, Fz\}$,

Kvadrant 3 = $\{C3, P3, O1, T3, T5, Pz\}$,

Kvadrant 4 = $\{C4, P4, O2, T4, T6, Pz\}$.

Dále definujeme prahovou úroveň:

$$t_Q = k_1 \mu,$$

kde konstanta k_1 je reálné nezáporné číslo a μ je střední hodnota klasifikovaného profilu. Pokud alespoň dvě elektrody ze skupiny mají ohodnocení vyšší než je prahová úroveň, pak je skupina označena jako pozitivní. V opačném případě je skupina označena jako negativní.

Definované množiny odpovídají svou polohou mozkovým lalokům. Obě hemisféry jsou rozděleny do čtyř částí. Část frontální, temporální, parietální a okcipitopolární. K těmto částem jsou přiřazeny následující skupiny elektrod:

Frontální vlevo = $\{Fp1, F3, C3, F7, Fz, Cz\}$,

Frontální vpravo = $\{Fp2, F4, C4, F8, Fz, Cz\}$,

Temporální vlevo = $\{T3, T5\}$,

Temporální vpravo = $\{T4, T6\}$,

Parietální vlevo = $\{P3, Pz\}$,

Parietální vpravo = $\{P4, Pz\}$,

Okcipitopolární vlevo = $\{O1\}$,

Okcipitopolární vpravo = $\{O2\}$,

I v tomto případě je definována prahová úroveň:

$$t_L = k_2\mu,$$

kde konstanta k_2 je reálné nezáporné číslo a μ má stejný význam jako v případě klasifikace vůči kvadrantům. Pokud alespoň dvě elektrody ze skupiny mají ohodnocení vyšší než je prahová úroveň, pak je skupina označena jako pozitivní. V opačném případě je skupina označena jako negativní. Výjimku tvoří *Okcipitopolární* skupiny, u kterých je požadována alespoň jedna elektroda s vyšším ohodnocením než je prahová úroveň.

V obou případech může být označeno jako pozitivních 0 až n skupin. V případě kvadrantů je $n = 4$, u laloků je $n = 8$.

3.5 Statistické zpracování získaných dat

Za referenční rozhodnutí je považována klasifikace MRI modelu. Tato klasifikace je porovnávána s klasifikací váženého EEG profilu a s klasifikací EEG profilu ze shlukové analýzy. Porovnání může vyústit ve čtyři různé výsledky.

True Positive (TP) Reference je pozitivní a rozhodnutí je taktéž pozitivní.

True Negative (TN) Reference je negativní a rozhodnutí je taktéž negativní.

False Positive (FP) Reference je negativní, ale rozhodnutí je pozitivní.

False Negative (FN) Reference je pozitivní, ale rozhodnutí negativní.

Součet rozhodnutí v daných kategoriích je roven počtu pokusů, v případě kvadrantů jsou prováděna čtyři porovnávání a u laloků celkem osm.

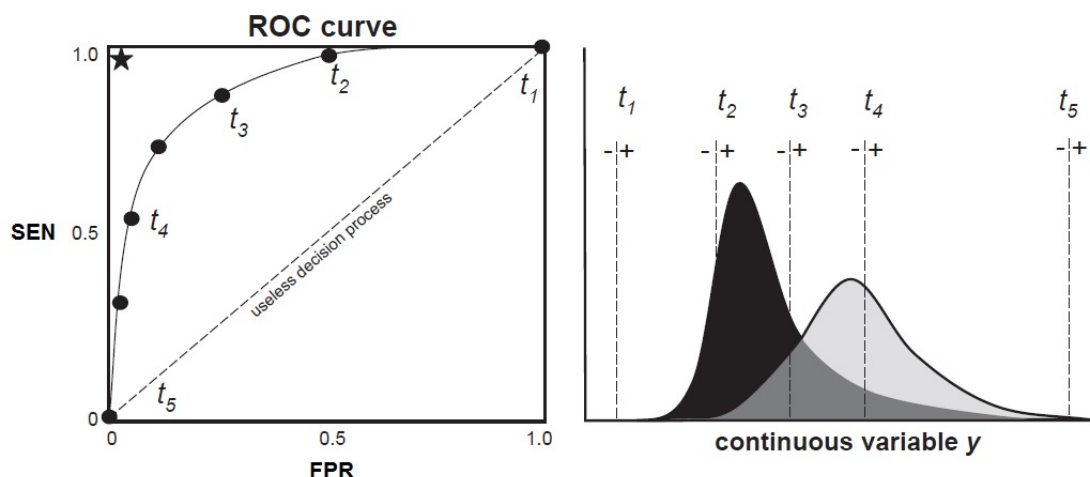
$$TP + TN + FP + FN = n$$

3.5.1 Hodnocení diagnostických testů

Kvalita testu se často hodnotí pomocí statistických nástrojů jako jsou senzitivita, specificita, prediktivní hodnota pozitivního testu a přesnost.

Senzitivita je definována jako pravděpodobnost pozitivního rozhodnutí v případě, že oblast je skutečně pozitivní:

$$SEN = \frac{TP}{TP + FN}.$$



Obrázek 5 V levém grafu je vykreslena ROC křivka pro různé úrovně prahového koeficientu t . Ideální bod s jedničkovou senzitivitou a nulovou FPR je v grafu označen hvězdou [12].

Specifická je definována jako pravděpodobnost, že test bude negativní, když je referenční oblast také negativní:

$$SPEC = \frac{TN}{TN + FP}.$$

Prediktivní hodnota pozitivního testu je pravděpodobnost, že referenční oblast je skutečně pozitivní, když je test pozitivní.

$$PPV = \frac{TP}{TP + FP}.$$

Přesnost (celková správnost) testu je definována jako míra toho jak dobře algoritmus označuje pozitivní a negativní oblasti:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} = \frac{TP + TN}{n}.$$

3.5.2 ROC křivky

ROC (*Receiver Operator Characteristic*) křivky byly navrženy za účelem grafického znázornění kompromisu mezi chybou *I.* a *II.* druhu pro různé rozhodovací úrovně algoritmu. ROC křivka vyjadřuje vztah mezi senzitivitou a $1 - Specificita$, také označovaná jako FPR (*False Positive Rate*). Ideální test by měl mít co možná nejvyšší senzitivitu při co možná nejmenším FPR. Příklad typické ROC křivky je uveden na Obrázku 5. Plocha pod křivkou má význam síly testu [12, 13].

4 Výsledky

Způsobem naznačeným na Obrázku 2 je provedena analýza záznamů pacientů a zdravých novorozenců. Po provedení četnostní analýzy je z dat extrahována hlavní komponenta. Z hlavní komponenty dat novorozenců je určena chybovost detektoru na tomto typu dat. Hlavní komponenta dat pacientů je vážena na základě chybovosti detektoru a variability vlastních dat. Paralelně s automatickou analýzou dat byla provedena shluková analýza dat pacientů. Z výsledných shluků byly vybrány pouze shluky obsahující epileptiformní výboje. Vážené EEG profily pacientů a EEG profily pacientů získané ze shlukové analýzy jsou klasifikovány vůči kvadrantům a lalokům.

Kódy reprezentující lokalizace tuberózních skleróz v mozku jsou pomocí navrženého modelu převedeny na MRI profily pacientů. Profily jsou následně klasifikovány stejným způsobem jako oba EEG profily. Kvalita algoritmu je popsána senzitivitou, specificitou, prediktivní hodnotou pozitivního testu a přesností.

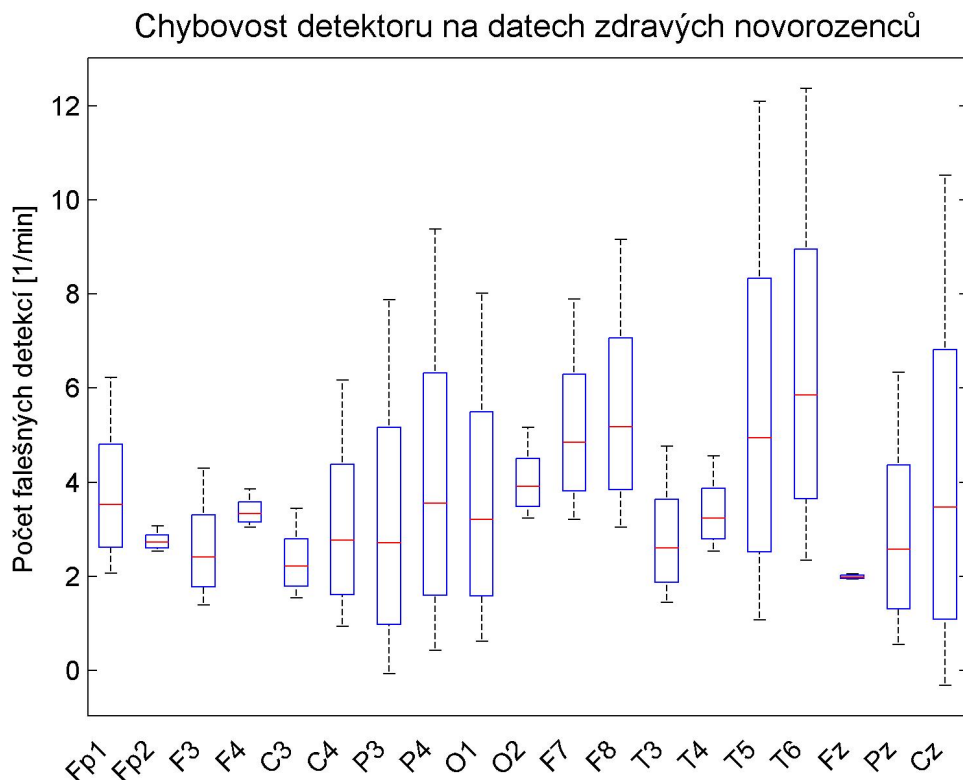
V případě vážených EEG profilů je navržený postup zkoumán pro několik hodnot parametru k . Parametr k určuje zda je kladen větší důraz na váhu určenou z chybovosti detektoru nebo na váhu určenou z vlastních dat.

4.1 Chybovost detektoru

Data *Novorozence 3* a *Novorozence 4* byla vyřazena ze zpracování z důvodu zhoršené kvality. Ze zbylých šesti záznamů pěti novorozenců je určena chybovost detektoru, počet falešně pozitivních detekcí za minutu na kanál, na těchto specifických datech. Výsledky zpracování jsou uvedeny na Obrázku 6. V grafu je červeně vyznačen medián kanálu, modře 25-percentil a černou barvou 75-percentil.

4.1.1 Váhy určené z dat zdravých novorozenců w_j^I

Váhy kanálů w_j^I byly určeny podle vztahu 1 definovaného v kapitole 3.3.3. Váhy jsou nezávislé na datech pacientů a magnetických rezonancích, a proto je tato komponenta celkové váhy vždy stejná. Na Obrázku 7 jsou uvedeny vypočítané váhy w_j^I jednotlivých kanálů.



Obrázek 6 Chybovost detektoru určená z dat zdravých novorozenců. Červenou čarou je vyznačen medián chybovosti kanálu. První značka je 25-percentil a druhá 75-percentil.

4.2 Porovnání modelu s manuální klasifikací

Na základě dat magnetických rezonancí byla provedena manuální klasifikace do skupin popsanych v kapitole 3.4. Manuální klasifikace jsou porovnávány s klasifikací podle modelu. Procedura je rozdělena na klasifikaci prvního a druhého vyšetření magnetickou rezonancí.

4.2.1 První vyšetření magnetickou rezonancí

Do této kategorie spadá celkem 7 vyšetření. Z porovnání manuální klasifikace s klasifikací na základě modelu jsou určeny střední hodnoty a směrodatné odchylky hodnotících kritérií.

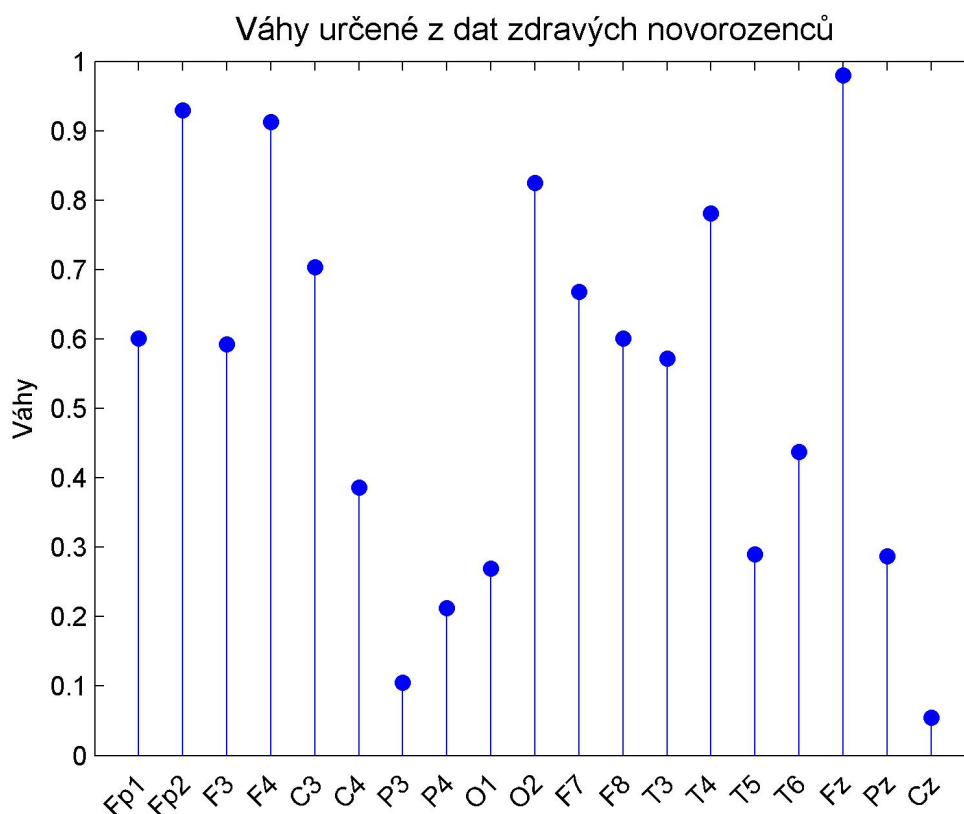
Klasifikace vůči kvadrantu:

$$\text{SEN} = 1$$

$$\text{SPEC} = (0,95 \pm 0,13)$$

$$\text{PPV} = (0,93 \pm 0,19)$$

$$\text{ACC} = (0,96 \pm 0,09)$$



Obrázek 7 Složka vah w_j^I určená z chybovosti detektoru na datech zdravých novorozenců.

Klasifikace vůči laloku:

$$\mathbf{SEN} = (0,93 \pm 0,19)$$

$$\mathbf{SPEC} = 1$$

$$\mathbf{PPV} = 1$$

$$\mathbf{ACC} = (0,98 \pm 0,05)$$

4.2.2 Druhé vyšetření magnetickou rezonancí

Do této kategorie spadá celkem 6 vyšetření. Z porovnání manuální klasifikace s klasifikací na základě modelu jsou určeny střední hodnoty a směrodatné odchylky hodnotících kritérií.

Klasifikace vůči kvadrantu:

$$\mathbf{SEN} = 1$$

$$\mathbf{SPEC} = (0,94 \pm 0,14)$$

$$\mathbf{PPV} = (0,92 \pm 0,20)$$

$$\mathbf{ACC} = (0,96 \pm 0,10)$$

Klasifikace vůči laloku:

$$\mathbf{SEN} = (0,92 \pm 0,20)$$

$$\mathbf{SPEC} = (0,95 \pm 0,12)$$

$$\mathbf{PPV} = (0,89 \pm 0,27)$$

$$\mathbf{ACC} = (0,94 \pm 0,10)$$

4.3 Váhy w_j^{II} EEG dat pacientů

Data pacientů jsou rozdělena do dvou skupin. První skupina náleží prvnímu vyšetření magnetickou rezonancí a jsou do ní zařazeny záznamy se stářím pacienta menším než dva roky. Druhá skupina náleží druhému vyšetření magnetickou rezonancí a obsahuje záznamy se stářím pacienta nad dva roky. Pokud některé vyšetření magnetickou rezonancí chybí nebo není provedeno, záznamy které by tomuto vyšetření náležely, nejsou při analýze použity. Z dat jsou podle vztahu 2 vypočteny váhy w_j^{II} kanálů pro obě vyšetření magnetickou rezonancí.

4.3.1 Pacient 1

Z celkového počtu 15 záznamů je prvních 8 zařazeno do první skupiny a zbylých 7 záznamů tvoří druhou skupinu. *Pacient 1* má provedené obě vyšetření magnetickou rezonancí. Vypočítané váhy kanálů jsou uvedeny na Obrázku 8.

4.3.2 Pacient 2

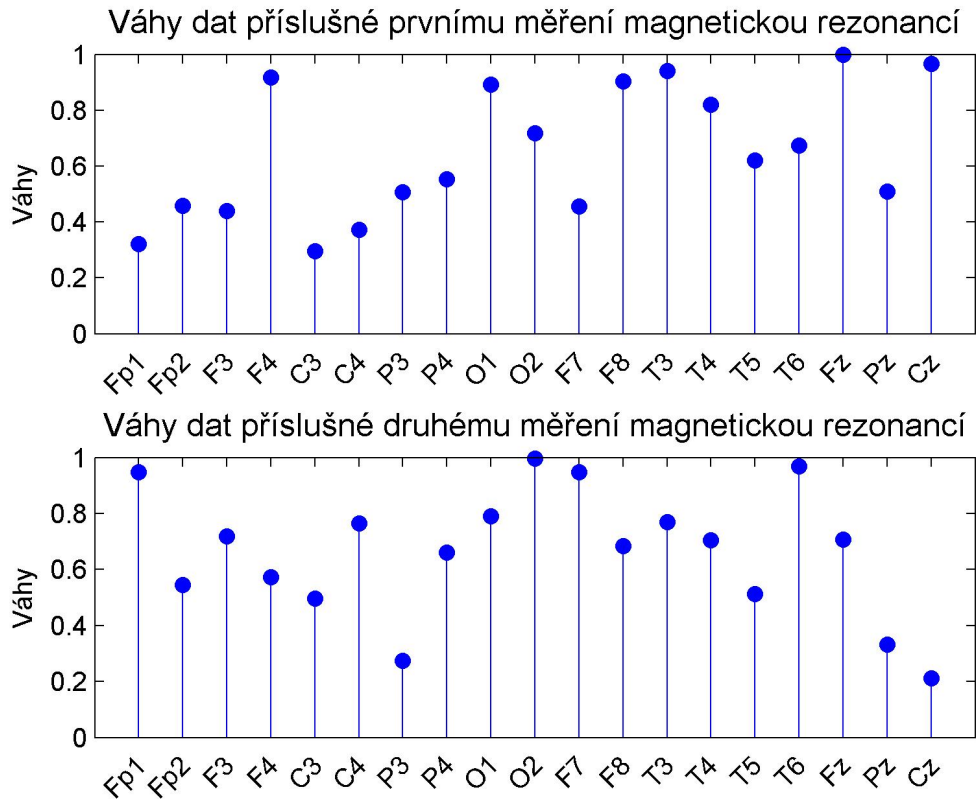
Pacient 2 má celkem 18 záznamů, z toho prvních 14 je zařazeno do první skupiny a zbývající 4 záznamy do druhé. Obě vyšetření magnetickou rezonancí byla provedena a ohodnocena. Vypočítané váhy kanálů jsou uvedeny na Obrázku 9.

4.3.3 Pacient 3

Pacient 3 má celkem 11 záznamů, z toho prvních 9 je zařazeno do první skupiny a zbývající 2 záznamy do druhé skupiny. Obě vyšetření magnetickou rezonancí byla provedena a ohodnocena. Vypočítané váhy kanálů jsou uvedeny na Obrázku 10.

4.3.4 Pacient 4

Pacient 4 má celkem 8 záznamů, z toho prvních 6 je zařazeno do první skupiny a zbylé 2 záznamy do druhé. V tomto případě bylo provedeno a ohodnoceno pouze první vyšetření magnetickou rezonancí. Druhá skupina záznamů proto není v analýze využívána. Vypočítané váhy kanálů jsou uvedeny na Obrázku 11.



Obrázek 8 Váhy w_j^{II} pro *Pacienta 1*. V horním grafu jsou váhy kanálů pro první vyšetření magnetickou rezonancí. Ve spodním grafu jsou váhy kanálů pro druhé vyšetření magnetickou rezonancí.

4.3.5 Pacient 5

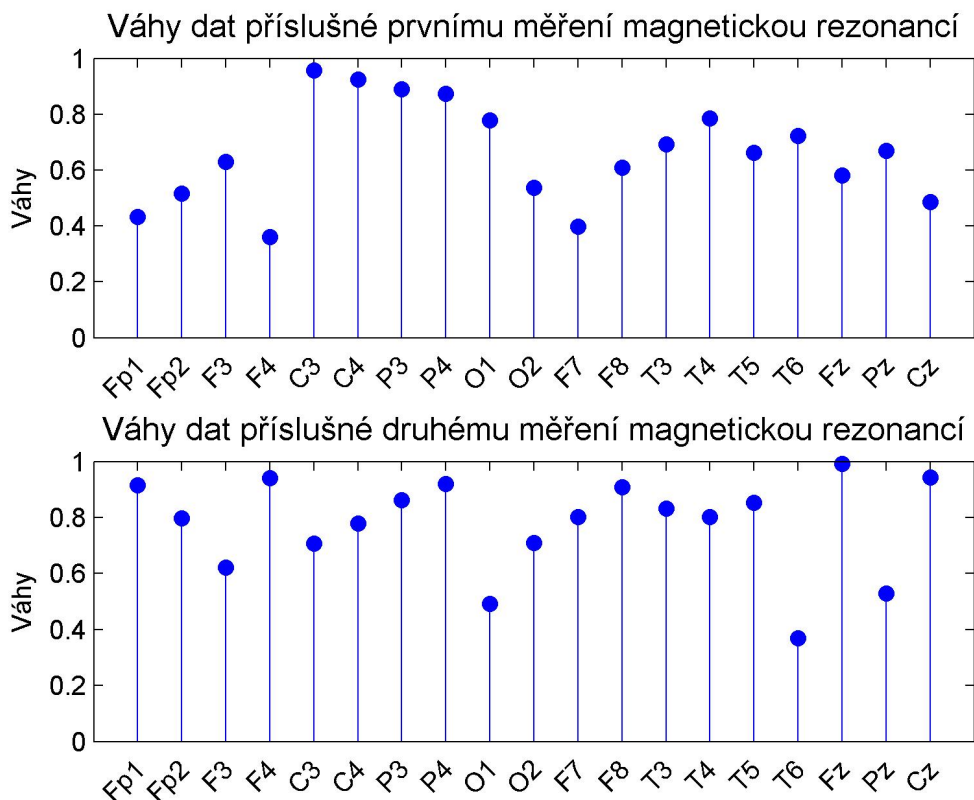
Pacient 5 má celkem 7 záznamů, z toho prvních 6 je zařazeno do první skupiny a zbylý záznamy do druhé. *Pacient 5* absolvoval obě vyšetření magnetickou rezonancí. Pouze první měření je ohodnoceno, ohodnocení druhého měření chybí. Záznam zařazený do druhé skupiny není v analýze použit. Vypočítané váhy kanálů jsou uvedeny na Obrázku 12.

4.3.6 Pacient 6

Pacient 6 má k dispozici 9 záznamů a všechny tyto záznamy jsou zařazeny do první skupiny. *Pacient 6* absolvoval pouze první měření magnetickou rezonancí. Vypočítané váhy kanálů jsou uvedeny na Obrázku 13.

4.3.7 Pacient 7

Pacient 7 má celkem 14 záznamů, z toho prvních 7 je zařazeno do první skupiny a zbývajících 7 záznamů do druhé. Z první skupiny byl vyřazen první záznam kvůli zhoršené



Obrázek 9 Váhy w_j^{II} pro *Pacienta 2*. V horním grafu jsou váhy kanálů pro první vyšetření magnetickou rezonancí. Ve spodním grafu jsou váhy kanálů pro druhé vyšetření magnetickou rezonancí.

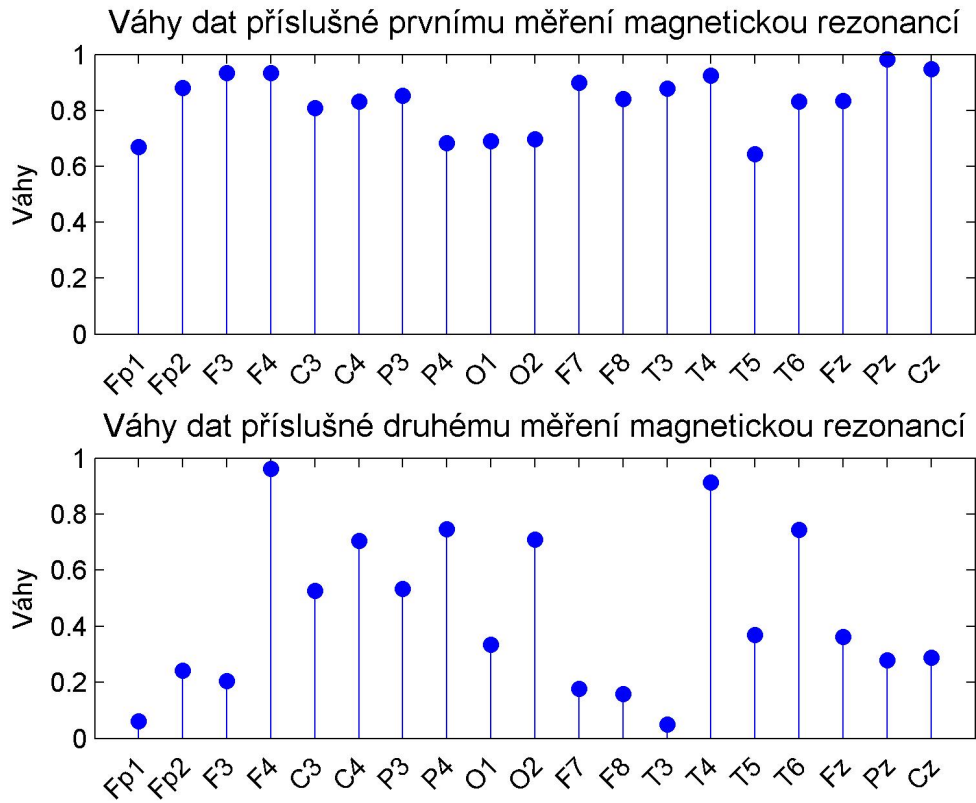
kvalitě. Obě vyšetření magnetickou rezonancí byla provedena a ohodnocena. Vypočítané váhy kanálů jsou uvedeny na Obrázku 14.

4.3.8 Pacient 8

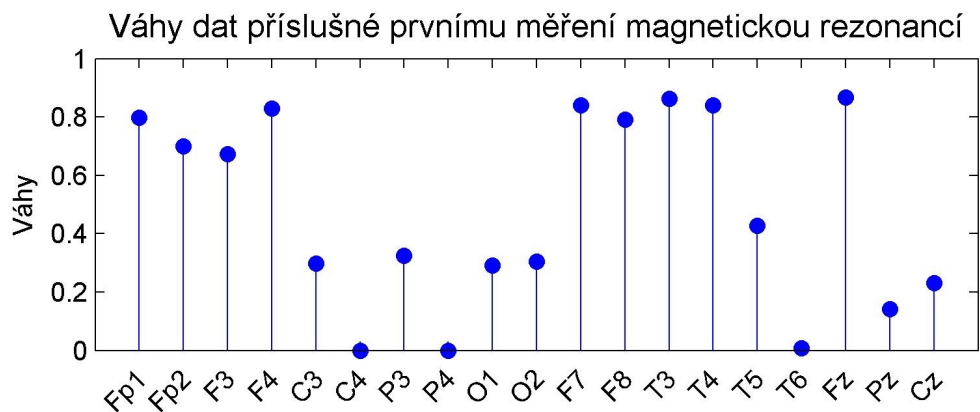
Pacient 8 má celkem 8 záznamů, z toho prvních 5 je zařazeno do první skupiny a zbývající 3 záznamy do druhé. Obě vyšetření magnetickou rezonancí byla provedena a ohodnocena. Expert nenalezl v prvním vyšetření magnetickou rezonancí žádné tuberozní sklerózy, a proto není první skupina dat používána při analýze. Vypočítané váhy kanálů jsou uvedeny na Obrázku 15.

4.3.9 Pacient 9

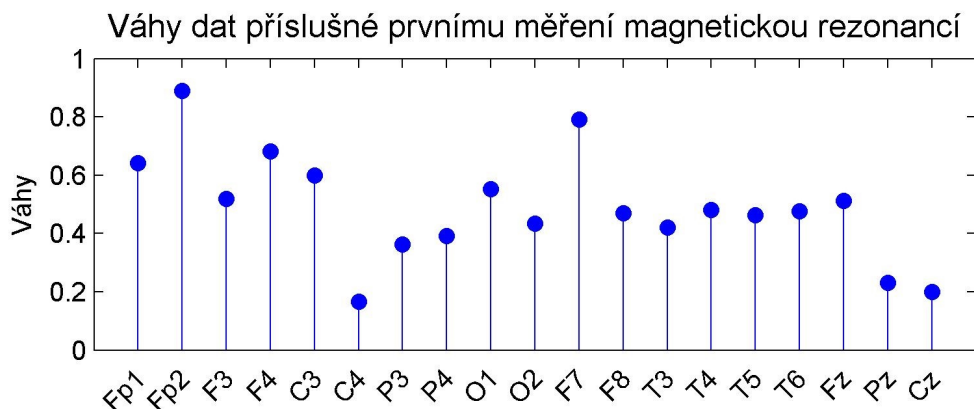
Pacient 9 má celkem 5 záznamů a všechny byly zařazeny do druhé skupiny. Dostupné záznamy byly pořízeny ve věku nad dva roky. Obě vyšetření magnetickou rezonancí byla provedena a ohodnocena. Vypočítané váhy kanálů jsou uvedeny na Obrázku 16.



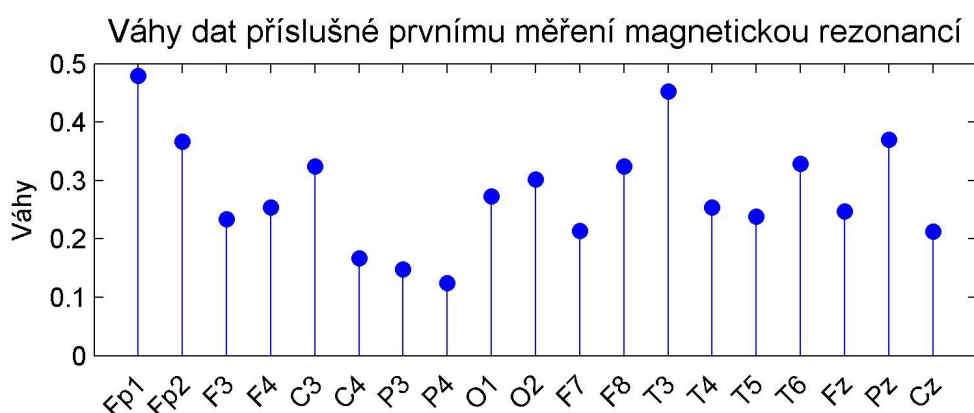
Obrázek 10 Váhy w_j^{II} pro *Pacienta 3*. V horním grafu jsou váhy kanálů pro první vyšetření magnetickou rezonancí. Ve spodním grafu jsou váhy kanálů pro druhé vyšetření magnetickou rezonancí.



Obrázek 11 Váhy w_j^{II} pro *Pacienta 4*. V grafu jsou zachyceny váhy kanálů pro první vyšetření magnetickou rezonancí.



Obrázek 12 Váhy w_j^{II} pro *Pacienta 5*. V grafu jsou zachyceny váhy kanálů pro první vyšetření magnetickou rezonancí.

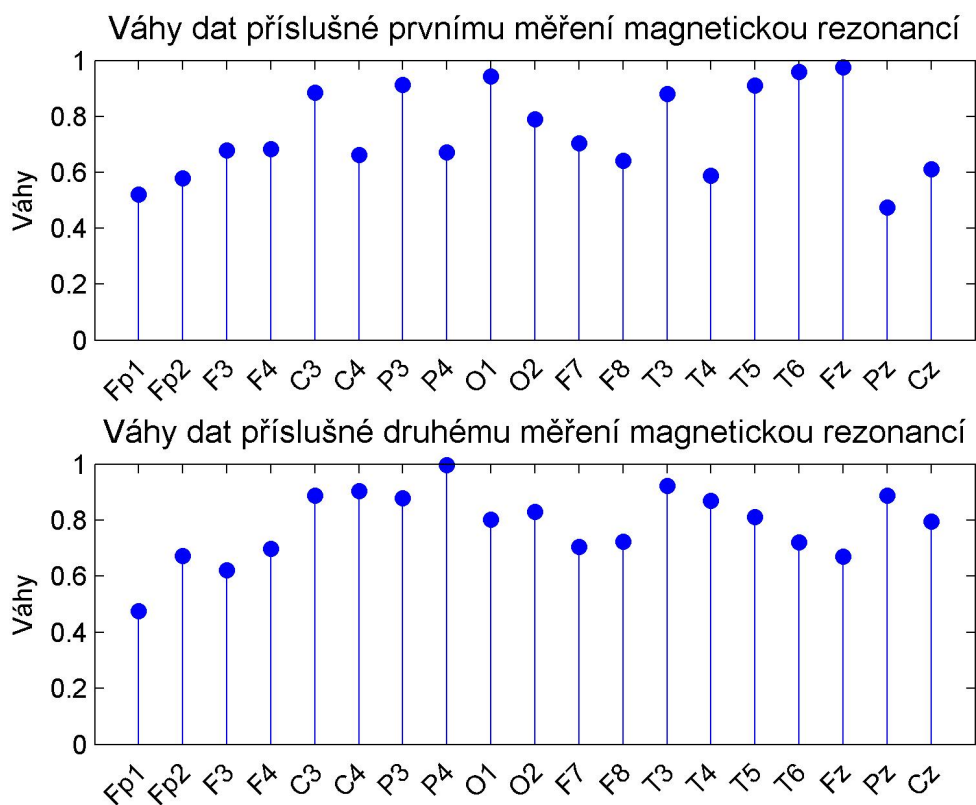


Obrázek 13 Váhy w_j^{II} pro *Pacienta 6*. V grafu jsou zachyceny váhy kanálů pro první vyšetření magnetickou rezonancí.

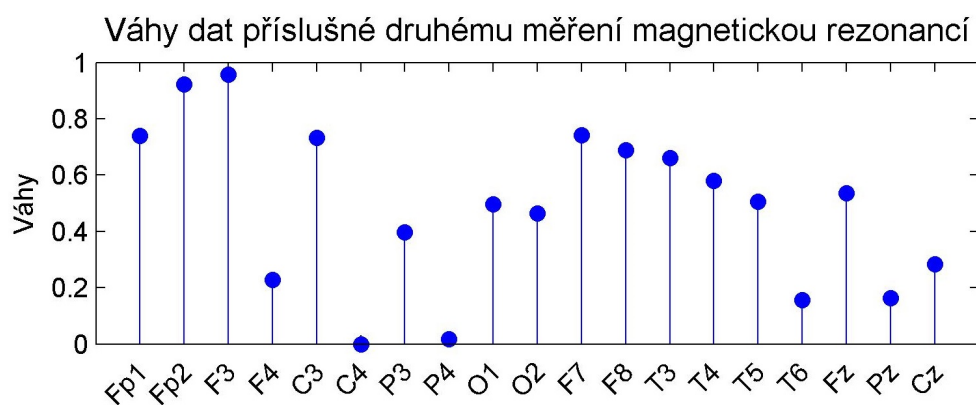
4.4 ROC křivka jako funkce parametru k

Z dílčích vah určených v kapitole 4.1.1 a 4.3 jsou podle vztahu 3 vypočítány celkové váhy příslušných skupin dat pro různé hodnoty parametru k . Vlastnosti algoritmu byly testovány pro celkem 11 hodnot parametru, od nuly do jedné s přírůstkem 0,1. V případě, kdy $k = 0$, jsou data vážena pouze vahou w_j^I . Naopak je-li $k = 1$, jsou data vážena pouze vahou w_j^{II} . Výsledné vážené profily pacientů jsou klasifikovány vůči kvadrantům a lalokům. Získané klasifikace jsou porovnány s klasifikacemi modelu a jsou určeny statistické parametry definované v kapitole 3.5.1. Popsaný postup je zopakován pro všechny hodnoty parametru k .

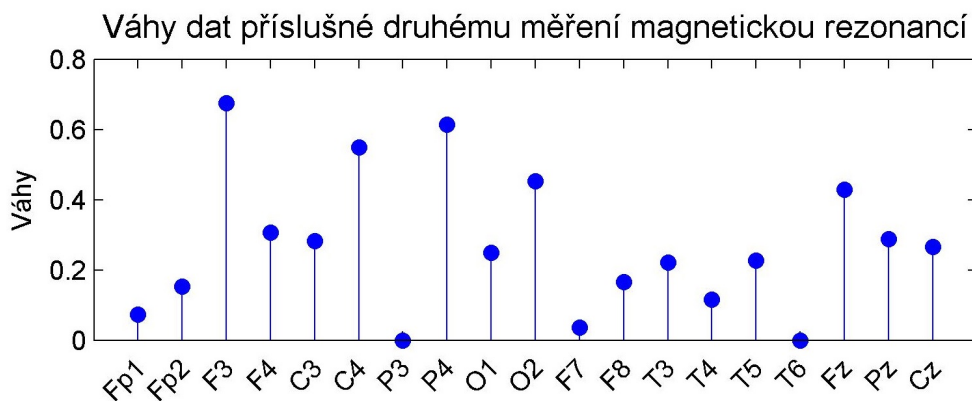
Hodnoty statistických parametrů jsou vyneseny do grafů. Zobrazená data jsou rozdělena do čtyř skupin. První dvě skupiny tvoří výsledky zpracování záznamů přiřazených prvnímu vyšetření magnetickou rezonancí, které jsou klasifikovány vůči kvadrantům a lalokům. Obdobně druhé dvě skupiny tvoří výsledky zpracování záznamů přiřazených druhému vyšetření magnetickou rezonancí, taktéž klasifikované vůči kvadrantům



Obrázek 14 Váhy w_j^{II} pro *Pacienta 7*. V horním grafu jsou váhy kanálů pro první vyšetření magnetickou rezonancí. Ve spodním grafu jsou váhy kanálů pro druhé vyšetření magnetickou rezonancí.



Obrázek 15 Váhy w_j^{II} pro *Pacienta 8*. V grafu jsou zachyceny váhy kanálů pro druhé vyšetření magnetickou rezonancí.



Obrázek 16 Váhy w_j^{II} pro *Pacienta 9*. V grafu jsou zachyceny váhy kanálů pro druhé vyšetření magnetickou rezonancí.

a laloků. Ve všech třech grafech je klasifikace záznamů přiřazených prvnímu vyšetření magnetickou rezonancí vůči kvadrantu označena červenou barvou. Zeleně je vyznačena klasifikace skupiny dat přiřazené druhému vyšetření magnetickou rezonancí vůči kvadrantům. Modrou barvou jsou značeny klasifikace první skupiny dat vůči lalokům a fialově druhé skupiny dat.

Na Obrázku 17 jsou uvedeny ROC křivky pro jednotlivé skupiny dat. Skupiny jsou barevně odlišeny a odstíny barev reprezentují různé hodnoty parametru k , pro které byla data analyzována. Zároveň jsou pro použité hodnoty parametru k vypočteny prediktivní hodnota pozitivního testu a přesnost. Prediktivní hodnota pozitivního testu je uvedena na Obrázku 18 a přesnost na Obrázku 19.

4.5 Kvalita klasifikace dat získaných shlukovou analýzou

Obdobně jako v kapitole 4.2 je provedeno porovnání klasifikací na základě navrženého modelu s klasifikacemi dat zpracovaných shlukovým algoritmem.

4.5.1 První vyšetření magnetickou rezonancí

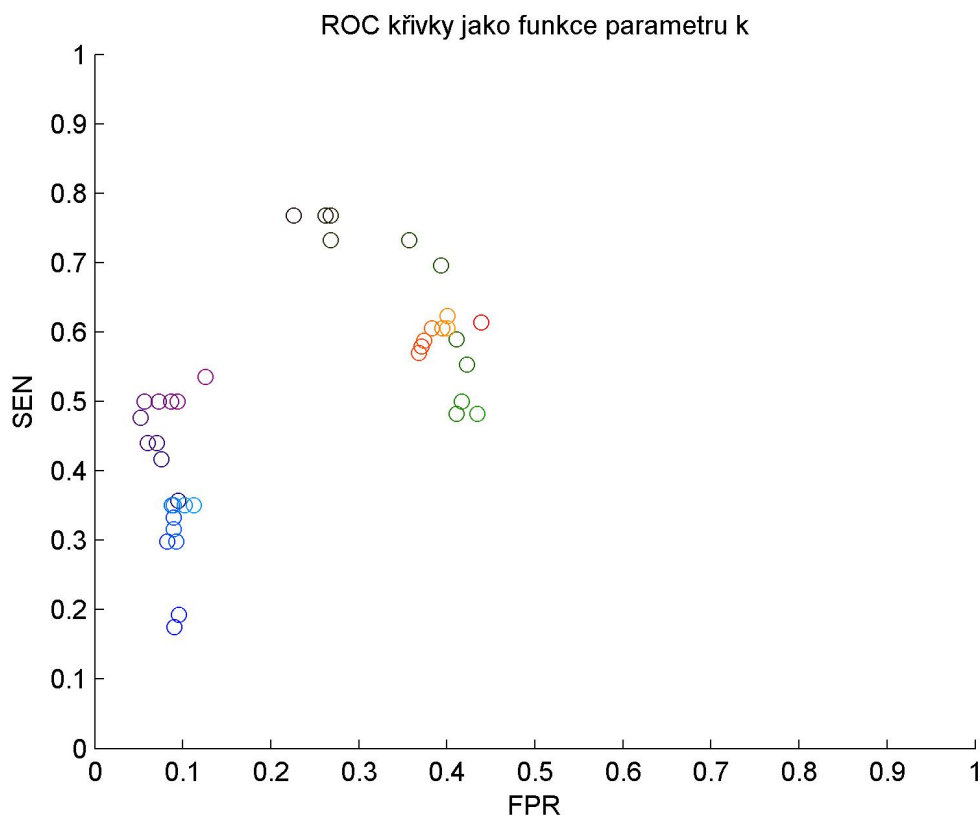
Do této kategorie spadá celkem 7 vyšetření. Z porovnání klasifikace dat zpracovaných shlukovým algoritmem s klasifikací na základě modelu jsou určeny střední hodnoty a směrodatné odchylky hodnotících kritérií.

Klasifikace vůči kvadrantu:

$$\text{SEN} = (0,09 \pm 0,29)$$

$$\text{SPEC} = (0,89 \pm 0,18)$$

$$\text{PPV} = (0,05 \pm 0,18)$$



Obrázek 17 Senzitivita a pravděpodobnost falešných detekcí pro různé hodnoty parametru k . Červenou, respektive zelenou barvou, jsou označeny vlastnosti klasifikace vůči kvadrantům skupiny dat přiřazené prvnímu vyšetření magnetickou rezonancí, respektive druhému vyšetření magnetickou rezonancí. Modrou, respektive fialovou barvou, jsou označeny vlastnosti klasifikace vůči lalokům skupiny dat přiřazené prvnímu vyšetření magnetickou rezonancí, respektive druhému vyšetření magnetickou rezonancí. Tmavší odstíny barev odpovídají nižším hodnotám parametru k .

$$\mathbf{ACC} = (0,69 \pm 0,13)$$

Klasifikace vůči laloku:

$$\mathbf{SEN} = (0,09 \pm 0,29)$$

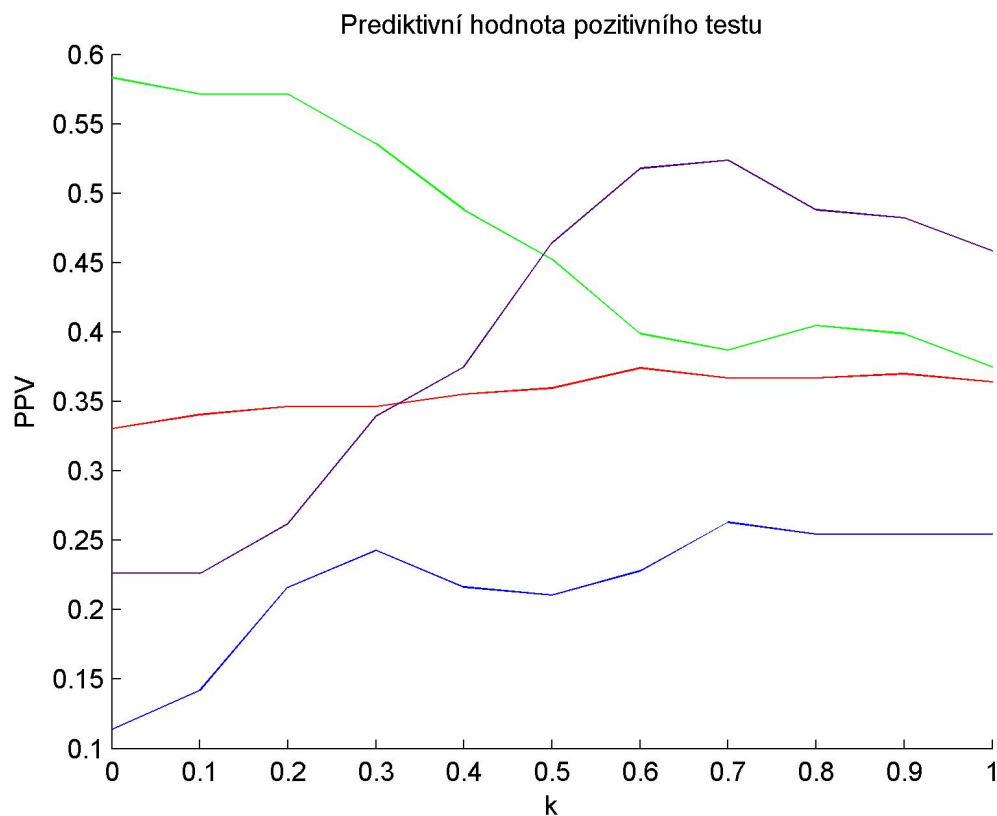
$$\mathbf{SPEC} = (0,96 \pm 0,07)$$

$$\mathbf{PPV} = (0,07 \pm 0,24)$$

$$\mathbf{ACC} = (0,86 \pm 0,07)$$

4.5.2 Druhé vyšetření magnetickou rezonancí

Do této kategorie spadá celkem 6 vyšetření. Z porovnání klasifikace dat zpracovaných shlukovým algoritmem s klasifikací na základě modelu jsou určeny střední hodnoty a směrodatné odchylky hodnotících kritérií.



Obrázek 18 Vývoj prediktivní hodnoty pozitivního testu pro hodnoty parametru k jdoucího od nuly do jedné. Červenou barvou jsou vykresleny výsledky pro klasifikaci první skupiny dat vůči kvadrantům, zeleně druhé skupiny dat vůči kvadrantům, modře první skupiny dat vůči lalokům a konečně fialově druhé skupiny dat vůči lalokům.

Klasifikace vůči kvadrantu:

$$\mathbf{SEN} = (0,07 \pm 0,26)$$

$$\mathbf{SPEC} = (0,87 \pm 0,19)$$

$$\mathbf{PPV} = (0,07 \pm 0,26)$$

$$\mathbf{ACC} = (0,67 \pm 0,17)$$

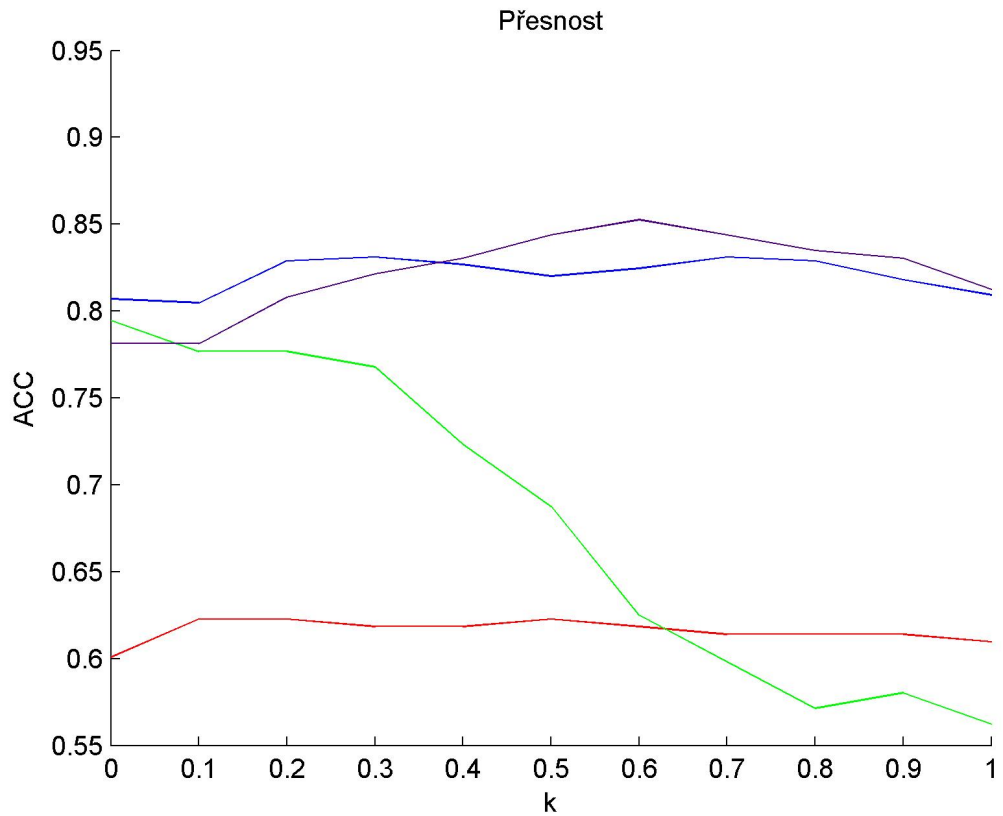
Klasifikace vůči laloku:

$$\mathbf{SEN} = (0,07 \pm 0,26)$$

$$\mathbf{SPEC} = (0,94 \pm 0,08)$$

$$\mathbf{PPV} = (0,07 \pm 0,26)$$

$$\mathbf{ACC} = (0,83 \pm 0,08)$$



Obrázek 19 Vývoj přesnosti algoritmu pro hodnoty parametru k jdoucího od nuly do jedné. Červenou barvou jsou vykresleny výsledky pro klasifikaci první skupiny dat vůči kvadrantům, zeleně druhé skupiny dat vůči kvadrantům, modře první skupiny dat vůči lalokům a konečně fialově druhé skupiny dat vůči lalokům.

5 Diskuze

Neinvazivní EEG je v mnohém odlišné od invazivního EEG, a proto bylo nutné prozkoumat vlastnosti detektoru na EEG datech zdravých novorozenců. Četnostní analýzou dat zdravých novorozenců bylo zjištěno, že se chybovost pohybuje v průměru od 2 do 6 falešných detekcí za minutu. U zdravých pacientů by se neměly vyskytovat žádné epileptiformní výboje, a proto se dá předpokládat, že veškeré detekce jsou falešně pozitivní. Četnostní analýza také ukázala, že některé elektrody mají velký rozptyl počtu falešných detekcí, a lze je považovat za méně věrohodné. Jsou to například elektrody $P3, P4, O1$ a zvláště pak elektrody $T5, T6, Cz$. Poslední tři jmenované elektrody jsou v těsné blízkosti úponu krčních svalů a falešné detekce jsou tedy způsobeny svalovými artefakty. Ze směrodatných odchylek kanálů jsou určeny váhy w_j^I . Ve většině případů je přiřazena větší váha elektrodám nad pravou hemisférou než jejich protikladům nad levou hemisférou. Výše jmenované elektrody mají kvůli jejich variabilitě přiřazené nižší váhy.

Pro převod lokalizací tuberózních skleróz z vyšetření MRI do vhodnější podoby k porovnávání, byl navržen model popsáný v kapitole 3.1. MRI profily pacientů z výstupu modelu byly klasifikovány vůči kvadrantům a lalokům. Tyto klasifikace byly porovnány s manuálními klasifikacemi vůči kvadrantům a lalokům. Bylo ukázáno, že klasifikace navrženým algoritmem poměrně dobře odpovídá skutečnosti. Falešné detekce u klasifikace vůči kvadrantům byly způsobeny označením protilehlého kvadrantu. Namísto označení pouze 1. kvadrantu, byl označen i kvadrant 2. V případě klasifikace vůči lalokům byly falešné detekce způsobeny označením temporálního a okcipitopolárního laloku.

Data pacientů jsou z důvodu vylepšení vlastností algoritmu vážena, první složka vah byla určena z dat zdravých novorozenců, druhá z vlastností dat samotných. Druhé složky vah dat jsou určeny v kapitole 4.3. Druhá složka vah určená ze skupiny dat přiřazených druhému vyšetření MRI *Pacienta 3* vykazuje nízké hodnoty u některých elektrod. Složka vah byla určena ze dvou záznamů, které jsou od sebe vzdáleny dva měsíce. Nízké váhy elektrod by mohly být způsobeny špatným odhadem prvních dvou momentů náhodného vektoru. Podobná situace nastává u *Pacienta 8*, v tomto případě je složka vah určena ze tří záznamů. U *Pacienta 9* je složka vah počítána z pěti záznamů, které se však pohybují v rozmezí od dvou do osmi let. Nízké váhy kanálů by mohly být způsobeny přirozeným vývojem mozku v průběhu šesti let.

Celková váha dat je dána kombinací složky určené ze záznamů zdravých novorozenců

a příslušné složky určené z vlastních dat. Poměr zastoupení těchto složek je možné ovlivnit parametrem k z intervalu $\langle 0, 1 \rangle$. Pro $k = 0$ je celková váha rovna váze určené z vlastních dat. Naopak pro $k = 1$ je celková váha rovna váze určené ze záznamů zdravých pacientů. Závislost kvality klasifikace na parametru k je analyzováno pomocí ROC křivek. Pro každou hodnotu parametru k je zároveň zaznamenána prediktivní hodnota pozitivního testu a přesnost. Ideální algoritmus by měl senzitivitu rovnou jedné a nulovou pravděpodobnost falešné detekce. Tento ideální stav je v ROC grafu reprezentován bodem $[0, 1]$. Odtud plyne, že nejvhodnější volba parametru k je taková, když má algoritmus co možná největší senzitivitu při co možná nejmenší pravděpodobnosti falešných detekcí. Skupiny dat přiřazených prvnímu nebo druhému vyšetření MRI byly klasifikovány vůči kvadrantům a lalokům. Tabulka 6 zachycuje optimální hodnotu parametru k pro takto definované skupiny. Pro úplnost uvedme hodnoty prediktivní hodnoty pozitivního testu a přesnosti v optimálních hodnotách parametru k , viz. Tabulka 7.

Paralelně s automatickou analýzou dat navrženým algoritmem byla provedena shluková analýza. Klasifikace získaných výsledků byla provedena stejným způsobem jako v případě automatické analýzy. Ve všech čtyřech případech se klasifikace vyznačovaly vysokou specificitou a nízkou senzitivitou.

Na závěr porovnejme výsledky klasifikace automatického zpracování a výsledky klasifikace po provedení shlukové analýzy. Klasifikace první množiny dat vůči kvadrantu, automatické zpracování dosahuje pro případ optimální hodnoty parametru k senzitivity 0,57 a selektivity 0,63, klasifikace po provedení shlukové analýzy má senzitivitu 0,09 a specificitu 0,89. Klasifikace první množiny dat vůči laloku, automatické zpracování dosahuje pro případ optimální hodnoty parametru k senzitivity 0,35 a selektivity 0,91, klasifikace po provedení shlukové analýzy má senzitivitu 0,09 a specificitu 0,96. Klasifikace druhé množiny dat vůči kvadrantu, automatické zpracování dosahuje pro případ optimální hodnoty parametru k senzitivity 0,77 a selektivity 0,77, klasifikace po provedení shlukové analýzy má senzitivitu 0,07 a specificitu 0,87. Klasifikace druhé množiny dat vůči laloku, automatické zpracování dosahuje pro případ optimální hodnoty parametru k senzitivity 0,50 a selektivity 0,94, klasifikace po provedení shlukové analýzy má senzitivitu 0,07 a specificitu 0,94. Klasifikace po provedení shlukové analýzy dosahuje lepší nebo stejné specificity ve všech případech. Naopak senzitivita je lepší u klasifikace automatického zpracování ve všech případech. Klasifikace po provedení shlukové analýzy je konzervativnější než klasifikace automaticky zpracovaných dat. Za nízkou pravděpodobnost falešné detekce je zapláceno nízkou senzitivitou.

Problémy mohou nastávat při klasifikaci vůči laloku. Temporální, parietální a okcipitopolární lalok jsou popsány velice úzkou skupinou elektrod. V některých případech je nutné se rozhodnout podle ohodnocení jediné elektrody. Navíc pokud temporální elektrody obsahují velké množství svalových artefaktů, jak bylo ukázáno zpracováním

	Q	L
MRI 1	0,4	0,3
MRI 2	0	0,6

Tabulka 6 Optimální hodnota parametru k pro definované skupiny dat klasifikované dvěma způsoby. Sloupec **Q** obsahuje optimální hodnoty parametru pro klasifikaci vůči kvadrantu. Sloupec **L** udává optimální hodnoty parametru pro klasifikaci vůči lalokům.

	PPV	ACC
MRI 1 vs. Q	0,36	0,62
MRI 1 vs. L	0,24	0,83
MRI 2 vs. Q	0,58	0,79
MRI 2 vs. L	0,52	0,85

Tabulka 7 Hodnoty prediktivní hodnoty pozitivního testu a přesnosti pro nalezené optimální hodnoty parametru k . Výraz *MRI 1 vs. Q* značí klasifikaci skupiny dat přiřazené prvnímu vyšetření MRI vůči kvadrantu. Význam ostatních výrazů je analogický.

záznamů zdravých novorozenců, může být klasifikace v této oblasti poměrně obtížná. Přejít na EEG záznamy s větším počtem kanálů, by usnadnil klasifikaci v těchto oblastech.

Příloha A

Appendix

Na přiloženém CD jsou nahrány vytvořené skripty a funkce pro Matlab. Navržený algoritmus je spouštěn hlavním skriptem se jménem *detector_vs_MRI.m*, který je umístěn ve složce *DP_batch_analysis*. Skript předpokládá, že složka *DP_batch_analysis* je umístěna na stejné úrovni jako složky s daty pacientů a zdravých novorozenců. Data zdravých novorozenců jsou jednoduše umístěna ve složce *detektorTest*, která nemá žádnou vnitřní strukturu. Data pacientů jsou umístěna ve složce *pacienti*. Detektor vyžaduje, aby data v této složce, byla uspořádána následujícím způsobem:

```
pacienti>Jméno_Pacienta>Datum_Měření>raw>interictal.
```

Ve složce se skripty jsou také dva *.mat* soubory. První z nich, *paramMatrix.mat*, obsahuje nastavení detektoru a navrženého algoritmu. Druhý soubor, *localizationMRI.mat*, obsahuje datovou strukturu definující příslušnost záznamů k vyšetřením magnetickou rezonancí.

```
localizationMRI(i).name  
localizationMRI(i).position_MRI_1  
localizationMRI(i).position_MRI_2  
localizationMRI(i).record_MRI_1  
localizationMRI(i).record_MRI_2
```

Index *i* jde přes zpracovávané pacienty. Členská proměnná *name* obsahuje *Jméno_Pacienta*. Proměnná *position_MRI_1* kódy lokalizací tuberózních skleróz při prvním vyšetření MRI, pokud pacient vyšetření podstoupil. Proměnná *position_MRI_2* má analogický význam. Proměnná *record_MRI_1* obsahuje seznam s *Daty_Měření* přiřazenými prvním vyšetření MRI. Proměnná *record_MRI_2* má opět analogický význam.

Literatura

- [1] D. W. Webb, R. D. Thomas, J. P. Osborne. *Cardiac rhabdomyomas and their association with tuberous sclerosis*. Archives of Disease in Childhood, 1993.
- [2] Návrh studentského projektu 2. lékařské fakulty Univerzity Karlovy GAUK 2014 podávaný pracovištěm kliniky dětské neurologie Fakultní Nemocnice Motol.
- [3] Radek Janča, Petr Ježdík, Roman Čmejla, Martin Tomášek, Gregory A. Worrell, Matt Stead, Joost Wagenaar, John G. R. Jefferys, Pavel Krsek, Vladimír Komárek, Přemysl Jiruška, Petr Marusic . *Detection of Interictal Epileptiform Discharges Using Signal Envelope Distribution Modelling: Application to Epileptic and Non-Epileptic Intracranial Recordings*. Brain Topography, 2015. doi: 10.1007/s10548-014-0379-1.
- [4] Trans Cranial Technologies. *10/20 System Positioning*, 2012.
- [5] Michal Friesl. *Pravděpodobnost a statistika HYPERTEXTOVĚ*, 2014. Online: <http://home.zcu.cz/friesl/hpsb/tit.html> (visited on 28/4/2015).
- [6] P. Gloor. *Contributions of electroencephalography and electrocorticography to the neurosurgical treatment of the epilepsies*. Advances in Neurology, 8(1975), pp. 59–105.
- [7] E. Niedermeyer, F. H. Lopes da Silva. *Electroencephalography. Basic principles, clinical applications and related fields*. Baltimore : Williams & Wilkins, 1999.
- [8] A. J. E. Geerts. *Detection of Interictal Epileptiform Discharges in EEG*. Master Thesis, University of Twente, 2012.
- [9] Seungduk Lee. Mina Lee. Dalkwon Koh. Beop-Min Kim. Jee Hyun Choi. *Cerebral hemodynamic responses to seizure in the mouse brain: simultaneous near-infrared spectroscopy–electroencephalography study*. Journal of Biomedical Optics 15(3), 2008. doi:10.1117/1.3365952.
- [10] Jon Shlens. *A Tutorial on Principal Component Analysis*. 2003.
- [11] Duncan Fyfe Gillies. *Intelligent Data Analysis and Probabilistic Inference*, 2015. Online: www.doc.ic.ac.uk/~dfg/ProbabilisticInference/Bayesian.html (visited on 1/5/2015).

- [12] Christopher D. Brown, Herbert T. Davis. *Receiver operating characteristics curves and related decision measures: A tutorial*. Chemometrics and Intelligent Laboratory Systems, 80(2006), pp. 24–38.
- [13] Tom Fawcett. *An introduction to ROC analysis*. Pattern Recognition Letters, 27(2006), pp. 861–874.
- [14] Dohnalová L. *Dynamika interiktálních výbojů v nezávislých epileptických zónách*. Diplomová práce, ČVUT v Praze, 2014.
- [15] Polanský Š. *Zpracování Elektroencefalografických Signálů*. Diplomová práce, VUT Brno, 2001.
- [16] Faber J. *QEEG*. GALÉN UK v Praze, 2005.
- [17] Brázdil M, Marusic P. a kol. *Epilepsie temporálního laloku*. TRITON, 2006.