Vilniaus universitetas
Matematikos ir informatikos fakultetas
L I E T U V A

INFORMATIKA (09 P)

ELEKTROENCEFALOGRAMŲ ANALIZĖS METODŲ TYRIMAS

Andrius Vytautas Misiukas Misiūnas

2018 m. spalis

Mokslinė ataskaita DMSTI-DS-09P-18-3

VU Matematikos ir informatikos fakultetas, Didlaukio g. 47, LT-08303 Vilnius
https://mif.vu.lt

Santrauka

Šioje ataskaitoje aprašomi darbai už trečiuosius doktorantūros metus. Remiantis ankstesniais metais sukurtu ir šiemet patobulintu trijų žingsnių (1. EEG pikų detekcija, 2. Aptiktų pikų parametrų nustatymas, 3. Klasifikavimas apgal diagnozę mašinų mokymosi metodais) EEG klasifikavimo pagal diagnozę algoritmu buvo išbandyti aštuoni mašinų mokymosi algoritmai, taikomi sukurto algoritmo trečiajame žingsnyje: 1) dirbtinis neuroninis tinklas (ANN), 2) logistinė regresija, 3) sprendimų medis, 4) atsitiktinis miškas, 5) labai atsitiktinis medis, 6) tiesinė diskriminantinė analizė (LDA), 7) adaptyvusis stiprinimas (AdaBoost), 8) palaikančiųjų vektorių mašina (SVM). Šiuo metu pasiektas patikimumas yra 73%.

Raktiniai žodžiai: EEG, elektroencefalogramos, EEG pikai, epilepsija, mašinų mokymasis, automatinis klasifikavimas

Turinys

1 Įvadas

Pagrindinis disertacijos tikslas – elektroencefalogramų (EEG) kompiuterinė analizė. Šioje ataskaitoje aprašomi darbai, atlikti trečiaisiais doktorantūros metais ir pateikiama kontekstiškai svarbi informacija iš ankstesniais metais atliktų darbų. Pagrindiniai ataskaitiniais metais atlikti darbai buvo:

1. Ankstesniais metais pasiūlyto EEG klasifikavimo pagal diagnozę algoritmo patobulinimas.

2. Įvairių mašinų mokymosi metodų išbandymas klasifikavimo algoritmo trečiajame žingsnyje.

3. Ankstesniais metais parašyto straipsnio publikavimas (straipsnio pataisymas pagal recenzentų pateiktas pastabas, straipsnis priimtas 2018 metų spalio 6 dieną žurnale "Biomedical Signal Processing and Control", IF=2.783, Cite score = 3.62).

4. Šių metų rezultatų pristatymas tarptautinėje konferencijoje ir paskelbimas konferencijos medžiagoje ("The 9th International Conference on Numerical Methods and Applications", 2018 rugpjūčio 20-24, Borovets, Bulgarija. ISI konferencijų medžiaga, publikuojama Springer Lecture Notes in Computer Science).

Pagrindinis šiuo metu gautas doktorantūros mokslinių tyrimų rezultatas – trijų žingsnių EEG klasifikavimo pagal diagnozę algoritmas, sukurtas ankstesniais doktorantūros metais ir patobulintas bei detaliau ištirtas pastaraisiais mokslo metais. Algoritmo pagrindiniai žingsniai: 1) EEG pikų detekcija panaudojant matematinį morfologinį filtrą [MMJ15], 2) aptiktų EEG pikų parametrų nustatymas [MMS16], 3) EEG pikų klasifikavimas pagal diagnozę.

Šiais metais buvo ištirtas įvairių mašinų mokymosi klasifikavimo algoritmų veikimą trečiajame klasifikavimo algoritmo žingsnyje:

- Dirbtinis neuroninis tinklas (ANN)
- Logistinė regresija
- Naivus Bayes
- Sprendimų medis
- Atsitiktinis miškas
- Labai atsitiktinis medis
- Tiesinė diskriminantinė analizė (LDA)
- Adaptyvusis stiprinimas (AdaBoost)

- Palaikančiųjų vektorių mašina (SVM)

Kai kurie šioje ataskaitoje pateikiami duomenys nepasikeitė lyginant su praeitais mokslo metais, tačiau juos būtina aprašyti tam, kad būtų aiškūs šiais metais atlikti darbai. Dėl šios priežasties 2 ir 3 skyriuose aprašyti tie patys duomenys, kaip ankstesnių metų ataskaitoje.

2 Tyrimo duomenys

Šiame darbe naudoti VšĮ Vilniaus Universiteto ligoninės Santaros klinikų filialo Vaikų ligoninės neurologų pateikti EEG tyrimų duomenys. Darbe analizuojami EEG tyrimai buvo atlikti 2010-2017 metais, atrinkti duomenys pasižymintys gėrybinėms vaikų epilepsijoms būdingais EEG pikais. 2018 metų spalį buvo gauta naujų duomenų: naujų EEG abiems diagnozių grupėms ir EEG su neurologų sužymėtais pikais. Čia aprašomi tyrimai buvo atlikti dar be šių duomenų, tačiau jie bus naudojami kitų metų tyrimams.

2.1 EEG parinkimas

Reikėtų pastebėti, kad šiame tyrime trivialūs atvejai nebuvo analizuojami. Grupės I EEG visada pasižymi panašios formos pikais, kurie daug nesikeičia EEG eigoje. Grupės II EEG kai kuriais atvejais akivaizdžiai skiriasi nuo Grupės I EEG net ne profesionalams, tokie atvejai šiame darbe analizuojami nebuvo. Buvo tiriami tik tokie Grupės II duomenys, kurie yra (beveik arba visiškai) neatskiriami neurologams be ligos anamnezės ar kitų duomenų, neesančių EEG failuose.

Šiame darbe buvo apdorotos 94 EEG, gautos iš 86 skirtingų pacientų, padalintų į tokias grupes:

- **Grupė I:** gėrybinė vaikų epilepsija su EEG pikais, šiuo atveju Rolando epilepsija (62 EEG, arba apie 66% visų duomenų).
- **Grupė II:** struktūrinė židininė epilepsija pacientams, kenčiantiems nuo cerebrinio paralyžiaus, smegenų žievės displazijos ir pan. (32 EEG, arba apie 34% iš visų EEG).

2.2 Klasifikatorių mokymo ir testavimo strategija

Šiame darbe buvo griežtai taikomas principas, kad pacientas (su visomis jo EEG ir visais jų pikais) priskiriamas tik vienai: mokymo arba testavimo imčiai. Jei vienas pacientas turi kelias EEG (įrašytas skirtingais laikais) visos jos priskiriamos tai pačiai (mokymo arba testavimo) imčiai. Tai yra atlikta dėl akivaizdžios priežasties: ANN persimokymo (angl. *overfitting*) išvengimo. Visi tyrime aptarti klasifikatoriai buvo mokomi ir testuojami su tomis pačiomis mokymosi ir testavimo duomenų imtimis.

Mokymo ir testavimo imtys (visiems klasifikatoriams) buvo apibrėžtos taip:

- 62 Grupės I EEG, 21 algoritmo mokymui (apie 34%) ir 41 testavimui (apie 66%);

- 32 Grupės II EEG, 11 algoritmo mokymui (apie 34%) ir 21 testavimui (apie 66%).

Mokymo imtys buvo rankiniu būdu (gydytojų neurologų) išvalytos nuo EEG artefaktų (kaip paciento judesiai, mirksėjimas, ir pan.) todėl, kad klasifikatoriai apmokyti su kuo švaresniais duomenimis maksimaliam rezultatui.

3 EEG klasifikavimo algoritmas

Šiame darbo skyrelyje bus detaliau aprašytas EEG (pasiūlytas ankstesniais studijų metais) algoritmas. Jis sudarytas iš tokių esminių žingsnių:

1. EEG pikų aptikimas (plačiau 3.1 skyrelyje ir [MMJ15]);
2. EEG pikų parametrų nustatymas (plačiau 3.2 skyrelyje ir [MMS16]) ir validavimas (plačiau 3.2.1 skyrelyje ir [MMS16]);
3. EEG klasifikatorius (plačiau 3.3 skyrelyje).

3.1 EEG pikų aptikimas

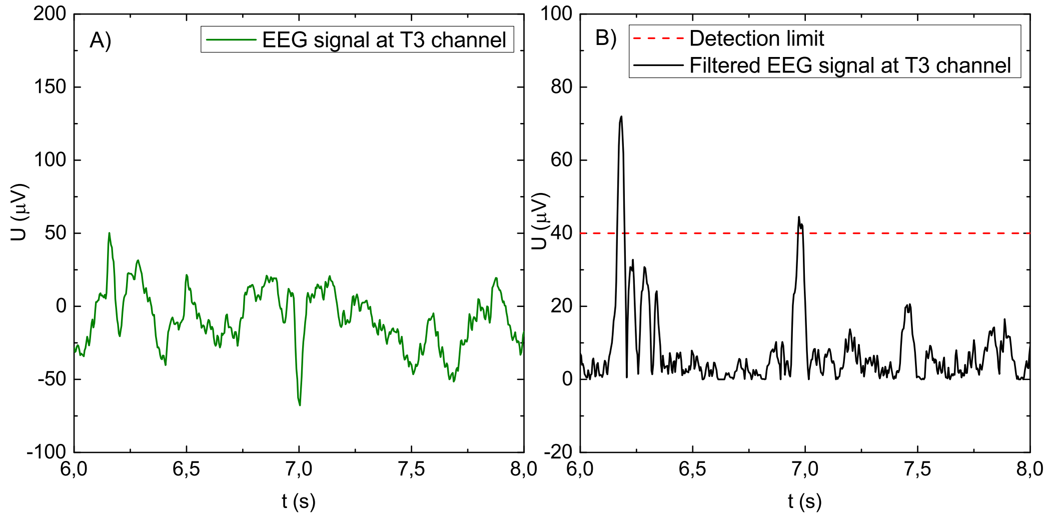
Šiame skyrelyje aprašomas pirmasis EEG klasifikavimo algoritmo žingsnis. Šio algoritmo studija ir įgyvendinimas iš esmės yra ankstesnių doktorantūros metų darbas, tačiau jo veikimas primenamas tam, kad geriau suprasti bendrą EEG klasifikatoriaus veikimą. Šio algoritmo paskirtis yra surasti pikų vietas EEG signale. Algoritmas yra jau žinomas [NNIS99, XWZZ06, XWZ+07, JBBS11, MMJ15].

Algoritmas veikia morfologinių filtrų pagrindu, signalas yra filtruojamas 4 sekundžių slenkančio lango filtru. Morfologinio filtro struktūrinis elementas yra sukonstruotas taip, kad filtruojant signalą pašalintų žinomą normalią smegenų veiklą, kaip pvz. smegenų ritmai, palikdamas tik nenormalią (žr. 1 pav.), pvz EEG pikus. Tačiau nenormaliai atrodanti veikla nebūtinai yra tik EEG būdingi pikai, o ir nemažai kitų darinių, todėl algoritmas ankstesniais metais buvo tobulinamas kuriant pikų parametrų nustatymo ir validavimo algoritmus.

Šiame žingsnyje parenkamas tas EEG kanalas, kuriame randama daugiausiai pikų ir toliau nagrinėjamas būtent šis kanalas. Ankstesnių metų tyrimai rodo, kad būtent šis kanalas būna arčiausiai epilepsijos židinio, todėl pikų parametrų nustatymo algoritmas jame veikia tiksliausiai.

3.2 EEG pikų parametrų nustatymo algoritmas

Šiame skyrelyje aprašomas antrasis EEG klasifikavimo algoritmo žingsnis. Šio algoritmo paskirtis yra gavus EEG pikų vietas išmatuoti jų parametrus. Pradinė šio algoritmo paskirtis buvo išmatuoti EEG pikų parametrus tam, kad juos galima būtų validuoti. Vėliau buvo pastebėta, kad šiuos parametrus būtų galima taikyti ir EEG klasifikavimui

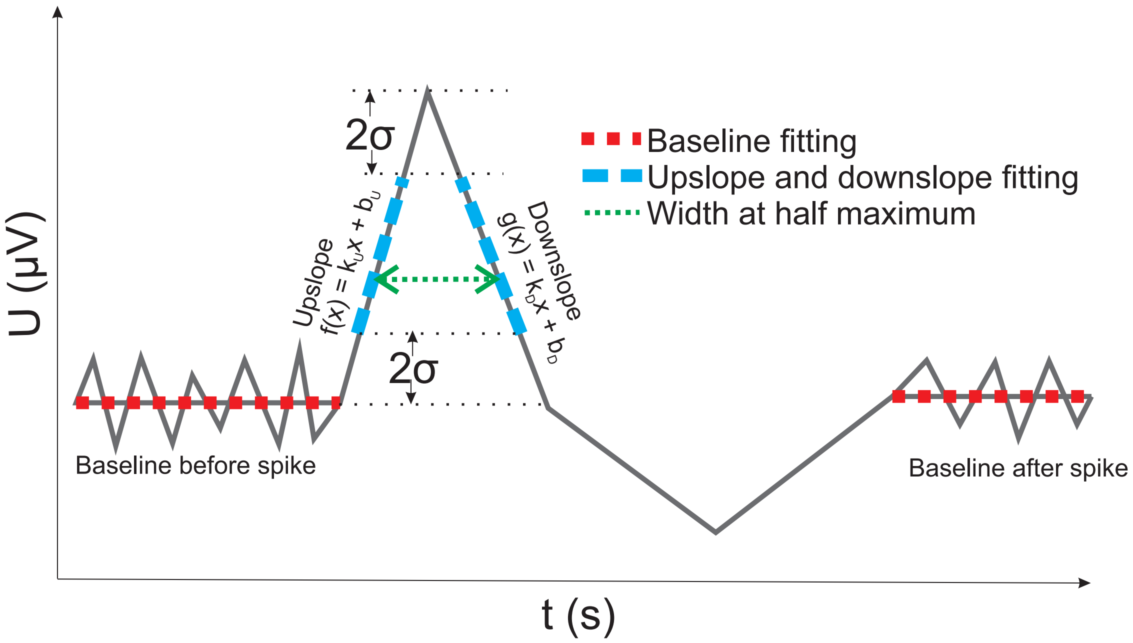


1 pav.: Matematinės morfologijos pagrindu veikiančio filtro veikimo demonstracija. Signalas turi du pikus ties 6.2 ir 7.0 sekundės. A) Nefiltruotas signalas, B) Morfologiniu filtru nufiltruotas signalas.

pagal diagnozę. Algoritmas išmatuoja tokias charakteristikas kaip EEG piko pakilimo kampas (k_U), nusileidimo kampas (k_D), bazinė linija, aštrios bangos ir piko trukmė.

Šiame darbe svarbiausi parametrai yra piko pakilimo ir nusileidimo kampai, todėl lyginant su ankstesniais metais buvo patobulinta jų nustatymo metodika. Abiejuose metodikose piko pakilimo ir nusileidimo kampas yra nustatomas prie atitinkamos piko dalies aproksimuojant tiesę. Pagrindinė kliūtis išmatuoti šį parametrą – neaišku kurioje tiksliai piko dalyje vykdyti aproksimaciją. Todėl buvo peržiūrėtas būtent šis kriterijus: dabar vietoje rišimosi prie piko maksimalios (absoliutinės) vertės, yra rišamasi prie signalo, esančio aplink piką svyravimo (žr. 2 pav.). Dabar imamas dvigubas standartinis nuokrypis nuo signalo aplink piką maksimumų ir atkerpami du šie dydžiai nuo piko apačios ir viršaus, o likusi dalis naudojama aproksimavimui. Tai leido tiksliau nustatyti piko parametrus esant tiek mažiems šalutiniams triukšmams, tiek ir esant dideliems triukšmams.

Kaip jau minėta 1 skyrelyje, pradinė hipotezę buvo, kad vienos EEG eigoje Grupės I pacientų pikai tarpusavyje skiriasi labiau, negu Grupės II pikai. prieš atliekant tolimesnius darbus ši hipotezė buvo patikrinta atliekant gautų EEG pikų parametrus statistiniais metodais. Pastebėta, kad egzistuoja skirtumas tarp EEG piko pakilimo ir nusileidimo kampų, jų vidurkiai yra panašūs, tačiau standartiniai nuokrypiai skiriasi. Toks skirtumas yra per sudėtingas banaliam *if-else* klasifikatoriui, tačiau turėtų būti klasifikuojami kokiais nors algoritmais kaip palaikomųjų vektorių mašinų (SVM) arba ANN pagrindu veikiančiais klasifikatoriais.



2 pav.: EEG piko parametrų vizualinė schema: pakilimo kampas, nusileidimo kampas, signalo bazinė linija ir plotis pusaukštyje

3.2.1 EEG pikų validavimas

Kaip jau minėta 3.2 pradinė EEG pikų parametrų paskirtis buvo jų validavimas. Yra tikrinami tokie parametrai [Sam13]:

- Piko pakilimo ir nusileidimo kampai turi turėti priešingus ženklus, po šios operacijos dirbama tik su absoliutinėmis jų reikšmėmis.
- Piko aštri banga turi trukti tarp 20 ms ir 80 ms
- Visas pikas turi trukti ne daugiau, negu 200 ms.

Tik tie pikai, kurie atitinka visus šiuos kriterijus yra nagrinėjami toliau. Piko pakilimo ir nusileidimo kampų ženklai yra toliau neanalizuojami, nes jie yra priklausomi vien tik nuo EEG matavimo aparato konfigūracijos, todėl neneša jokios svarbios informacijos, todėl patikrinus ar jie yra skirtingi, toliau imamos jų absoliučios reikšmės.

3.3 EEG klasifikavimo algoritmas

Paskutinis EEG klasifikavimo algoritmo žingsnis yra mašinų mokymosi metodų taikymas EEG klasifikavimui. Šiam tikslui buvo išbandyti įvairūs mašinų mokymosi algoritmai, aprašyti 4 ataskaitos skyriuje.

4 Mašinų mokymosi algoritmų taikymas EEG klasifikavimui

4.1 Patikimumo analizė

Nors klasifikavimo patikimumas nėra universalus algoritmo tinkamumo matas, jis yra naudingas atpažinti tam tikroms klasifikavimo algoritmo problemoms. Žemas patikimumas visada reiškia klasifikavimo problemas, aukštas patikimumas – nebūtinai reiškia gerą klasifikavimą. Dėl šios priežasties šiame darbe žemą klasifikavimo patikimumą laikėme kaip pakankamą priežastį algoritmo toliau nebenagrinėti darbe.

4.1.1 LDA klasifikatorius

Tiesinė diskriminantinė analizė (LDA) yra vienas iš klasikinių klasifikavimo algoritmų. Dėl šios priežasties ji buvo išbandyta šiame darbe, tačiau suteikė mažą patikimumą (53%) todėl toliau nagrinėta nebuvo.

4.1.2 Logistinė regresija

Logistinė regresija nėra tradicinis klasifikavimo algoritmas, kaip visi kiti, kurie aprašomi šiame darbe. Tai yra regresinis algoritmas, kas reiškia, kad yra nustatoma ne kiekvieno objekto klasė, o tikimybės, kad minėtas objektas priklauso vienai ar kitai klasei. Tačiau nustatyta tikimybė gali būti labai paprastai redukuojama į klasę: objektas gali būti priskirtas klasei, kurios tikimybė yra didžiausia. Ši procedūra buvo atlikta šiame darbe ir nustatyta, kad logistinės regresijos klasifikavimo patikimumas yra 59%, dėl ko ši metodika toliau nagrinėta šiame darbe nebuvo.

4.2 Grupės I ir Grupės II aptikimo patikimumo analizė

4.2.1 SVM klasifikatorius

Atsižvelgiant į tai, kad SVM yra klasikinis binarinio klasifikavimo uždavinio sprendimo būdas, pirmiausia buvo išbandytas būtent šis klasifikatorius. Pradinė SVM klasifikatorius versija tepasiekė 50-52% patikimumą, todėl buvo išbandytos įvairios SVM klasifikatoriaus konfigūracijos.

Buvo atlikti bandymai su įvairiais SVM klasifikatoriaus branduoliais:

1. Tiesiniu (N=1)
2. Kvadratiniu (N=2)
3. Kubiniu (N=3)
4. Sigmoidiniu
5. Radial Base Function (RBF)

Taip pat buvo išbandytos įvairios baudos parametro reikšmės:

- $C = 0.1$
- $C = 1$
- $C = 10$
- $C = 100$
- $C = 1000$

Atlikus eksperimentus buvo nustatyta, kad SVM klasifikatorius su RBF ir sigmoidiniu branduoliais suteikia 75% patikimumą, tačiau to pasiekia "sukčiaudami". 75% visų turimų EEG pikų yra I Grupėje (žr. 2), o minėtos SVM konfigūracijos klasifikuoja visas įvestis kaip Grupę I. Iš polinominių branduolių, nustatyta, kad aukščiausią patikimumą suteikia kubinis branduolys, dėl ko tik ši SVM konfigūracija buvo nagrinėta toliau.

4.2.2 Naivus Bayes

Atlikus eksperimentus su naiviu Bayes klasifikatoriumi nustatyta, kad jis visas įvestis klasifikuoja kaip Grupę I, pasiekdamas 75% patikimumą. Dėl šio akivaizdaus trūkumo šis algoritmas toliau nagrinėtas nebuvo.

4.3 Kitų klasifikavimo metrikų analizė

Šiame ataskaitos skyriuje pateikiami duomenis apie kitas mašinų mokymosi metodikas, kurios nebuvo atmestos dėl akivaizdžių problemų kaip pvz. LDA ir logistinė regresija (žr. 4.1 skyrelį), naivus Bayes ir kai kurios SVM konfigūracijos (žr. 4.2 skyrelį). Svarbiausios klasifikatorių tinkamumo metrikos pateikiamos 1 lentelėje.

4.3.1 AdaBoost

AdaBoost algoritmas yra klasifikavimo metaalgoritmas. AdaBoost pasižymi didžiausiu patikimumu, tačiau prastu Grupės II aptikimu (52%). Kadangi turime žymiai mažiau Grupės II duomenų (25%) lyginant su Grupės I duomenimis, matuojant bendrą patikimumą, jame dominuoja Grupės I aptikimo patikimumas. Tačiau kuriamame algoritme svarbu tiek Grupės I tiek Grupės II aptikimas, dėl ko šis algoritmas nėra tinkamas pasirinkimas EEG klasifikavimui pagal diagnozę.

Vis dėl to AdaBoost metodika galėtų būti pritaikyta padidinat kokio nors kito algoritmo klasifikavimo patikimumą, kas yra svarstoma šiais mokslo metais.

4.3.2 SVM su kubiniu branduoliu

SVM su kubiniu branduoliu kenčia nuo tos pačios problemos kaip AdaBoost algoritmas, dėl ko taip pat nėra tinkamas EEG klasifikavimui pagal diagnozę.

1 lentelė: Klasifikatorių tinkamumo metrikos [SW17] algoritmams klasifikuojant pagal 100 EEG pikų. Idealaus klasifikatoriaus laukas parodo vertes, kurias surinktų teorinis idealus klasifikatorius.

Metrika/ Algoritmas	**Atsitiktinis miškas**	**Sprendimų medis**	**Labai atsitiktinis medis**	**AdaBoost**	**ANN**	**SVM N=3**	**Idealus klasi– fikatorius**
Patikimumas	0.78	0.76	0.80	0.81	0.75	0.69	**1.00**
TNR	0.79	0.76	0.83	0.90	0.79	0.79	**1.00**
TPR	0.74	0.77	0.71	0.52	0.74	0.48	**1.00**
F1 metrika	0.76	0.76	0.75	0.64	0.78	0.57	**1.00**
ROC AUC	0.53	0.49	0.56	0.69	0.64	0.49	**1.00**
Cohen kappa	0.06	-0.01	0.12	0.38	0.28	0.26	**1.00**
Hamming loss	0.48	0.52	0.45	0.32	0.37	0.38	**0.00**
Jaccard simmilarity score	0.52	0.48	0.55	0.68	0.63	0.62	**1.00**
Log loss	16.72	17.92	15.55	11.00	12.93	12.96	**0.00**
Matthews correlation coefficient	0.07	-0.01	0.15	0.42	0.38	0.28	**1.00**
Recall metrika	0.78	0.76	0.81	0.84	0.78	0.69	**1.00**
Zero one loss	0.48	0.52	0.45	0.32	0.25	0.38	**0.00**

4.3.3 Cohen kappa metrikos analizė

Atlikus analizę iki šio momento yra likę keturi algoritmai: atsitiktinis miškas, sprendimų medis, labai atsitiktinis medis ir ANN. Šių algoritmų patikimumas panašus, tačiau skiriasi Cohen kappa metrika. Cohen kappa metriką interpretuoti be konteksto yra gana sudėtinga, tačiau esant panašiam patikimumui ši užduotis yra lengvesnė. Šios metrikos reikšmė priklauso nuo tikimybės atspėti teisingą atsakymą.

Atsitiktinis medis, sprendimų medis ir labai atsitiktinis medis pasižymi gana aukštu patikimumu, tačiau mažomis Cohen kappa metrikos reikšmėmis. Tai leidžia daryti išvadą, kad minėti klasifikatoriai turimą klasifikavimo problemą traktuoja ne kaip klasifikavimo problemą, o kaip nesubalansuotos monetos metimą ir tokiu būdu teisingą atsakymą atspėja. Taigi, darytina išvada, kad geriauias EEG klasifikatorius pagal diagnozę yra dirbtinis neuroninis tinklas (ANN).

4.4 ANN klasifikatorius

Atsižvelgiant į tuos faktus, kad SVM ir kiti klasfikatoriai bei ANN klasifikatorius tokio tipo problemai veikia geriau [PM08], šiai problemai buvo pritaikytas ANN pagrįstas klasifikatorius.

Buvo sukurtas vieno paslėpto sluoksnio pereptroninis ANN be atgalinio klaidų sklidimo (angl. *backpropagation*). Įvesties sluoksnį sudarė lygiai tiek neuronų, kiek kiekvienu atveju buvo įvesties parametrų, paslėptame sluoksnyje buvo 20 neuronų, išvesties sluoksnį sudarė 1 neuronas (nes buvo vienintelis išviesties parametras 0 - Grupei I ir 1 - Grupei II). Apmokymui buvo naudotas normuotas konjugacinis gradientas (angl. *Scaled conjugate gradient*) algoritmas.

Kaip aktyvavimo funkcija pradžioje buvo naudota sigmoidinė funkcija, tačiau susidurta su tuo, kad šia funkcija pagrįsti neuronai apmokomi lėtai. ANN apsimokydavo tik per 300-500 epochų, galutinis patikimumas gana stipriai keisdavosi priklausomai nuo pradinių svorių, kurie buvo pasirinkti atsitiktinai. Sigmoidinio neurono lėtas mokymasis yra gerai žinoma problema [Nie15], todėl buvo pritaikyta kros entropijos kainos funkcija (kuri yra minimizuojama) [Nie15]:

$$C = -\frac{1}{n}\sum_x \Big[y \ln a + (1-y)\ln(1-a)\Big], \tag{1}$$

čia n yra mokymosi imties dydis, sumuojama per visus mokymosi imties elementus x, y yra norimas rezultatas ir a aktyvavimo funkcija. Buvo pritaikyta taip vadinama softmax aktyvavimo funkcija [Nie15]

$$a_j^L = \frac{e^{z_j^L}}{\sum_k e^{z_k^L}}, \tag{2}$$

kaip aktyvavimo funkcija j-ajam neuronui L-ajame sluoksnyje, čia

$$z_j^L = \sum_k w_{jk}^L\, a_k^{L-1} + b_j^L, \tag{3}$$

w_{jk}^L yra svoris nuo k-tojo neurono praėjusiame sluoksnyje ($L-1$-ajame) į j-ąjį neuroną esamame (L-ajame) sluoksnyje, b_L^j yra j-ojo neurono slenkstinis parametras (angl. *bias*) L-ajame sluoksnyje, a_L^j aktyvacijos funkcija j-ojo neurono L-ajame sluoksnyje, a_k^{L-1} yra aktyvacijos funkcija k-ajam neuronui ankstesniame ($L-1$) sluoksnyje. Eksperimentai įvairiomis aktyvavimo funkcijomis (pvz. sigmoidine) parodė, kad (1) ir (2) kombinacija veikia geriausiai.

Efektyvus ANN mokymasis šiame darbe yra svarbus, nes daugėjant tiriamos grupės pacientų ir kartu didėjant duomenų imčiai jis bus mokomas iš naujo, ateityje neatmetama pridėti ir daugiau pacientų grupių.

5 Morfologinio filtro optimizavimas

Nors daroma išvada, kad EEG galima klasifikuoti pagal diagnozę, tačiau pasiektas 73% patikimumas nėra itin didelis. Viena to priežasčių yra ta, kad turimas pikų aptikimo algoritmas vis dar duoda tiek signalo artefaktų detekcijų kaip pikų, tiek ir neaptiktų EEG pikų. Dėl šios priežasties buvo nutarta pabandyti patobulinti morfologinį EEG pikų aptikimo filtrą. Ši ataskaitos dalis atskleidžia preliminarius, dar nepublikuotus darbus, dėl to nebus pateikiamos visos detalės (nes rezultatai dar gali keistis atliekant tolimesnius eksperimentus).

EEG pikų aptikimo algoritmas [MMJ15] turi kelis parametrus. Svarbiausi jų yra struktūrinio elemento (parabolės) ilgis laike ir detekcijos riba, kurie buvo optimizuoti ataskaitos autoriaus rankiniu eksperimentiniu būdu. Taip pat parabolės koeficientų apskaičiavimo formose egzistuoja daugikliai 0.5 ir 1.5, kurių egzistavimas nėra gerai paaiškintas pirminių algoritmo autorių [JBBS11]. Reikėtų pastebėti, kad nuo šių parametrų reikšmių priklauso EEG pikų detekcijos algoritmo veikimo patikimumas. Dėl šios priežasties nutarta minėtus parametrus optimizuoti pasinaudojant genetiniu algoritmu.

5.1 Optimizavimo metrikos parinkimas

Tam, kad atlikti optimizavimą reikia pasirinkti tam tikrą reikšmę, kuri bus optimizuojama. EEG pikų aptikimo algoritmo rezultatas yra EEG laikų sąrašas, kuriuose algoritmas aptiko galimus pikus. Kaip minėta, dalis EEG pikų gali būti aptikti neteisingai, o dalis gali būti praleista. Dėl šios priežasties optimizuojamo parametro pasirinkimas nėra akivaizdus.

Buvo svarstytos kelios metrikos, pagal kurias būtų galima optimizuoti. Paprastas patikimumas turi vieną esminę problemą: optimizuojant pagal jį, gali laimėti algoritmas, aptinkantis per daug pikų, nes tokiam algoritmui didelė tikimybė aklai pataikyti į teisingą piko vietą, kartu sugeneruojant labai daug neteisingai teigiamai aptiktų pikų. Antra svarstyta metrika buvo specifiškumas (kuri dalis aptiktų EEG pikų yra teisingai aptikti), tačiau čia egzistuoja atvirkštinė problema, kai galima aptikti kelis aiškiai išsiskiriančius pikus, praleidžiant labai didelę jų dalį ir pasiekiant mažą patikimumą.

Dėl to buvo įvesta metrika, apibrėžta kaip patikimumo ir specifiškumo skaičių minimumas, ir ji buvo optimizuojama. Tokiu būdu pasiektas geras kompromisas tarp patikimumo ir specifiškumo optimizavimo, nes pagal šią metriką prastais būtų laikomi anksčiau minėti spėliojimo algoritmai.

Ateičiai svarstomos ir kitos metrikos, atsižvelgiančios į teisingai teigiamai ir teisingai neigiamai aptiktus EEG pikus, kaip F1 metrika.

5.2 Optimizavimo algoritmo pasirinkimas

Turint optimizuojamą metriką, svarbu pasirinkti tinkamą optimizavimo algoritmą. Daugelis paprastų ir gerai žinomų algoritmų kaip gradientinis nusileidimas ar nusileidimas pakoordinačiui yra godūs ir suras lokalų minimumą parametrų erdvėje. Be to jie reikalauja, kad analizuojama funkcija būtų glodi, bent vieną kartą diferencijuojama ir turi kitų reikalavimų.

Mes nežinome jokių analizinių savo tiriamos funkcijos (metrikos) savybių, ji gali būti kad ir osciliuojanti, be to norime surasti globalų (o ne lokalų) optimizuojamos metrikos maksimumą. Tokiais atvejais dažnai yra taikomi “paskutinės pagalvos” algoritmai, kaip genetinis algoritmas.

Negana to, vieno optimizuojamos reikšmės apskaičiavimo laikas yra labai ilgas - apie 5 minutes skaičiuojant nuosekliai – tokiais atvejais taip pat dažnai būna naudojamas genetinis algoritmas. Taip pat genetinis algoritmas gali būti lengvai išlygiagretintas paprasčiausiai kiekvieną populiacijos individą (ar jų grupę) atiduodant vertinti atskiriam procesui VU MIF superkompiuteryje.

Dėl šių priežasčių buvo nuspręsta naudoti genetinį algoritmą morfiloginio filtro parametrų optimizavimui.

5.3 Preliminarūs optimizavimo rezultatai

Čia aprašyti rezultatai yra preliminarūs ir gali bet kada pasikeisti, aprašoma ataskaitos rašymo metu turima situacija.

Optimizavimui buvo panaudotos 6 Grupės I EEG su gydytojos neurologės rankiniu būdu sužymėtais EEG pikais. Ateityje šią imtį planuojama plėsti, ypač jei optimizavimas pasirodys sėkmingas. Jie buvo lyginti su EEG pikų aptikimo algoritmo aptiktu pikų sąrašu. Atsižvelgiant į tai, kad gydytoja piko centro vietą gali pažymėti su nedidele paklaida, mūsų algoritmas taip pat gali turėti tam tikrų paklaidų piko centro aptikimo atžvilgiu. Dėl to pikai buvo laikomi sutampančiais, jei jie jų laikai skyrėsi ne daugiau, kaip 0.05 sekundės. Pagal sutampančių ir praleistų pikų skaičių buvo apskaičiuotas patikimumas ir specifiškumas.

Kadangi optimizuojamas yra morfologinis algoritmas, praleidžiamas EEG pikų validacijos procesas (žr. 3.2.1 skyrelį). Minėto proceso tikslas yra eliminuoti neteisingai teigiamai aptiktus EEG pikus, todėl jo naudojimas potencialiai gali iškreipti atliekamo eksperimento rezultatus.

Preliminarūs rezultatai rodo, kad EEG pikų aptikimo rezultatai pagerėja apie 10%, tačiau rezultatai skiriasi mažiau grąžinus validavimo žingsnį. Vis dėl to planuojama atlikti didesnės skalės eksperimentą VU MIF superkompiuteryje, nes šiuo metu buvo atliekami eksperimentai su gana mažomis imtimis.

6 Rezultatų aptarimas

Šių mokslo metų darbai buvo aprašyti dviejuose moksliniuose straipsniuose. Šiais ir ankstesniais metais kurtas EEG klasifikavimo algoritmas buvo aprašytas moksliniame straipsnyje "Algorithm for automatic EEG classification according to the epilepsy type: benign focal childhood epilepsy and structural focal epilepsy", priimtame į "Biomedical Signal Processing and Control" žurnalą 2018 spalio 6 dieną (A priedas). Taip pat buvo perskaitytas mokslinis pranešimas tema "Accuracy of different machine learning type methodologies for EEG classification by diagnosis" tarptautinėje recenzuojamoje konferencijoje "The 9th International Conference on Numerical Methods and Applications", vykusioje 2018 rugpjūčio 20-24 dienomis Borovets, Bulgarijoje. Taip pat buvo pateiktas, teigiamai recenzuotas ir priimtas straipsnis ta pačia tema ISI indeksuotoje konferencijos medžiagoje (B priedas).

Patvirtinta ankstesnių metų išvada, kad EEG galima klasifikuoti pagal diagnozę 73% tikslumu. Geriausias klasifikatorius šiai užduočiai atlikti yra ANN (dirbtinis neuroninis tinklas), nustatyta, kad pagal 1 lentelėje pateikiamus duomenis, ANN klasifikatorius yra tinkamiausias atlikti EEG klasifikavimą pagal diagnozę.

Nustatyta, kad SVM su sigmoidiniu ir RBF branduoliais ir naivus Bayes klasifikatorius visas įvestis klasifikuoja kaip Grupę I, dėl ko yra netinkami naudoti šiai problemai.

Nustatyta, kad SVM su tiesiniu ir kvadratiniu branduoliais, LDA ir logistinė regresija suteikia mažesnį, negu 60% klasifikavimo patikimumą, dėl ko yra netinkami naudoti spręsti šiai problemai.

Nustatyta, kad SVM su kubiniu branduoliu bei AdaBoost algoritmai pasižymi žemu Grupės II detekcijos patikimumu, dėl ko yra netinkami naudoti spręsti šiai problemai.

Nustatyta, kad atsitiktinis miškas, sprendimų medis ir labai atsitiktinio medžio algoritmai pasižymi prasta Cohen kappa metrika. Atsižvelgiant į tai, kad šie algoritmai pasižymi gana aukštu patikimumu, galima daryti išvadą, kad šių algoritmų atsakymai yra panašūs į spėliojančio (žiūrinčio į problemą kaip į nesubalansuotos monetos metimo problemą) algoritmo veikimą. Dėl to darytina išvada, kad šie algoritmai teisingą atsakymą atspėja ir yra netinkami spręsti turimai klasifikavimo problemai.

Atlikti preliminarūs tyrimai optimizuojant EEG pikų aptikimo algoritmo parametrus genetiniu algoritmu, šiuos tyrimus planuojama tęsti toliau.

Literatūros sąrašas

[JBBS11] A. Juozapavičius, G. Bacevičius, D. Bugelskis, and R. Samaitienė. EEG analysis – automatic spike detection. *Nonlinear Analysis: Modelling and Control*, 16(4):375 – 386, 2011.

[MMJ15] A. V. Misiukas Misiūnas, T. Meškauskas, and A. Juozapavičius. On the implementation and improvement of automatic EEG spike detection algorithm. *Proc. of the Lithuanian Mathematical Society*, 56(Ser. A):60–65, 2015.

[MMS16] A. V. Misiukas Misiūnas, T. Meškauskas, and R. Samaitienė. Derivative parameters of electroencephalograms and their measurement methods. *Proc. of the Lithuanian Mathematical Society*, 57(Ser. A):47–52, 2016.

[Nie15] Michael A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015.

[NNIS99] Shigeto Nishida, Masatoshi Nakamura, Akio Ikeda, and Hiroshi Shibasaki. Signal separation of background EEG and spike by using morphological filter. *IFAC Proceedings Volumes, 14th World Congress of IFAC*, 32(2):4301–4306, 1999.

[PM08] L. M. Patnaik and Ohil K. Manyam. Epileptic EEG detection using neural networks and post-classification. *Computer Methods and Programs in Biomedicine*, 91(2):100–109, 2008.

[Sam13] Rūta Samaitienė. Rolando epilepsija sergančių vaikų EEG pakitimų, miego bei elgesio sutrikimų ir klinikiniųcharakteristikų sąsajos. Vilniaus universitetas, 2013.

[SW17] Claude Sammut and Geoffrey I. Webb. *Encyclopedia of Machine Learning and Data Mining*. Springer, 2017.

[XWZ+07] Guanghua Xu, Jing Wang, Qing Zhang, Sicong Zhang, and Junming Zhu. A spike detection method in eeg based on improved morphological filter. *Computers in Biology and Medicine*, 37(11):1647 – 1652, 2007.

[XWZZ06] G. Xu, J. Wang, Q. Zhang, and J. Zhu. EEG spike detection algorithm using morphological filter. *Proceeding of the 2006 IEEE International Conference on Automation Science and Engineering Shanghai, China, October 7–10, 2006*, pages 170 – 175, 2006.

A Priedas

Algorithm for automatic EEG classification according to the epilepsy type: benign focal childhood epilepsy and structural focal epilepsy

Andrius Vytautas Misiukas Misiūnas[a,*], Tadas Meškauskas[a], Rūta Samaitienė[b,c]

[a] Institute of Computer Science, Faculty of Mathematics and Informatics, Vilnius University, Didlaukio 47, LT-08303 Vilnius, Lithuania
[b] Children's Hospital, Affiliate of Vilnius University Hospital Santaros Klinikos, Santariškių 7, LT-08406, Vilnius, Lithuania
[c] Clinic of Children's Diseases, Faculty of Medicine, Vilnius University, Santariškių 4, LT-08406 Vilnius, Lithuania

Abstract

Rationale: It is still not clear if there are EEG parameters that may be related to the epilepsy etiology in epilepsies presenting with rolandic spikes. Rolandic spikes are not pathognomonic for rolandic epilepsy and could be related to the area of discharges itself. The initial hypothesis was that even visually identical spikes have some difference, because of the different etiology.

Objective: The aim of the study was to find the differences in rolandic spike morphology in two epilepsy groups, different by etiology, but presenting with visually identical spikes.

Methods: A novel algorithm for automatic classification of interictal electroencephalogram (EEG) rolandic spikes according to the epilepsy type (Group I – patients with benign focal childhood epilepsy, self-limiting, with no causal lesion in the brain, Group II – patients with structural focal epilepsy) is proposed. The algorithm consists of three stages: 1) EEG spike detection, 2) determination of EEG spike parameters, 3) classification of EEG by epilepsy type based on estimated spike parameters. Automatic classification method is defined by artificial neural network. The algorithm has been trained and tested on a large data sample provided by Children's Hospital, Affiliate of Vilnius University Hospital Santaros Klinikos. Only those EEGs that were visually identical and inaccessible for manual clustering to the groups according the visual spike morphology and contained 50 or more spikes have been analyzed. Training and testing pools have been selected as non overlapping (containing different patients) data sets.

Results: The proposed methodology let us to achieve up to 75% of accuracy of classification of EEG.

Keywords: EEG, Epilepsy, Epileptiform discharge, Spike, Machine learning, Artificial neural network

1. Introduction

In this study, we attempt to apply machine learning techniques for finding the differences in rolandic spike morphology in two epilepsy groups, different by etiology, but presenting with visually identical spikes. There are multiple studies relevant to classifying EEGs into some categories like detecting normal, interictal and epileptic signals [1–3], or signals obtained from alcoholic vs non-alcoholic patients [4]. However, no studies (to our knowledge) have attempted to develop automatic algorithms for detecting and classifying interictal epileptiform discharge (ED, or spike) based on the epilepsy type.

We have focused on two different types of epilepsy (defined further), which could have the similar electroencephalogram (EEG) pattern. Manually distinguishing this type of spikes is time-consuming, and sometimes impossible (without clinical record of the illness), as spikes are visually similar. Cases, where one can distinguish these two groups by morphology of the discharge manually, are well investigated.

*Corresponding author.
Email address: andrius.misiukas@mif.vu.lt (Andrius Vytautas Misiukas Misiūnas)

The mechanism and precise localization of the functional generator of the rolandic discharges are not yet completely understood. Current studies show that similarities between rolandic discharges of two different epilepsy groups are related to the area of the discharges and not to the epileptic syndrome itself [5–7]. The difference between rolandic discharges could be due to the different origin (genetic versus lesional) and using the spatial analysis only, one could not find the subtle morphological difference.

Automatic detection algorithms mostly focus on detection of EEG spikes [8, 9]. Epileptiform sharps are detected using a variety of algorithms such as mathematical morphology used in this work [10–12], while wavelets and other methods used by other authors are summarised in [8, 9].

We are going to deal with two different groups of patients (for details see Section 2).

Group I is defined as benign childhood epilepsy patients with centrotemporal spikes (in literature reffered as BCECT, rolandic epilepsy or self-limited epilepsy with centrotemporal spikes) exhibiting the EEG hallmark – the benign epileptiform discharge (BED, further in text – spike) frequently appearing as so called rolandic discharge [13]. It is sharp wave or spike detected over the centrotemporal brain regions (T3, T4, C3, C4 regions), but other regions as central, centroparietal or centrofrontal, can also be involved [14]. A number of less well-defined syndromes of idiopatic focal epilepsies (IFE) have been proposed, including benign childhood epilepsy with parietal spikes, benign childhood seizures with frontal or midline spikes [15].

Group II is defined as structural focal epilepsy patients with cerebral palsy, dysplastic brain lesion, gliosis *etc.*, following ILAE (International League Against Epilepsy) criteria [16].

The problem is that benign epileptiform discharges seem not to be pathognomonic for rolandic epilepsy. They may represent both a functional focus [17] and an expression of a focus secondary to an organic brain damage. Both rolandic epilepsy and structural focal epilepsy with a lesion in rolandic area may present with BEDs [6, 13].

It is important to distinguish between these two epilepsy groups because different investigation and treatment strategies are used. Usually it is no need for magnetic resonance imaging (MRI) analysis in typical rolandic epilepsy. But as rolandic discharge is not patognomonic, sometimes one could miss the underlying lesion in the brain.

Many studies analyzed neurophysiologic aspects of rolandic discharges in this context (horizontal dipole discharges, double spike phenomenon, the extension of epileptiform discharges and background activity). There is no consensus about the results. Some studies have found a distinction between these two epilepsy groups, some – not [5, 6, 18–22].

In this study we have confirmed our initial hypothesis, that *even visually similar spikes are not really identical*, because of the different etiology and statistical differences between Group I and Group II spike parameters can be detected by machine learning type methods.

In the previous study [23] we have defined some quantifiable parameters (discussed in Section 3.2) of spikes. In this paper the tests of artificial neural network (ANN) based classifier (see Section 3.3.1) and the statistical analysis (see Table 1, Fig. 6 and Fig. 7) have shown that the array of spike upslopes and downslopes (see Fig. 5), further in text referred as spike array, is significant for EEG classification. Tests involving other parameters (defined in [23]) were also carried out, but such approach did not improve the accuracy of the classifier (for details see Section 3.2), therefore these additional parameters were discarded from the proposed methodology. We have also been experimenting with other (beside ANN) automatic classification methods (linear regression, linear discriminant analysis, support vector machine), but ANN prevailed among them (see Sections 3.3 and 4.2).

Summarising – we have proposed the algorithm (for automatic classification between two different epilepsy groups), based on three consecutive steps (as described in more detail in Section 3: 1) EEG spike detection, 2) EEG spike measurement, 3) ANN based classifier. In Section 4 we show that accuracy rate of the proposed classification methodology (trained and tested on not overlapping inputs, obtained from different patients' EEGs) is up to 75%.

The algorithm has been tested and deployed in high performance parallel computing environments, including supercomputer at Digital Science and Computing Center, Faculty of Mathematics and Informatics, Vilnius University.

2. Data and methods

A retrospective search has been done of the EEG database (provided by Children's Hospital, Affiliate of Vilnius University Hospital Santaros Klinikos) over a 2010 – 2017 period to identify children having benign focal epileptiform discharges (see Fig. 1 and Fig. 2). It has been defined as any focal epileptiform discharge that involved the central regions and mid-central or mid-parietal regions (T3, T4, C3, C4, Cz, Pz).

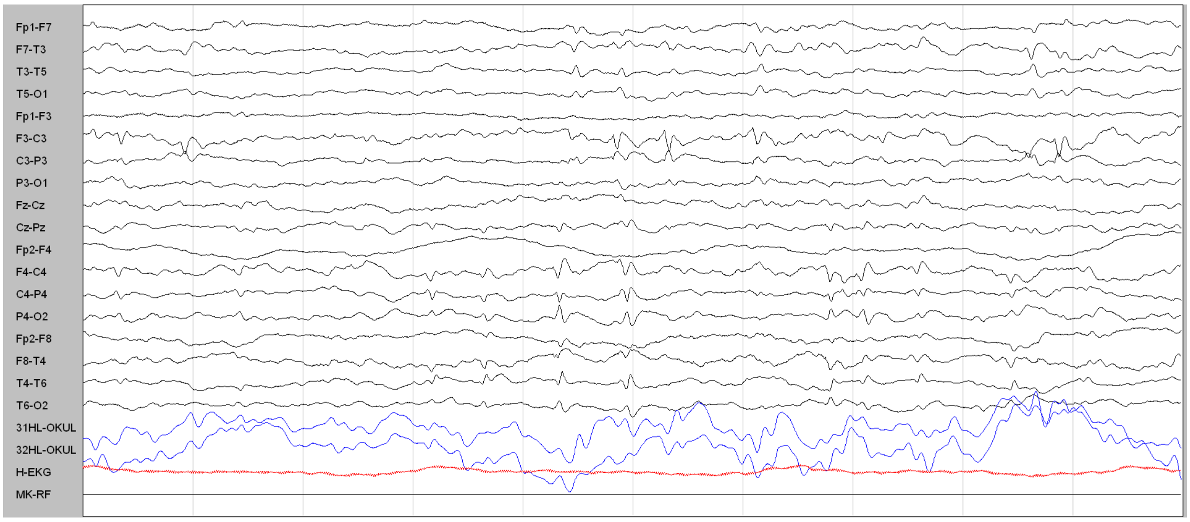


Figure 1: Spikes in C3, P3, T3, and T4 regions in a patient with rolandic epilepsy (Group I). Bipolar longitudinal montage.

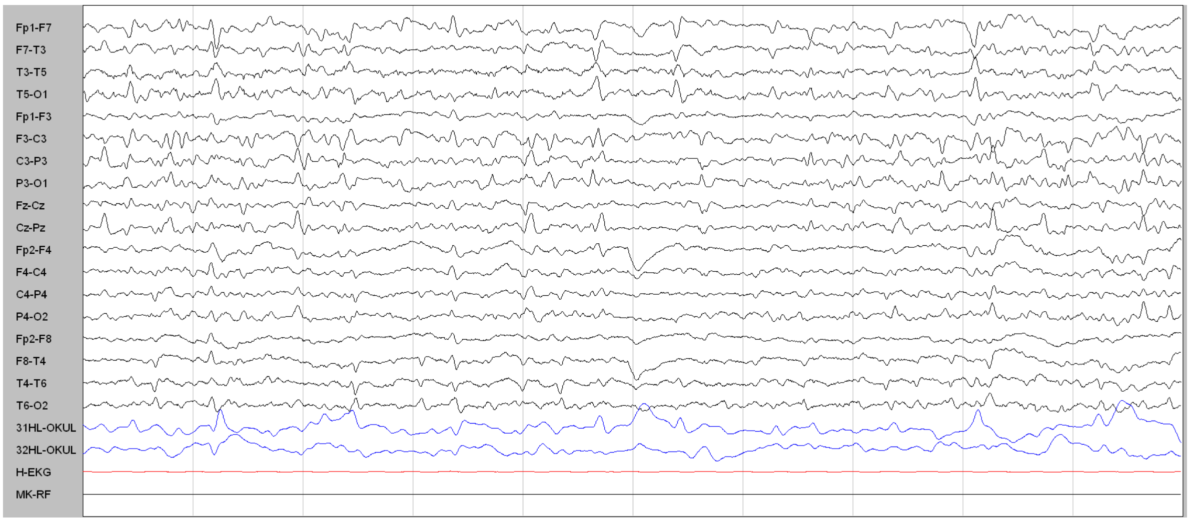


Figure 2: Spikes in T3, C3, P3, Cz, and T6 regions in a patient with cerebral palsy (Group II). Bipolar longitudinal montage.

Focal discharges that had maximal negativity at the mid-temporal regions (T5 or T6) or the biparietal regions (P4 or P3) were also included if the discharges had constant, stereotyped waveform, spread to central regions, no diffuse or localized slowing (see Fig. 3). Inclusion criteria were:

(a) hard to distinguish (visually identical) benign focal epileptiform discharges in benign focal childhood epilepsy and rolandic like discharges in structural focal epilepsy;
(b) exact diagnosis known from clinical records;

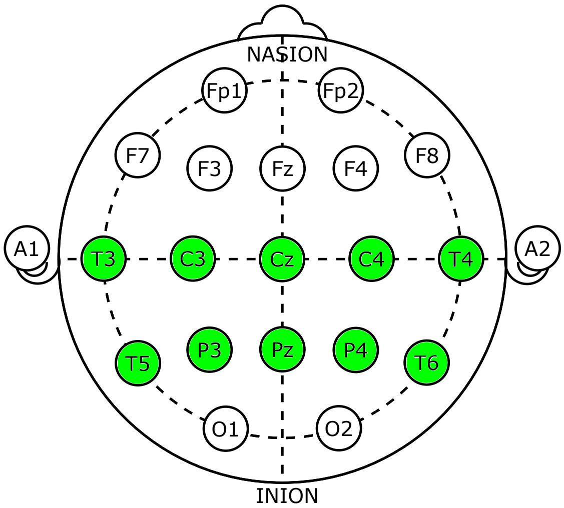


Figure 3: Schematic visualization of international 10 – 20 EEG system [24]. Electrodes over central regions and mid-central or mid-parietal regions (T3, T4, C3, C4, Cz, Pz, P3, P4, T5, T6) investigated in this study are highlighted in green color.

(c) artifact-free EEG recording of 2 – 13 min, we excluded physiologic and extraphysiologic artifacts, according to qualitative and quantitative criteria[1] for EEG artifacts, see [25].

(d) $\geqslant 50$ spikes in the raw EEG, none of which should be preceded or followed by artifacts.

EEG was performed according to the international 10 –20 system (see Fig. 3), with an average referential (Medtronic, Brain Explorer, 28 amplifier system, Galileo NT DEEG software). Interictal EEGs lasted for at least 20 min and were performed with the patient awake and asleep. We used a sampling rate of 200 Hz with the bandwidth set at 0.3 Hz to 70 Hz.

The proposed algorithm (for automatic classification between two different epilepsy groups) has been implemented with Python programming language, employing Scikit-learn, NumPy, SciPy, OpenCV, Mpi4Py and EEGTools libraries. The methodology was tested independently using MATLAB Pattern Recognition toolbox and Python Scikit-learn framework.

2.1. EEG selection

In this research we have processed 94 EEGs of 86 different patients (3 – 17 years old), divided into two groups:

- **Group I:** benign childhood epilepsy with centrotemporal spikes (62 EEGs or about 66% of all EEGs), 35 ($\approx$ 56.4%) of them boys;
- **Group II:** structural focal epilepsy patients with cerebral palsy, dysplastic brain lesion, gliosis *etc.* (32 EEGs, or about 34% of all EEGs), 18 ($\approx$ 56.3%) of them boys.

[1]**Criteria for EEG artifacts.** Electromyogram (muscle): waveforms with a frequency of 20 – 100 Hz. Tong movement artifact: potential field from frontal to occipital areas, amplitude is greater inferiorly than in parasagittal regions, frequency is variable, usually in the delta range. A blinking artifact: vertical eye movement, symmetric downward deflections in frontopolar (Fp1 – Fp2) electrodes. Downward eye movement produces an upward deflection best recorded in electrodes near the eye. Lateral eye movements mostly affect lateral frontal electrodes F7 and F8. The phase reversals at electrodes F7 and F8 are of opposite polarity. ECG artifact is recognized by its rhythmicity/regularity and coincidence with the ECG tracing. Pulse artifact – QRS complexes are recorded slightly ahead of the pulse waves (200 – 300 millisecond delay after ECG. Respiration: slow and rhythmic activity, synchronous with the body movements affecting the impedance of electrodes.

It should be emphasized, that we have omitted from further analysis trivial cases: Group II EEGs, whose spikes exhibit visual dissimilarities (easily detectable by neurologists) versus Group I spikes. By including such data into presented methodology one would only decrease the error of automatic classification. The reason for such decision was to demonstrate that discussed methodology is capable to classify even challenging data. Furthermore, cases with obviously different morphology are well investigated already [6, 19, 21, 22].

2.2. ANN training and testing strategy

We applied strict policy of assigning the same patient either to the training or the testing pool. If multiple EEGs (recorded at different times) of the same patient were available, then all of them went into the same (training or testing) set. This has been done in order to prevent biasing the results by ANN over-fitting [26]. An experiment to determine significance of over-fitting effect has also been carried out, resulting in approximately $95-99\%$ pseudo-accuracy of classifier (compared to about $72-75\%$ accuracy of non over-fitted methodology, see Section 4).

Training and testing (of ANN based classifier) pools have been defined in the following way:

- 62 EEGs from Group I, containing $\approx 75\%$ of all spikes – 21 for training (about 34%) and 41 for testing (about 66%);
- 32 EEGs from Group II, containing $\approx 25\%$ of all spikes – 11 for training (about 34%) and 21 for testing (about 66%).

Training dataset has been selected manually. Artifact-free parts of EEG were provided by neurologist to avoid ANN training errors. In visual artifact detection, neurologist visually inspected the data and identified the EEG data segments that are affected by artifacts of physiological and non-physiological or technical origin, caused by patient movements, eye movements, *etc.* These EEG parts were removed from the training data set only. The algorithm was tested on not cleaned data.

Note that the same training and testing strategy has been applied for both ANN and support vector machine (SVM, see Section 3.3.2) based classifiers.

3. Algorithm for EEG processing and automatic classification

In this section we discuss the three steps of the proposed EEG processing and classification algorithm in more detail. These steps are:

1. EEG spikes detection, for details see Section 3.1 and references [10–12].
2. Determining EEG spikes parameters, for details see Section 3.2 and our previous work [23].
3. ANN based classifier, for details see Section 3.3.1.

3.1. EEG spike detection

Now we provide some details of the first step of the algorithm. Before morphological filter is applied, the raw EEG signal is filtered with 50 Hz band-stop filter in order to eliminate noise generated by alternating electric current. Then mathematical morphological filter based method (first described in [10], later in [11], with further improvements in [12, 23]) is applied in order to find occurance of spikes in all channels of EEG. The morphological filter is designed in order to filter out known normal brain activity like brain rhythms, movement artefacts, *etc.*

Spike detection is based on combination of morphological filters and operations. The most basic component of morphological filter is erosion and dilation.

In this section we are going to employ following notations: the EEG channel under analysis is denoted by $f(t)$, the structuring element is defined by $g(t)$, while reflection of the structuring element is $g^s(t) = g(-t)$. D denotes the domain of signal $f(t)$. Then erosion can be defined as:

$$(f \ominus g^s)(t) = \min_{\tau \in D}\{f(\tau) - g(-(t-\tau))\}. \quad (1)$$

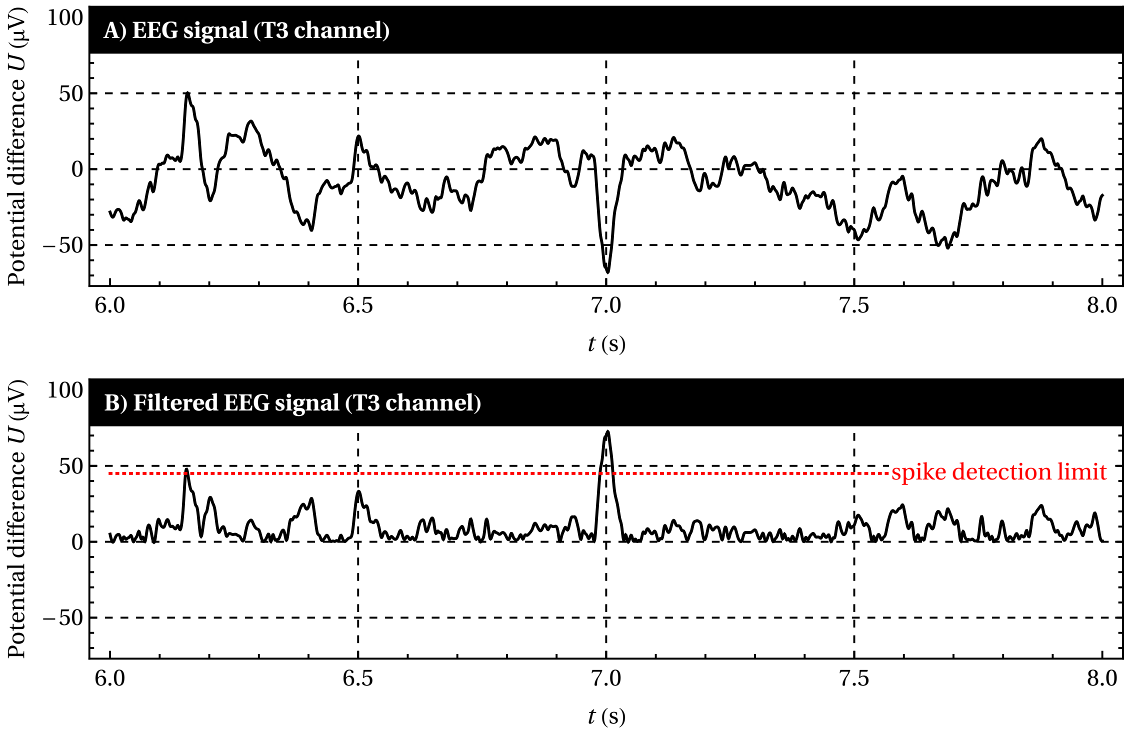


Figure 4: Demonstration of principles behind the morphological filter (for details see [10–12]) employment for spikes detection. The method has detected two spikes at 6.15 and 7 seconds. A) Original EEG at T3 channel. B) The same EEG signal with morphological filter applied – $f_{filtered}(t)$, defined by (11).

Dilation:

$$(f \oplus g^s)(t) = \min_{\tau \in D}\{f(\tau) + g(-(t-\tau))\}. \tag{2}$$

Employing these operators opening operators can be defined:

$$(f \circ g)(t) = [(f \ominus g^s) \oplus g](t). \tag{3}$$

Closing:

$$(f \bullet g)(t) = [(f \oplus g^s) \ominus g](t). \tag{4}$$

EEG spike can have both positive and negative amplitudes, hence both open-closing and close-opening operations are required. Open-closing:

$$OC(f(t)) = f(t) \circ g_1(t) \bullet g_2(t). \tag{5}$$

Close-opening:

$$CO(f(t)) = f(t) \bullet g_1(t) \circ g_2(t). \tag{6}$$

It should be noted that both OC and CO have an impact on average value of signal. The impact has same absolute value but different signs, so averaging out the value eliminates the change:

$$OCCO(f(t)) = \frac{OC(f(t)) + CO(f(t))}{2}. \tag{7}$$

3.1.1. Structuring element

Morphological filters apply a structuring element in order to filter any signal. Since the discussed method is based on filtering out normal brain activity, it is recommended to define the structuring element as two parabolas [11]:

$$g_i(t) = a_i\, t^2 + b_i, \qquad i = 1, 2. \tag{8}$$

g_1 and g_2 have different amount of incisiveness, which allows the filter to follow normal brain activity without dipping into sharp changes characteristic to EEG spikes [27]. Parabolas g_1 and g_2 are employed in both OC and CO operations, see equations (5) and (6), respectively.

Because both frequency (between 0.5 Hz and 100 Hz) and amplitude of brain rhythm vary throughout the EEG, the parameters a_i and b_i vary as well. Therefore a_i and b_i are recalculated for each selected section of EEG channel (for more details see Section 3.1.2). Values of these parameters are defined by

$$a_1 = \frac{2\, Median(|f|)}{Median(W)}, \qquad a_2 = \frac{2\, Median(|f|)}{3\, Median(W)}, \qquad b_1 = b_2 = Median(|f|), \tag{9}$$

here W is an array of EEG signal arc lengths [11].

In the selected section of EEG channel, the parameters of the structuring element are computed only once. Then, the structuring element is applied in all possible subintervals (number of such subintervals is finite, since we deal with discrete data) of the selected section. It was found (experimenting numerically) that optimal width of these subintervals (much shorter than length of the selected section) is equal to $Median(W)$ [12].

3.1.2. Spike detection limit

Sections of EEG channel, where parameters of structuring element (see Section 3.1.1) are recalculated, are defined by moving window.

We set the length (of moving window) 4 seconds (optimal value, based on our numerical analysis [12]). It should be noted that if structuring element is reevaluated less often (note that background activity of patients brain is constantly changing) – we get more errors (false positives and/or false negatives) detecting spikes.

Due to the argument that moving window is relatively short and signal is locally stationary (*i. e.* should not exhibit large changes of the average value), we define spike detection limit

$$L = 2\, Median(f_{filtered}), \tag{10}$$

here

$$f_{filtered}(t) = |f(t) - OCCO(f(t))|\,, \tag{11}$$

denotes the filtered signal. In other words, spike is detected if it rises more than two times above the background activity of the brain (see Fig. 4).

Further, to eliminate some false detections we require the spike to occur in at least two neighbouring EEG electrodes (see Fig. 3) and apply spike parameter post-processing (see Section 3.2.1).

3.2. Measuring of EEG spike derivative parameters

Here we briefly describe the second step (for details see our previous work [23]) of the algorithm.

In order to distinguish between Group I and Group II patients some quantifiable parameters (obtained by measuring spike geometrical characteristics) are employed. Experiments with measurements of many features (upslope, downslope, width at half maximum, baseline level) of EEG spike were done, but upslope and downslope of EEG spike have prevailed as the most significant parameters in ANN based classifier (see Section 3.3.1).

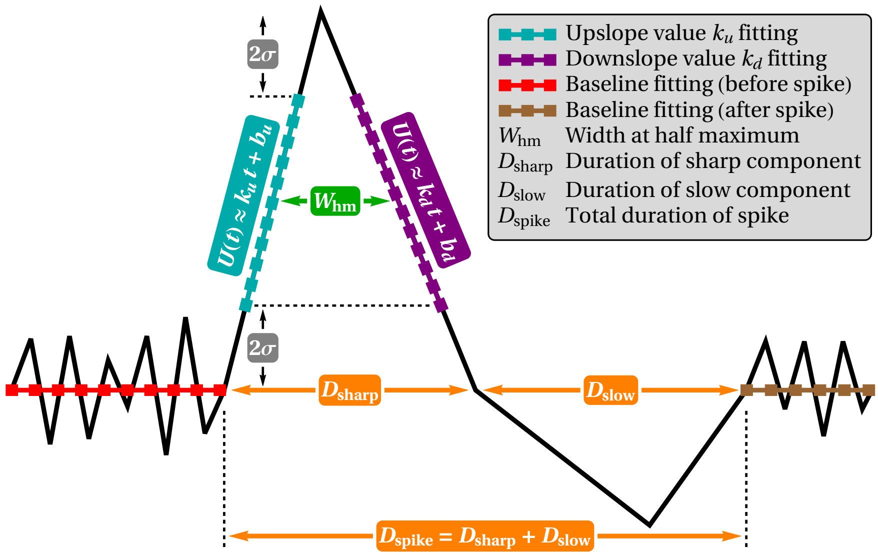


Figure 5: Schematic visualization of EEG spike and main metrics of it: upslope, downslope, baseline level, width at half maximum, durations of sharp and slow components of the spike and total spike duration. σ denotes a standard deviation of signal fluctuations at baseline.

We denote that k_u and k_d are coefficients at variable t in linear dependance $U(t) \approx kt + b$, when line is fitted to respective parts (upslope or downslope) of spike (see Fig. 5).

Upslope value k_u and downslope value k_d are measured by cutting of 2σ from both lower and upper parts of the sharp wave (of EEG spike) and by fitting the line $U(t) = k_u\, t + b_u$ or $U(t) = k_d\, t + b_d$, respectively (as shown in Fig. 5). Here σ is a standard deviation of signal fluctuations at baseline neighbourhood. Note that the signs of both upslope and downslope values are dependent purely on EEG measurement apparatus configuration. The values k_u and k_d should have different signs [28] (otherwise the spike is considered falsely detected and rejected from further analysis). Therefore only absolute values of upslope and downslope are dealt with further.

More detailed description of spike parameters measurement methods is available in our previous paper [23].

3.2.1. Post-processing of EEG spikes detections

The initial spike detection algorithm still had some problems – for example it might detect patient's rapid eye movement as EEG spike. This is remedied by post-processing as described in [23]. The basic idea behind the post-processing is to eliminate false detections of spikes based on their estimated characteristics (see Section 3.2). These characteristics are validated against their neurological properties:

- Upslope and downslope values (k_u and k_d) should have different signs;
- Duration of sharp component of the spike (D_{sharp}) should be between 20 ms and 80 ms;
- Total duration of the spike ($D_{spike} = D_{sharp} + D_{slow}$) should not exceed 200 ms [29].

Only the spikes meeting all three above conditions are considered detected successfully and the values of upslopes as well as downslopes (computed from these spikes) are included in further analysis. Such post-processing procedure improves specificity of EEG spike detection algorithm.

3.3. Machine learning type methods for automatic classification of EEGs

Now we describe the third (final) step of the proposed algorithm.

As signs of upslope and downslope values are omitted (see Section 3.2) – from here we denote (for the sake of convenience) $|k_u|$ as k_u and $|k_d|$ as k_d and refer to them as upslope and downslope values, respectively.

At first some relatively simple (less sophisticated than ANN or SVM) classification (between Group I and Group II, employing lists of k_u and k_d, analogously as in Section 4.1) methods were given a trial – with almost no success. Accuracy achieved by *linear regression* was $\approx 53\%$, while *linear discriminant analysis* (LDA) resulted in $\approx 59\%$ accuracy. Considering that we have more EEGs (and therefore detected spikes) in Group I than in Group II, accuracy alone can be misleading (other metrics, discussed in Section 4.2, should be analyzed too), however poor accuracy already tells that algorithm tested is not suited for the task at hand [30]. This led us to experiment with more advanced classifiers, presented in following sections.

3.3.1. ANN based classifier

A feed-forward perceptron based artificial neural network without backpropagation [26] has been applied in order to classify EEGs between Group I and Group II, using parameters of their spikes. The classifier has been defined by an input layer, a single hidden layer (including 20 hidden neurons), and an output layer (including 1 neuron), scaled conjugate gradient as a training algorithm (for minimization of some cost function), and cross entropy

$$C(W,B) = -\frac{1}{n}\sum_x \Big(y \ln a_j^L + (1-y)\ln(1-a_j^L)\Big), \tag{12}$$

as the cost function [26]. Here n is length of the list, containing training data, the sum is over all training inputs x, also y is the corresponding desired output and a is the activation function. The cost function is minimized in order to find such weights (W) and biases (B) for all neurons (described further) that give the lowest value of C. The cost function is only employed in training of the ANN, it is not used during testing of the ANN. We have applied the *softmax* function (normalized exponential function) [26]

$$a_j^L = \frac{e^{z_j^L}}{\sum_k e^{z_k^L}}, \qquad L = 2,3, \tag{13}$$

as the activation function for j-th neuron in L-th layer ($L=2$ in the hidden layer and $L=3$ in the output layer). Here

$$z_j^L = \sum_k w_{jk}^L\, a_k^{L-1} + b_j^L, \qquad L = 2,3, \tag{14}$$

w_{jk}^L is the weight from the k-th neuron in the previous ($L-1$-th) layer to the j-th neuron in the current (L-th) layer, and b_L^j is the bias of the j-th neuron in the L-th layer. Since input layer $L=1$ does not possess activation function,

$$a_j^1 = x_j, \qquad L = 1, \tag{15}$$

here x_j is input of neural network. The number of neurons in the input layer is always equal to the number of input values. For example, when dealing with array defined by 100 spikes, there are: 200 input neurons when using Strategy A (see Section 4.1); 100 input neurons when using Strategy B or C.

To choose optimal ANN structure, performance (see Section 4.2) of various ANN configurations has been tested. Single hidden layer ANNs, defined by $5-100$ neurons in the hidden layer, have been given a trial (with step size of 5 neurons). These experiments have shown that 20, 25 and 30 neurons single hidden layer ANNs perform within statistical margin of error. Therefore, the simplest configuration (ANN with 20 hidden neurons) has been chosen.

It should be noted that we have experimented with some other activation functions also (for example, sigmoid activation function), however the choice (13) has resulted in training within least amount of epochs, without any substantial impact on accuracy of classification.

This quite simple neural network is well suited for binary classification problems (note that we apply binary classification as well), as some related studies suggest, for example, see [4]. Increasing ANN complexity would result in higher computational time, not necessarily achieving better accuracy rate, compared to up to 75% (see Section 4.1) achieved by the classifier, dealt with in this work.

3.3.2. SVM based classifier

Support vector machine (SVM) techniques (*e. g.* [31]) are commonly used for binary classification problems, therefore we given a trial for our task.

After normalizing input data (lists of k_u and k_d, defined in Section 4.1), we have experimented with various configurations of SVM:

- Linear, polynomial (of order $n = 2$ and $n = 3$), sigmoid and radial basis function (RBF) kernels;
- Values of penalty parameter C: 0.1, 1, 10, 100 and 1000.

Linear, sigmoid and RBF kernel SVMs classified all (or almost all) inputs as Group I – therefore were excluded from further analysis. Accuracy achieved by SVM with quadratic kernel ($n = 2$) was $\approx 58\%$, which is far from satisfactory.

SVM with cubic kernel ($n = 3$) and penalty parameter $C = 100$ resulted in $\approx 69\%$ accuracy (see Table 1 in Section 4.2). At first glance, this SVM configuration is at least competitive. However, even if it classified Group I with $\approx 74\%$ accuracy, *accuracy within Group II was poor* – only $\approx 48\%$ (Table 1). Recall (Section 2.1) that we have much larger Group I than Group II. Due to these arguments we must conclude that SVM with cubic kernel fails, too.

Also note, overall performance of SVM with cubic kernel ($n = 3$) is worse than that of ANN (Table 1) – which is consistent with findings of other authors (*e. g.* [4], where ANN and SVM techniques are compared for other type of EEG binary classification task).

For readers with deeper interest in machine learning methods we discuss possible reasons for failure of SVM based classifiers in Section 4.3.

4. Results and discussion

We have proposed a novel three stage algorithm (defined in Section 3) for EEG classification, related to EEG epileptiform discharges and their parameters, according to the epilepsy type: benign focal childhood epilepsy with centrotemporal spikes (Group I) and structural focal epilepsy (Group II).

Our results suggest that the algorithm, based on: 1) spike detection, 2) derivative characteristics (upslopes, downslopes of spike) estimation, 3) machine learning method (ANN), can be employed for classifying EEGs (we discuss the accuracy of this methodology further). The mathematical morphology based EEG spike detection algorithm has already been explored in previous studies [10–12], while employment of EEG spike parameters is a novel approach for distinguishing between patients diagnosed with different epilepsy types (in this study Group I and Group II).

4.1. Accuracy of ANN based automatic classification

In order to evaluate the accuracy rate of the ANN based classifier (see Section 3.3.1) between Group I and Group II patients, we have experimented with the following strategies (defining lists, containing parameters of spikes, see Section 3.2, employed in training and testing of the algorithm):

- **Strategy A:** lists have been defined by upslope and downslope pairs of values $(k_u[1], k_d[1]), (k_u[2], k_d[2]), \ldots, (k_u[N], k_d[N])$;
- **Strategy B:** lists have been defined by the upslopes values $k_u[1], k_u[2], \ldots, k_u[N]$ only;
- **Strategy C:** lists have been defined by the downslopes values $k_d[1], k_d[2], \ldots, k_d[N]$ only.

Here N denotes length of available (obtained from concrete EEG) spike parameters data; N can vary for different EEGs, therefore we also denote by N_{min} the smallest (through all available database of different EEGs) N. In this analysis all the lists have been defined by partitions (of above data) of equal length N_{spikes} ($N_{spikes} \leqslant N_{min}$).

We are now in position to discuss the main result of this paper – represented in Fig. 6. It allows us to conclude that the proposed methodology can be used for classification of benign focal childhood epilepsy

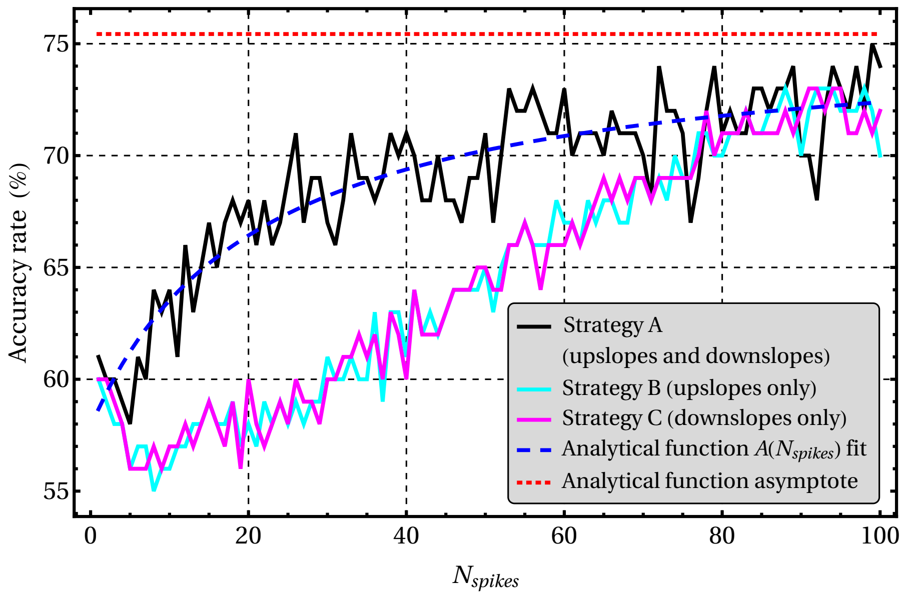


Figure 6: Accuracy of automatic ANN based classification (between Group I and Group II) vs N_{spikes} (length of lists, containing parameters of spikes, employed in training and testing) for different training and testing strategies.

and structural focal epilepsy patient EEGs, reaching up to 75% accuracy. Other conclusions coming from Fig. 6 are discussed in Section 5.

As we see, Strategy A (as compared to Strategy B or Strategy C) yields better or equal accuracy rate, especially with lower values of N_{spikes}. For larger values of N_{spikes} (if $80 \leqslant N_{spikes} \leqslant 100$), we conclude (after eliminating statistical error) that all three strategies reach the same rate of accuracy.

Additionally, to eliminate statistical fluctuations of computed accuracy rate (achieved by Strategy A) we have derived an analytical equation, by fitting (to the values of accuracy, obtained by Strategy A for different values of N_{spikes}) function

$$A(N_{spikes}) = \frac{1}{\dfrac{C_1}{N_{spikes} + C_2} + C_3}, \tag{16}$$

with estimated values $C_1 \approx 0.06492$, $C_2 \approx 16.13$ and $C_3 \approx 0.01326$, obtaining the averaged (with decreased impact of statistical error) rate of accuracy at $N_{spikes} = 100$:

$$A(100) \approx 72.4\%. \tag{17}$$

It is also worth poing out that, if $N_{spikes} \to \infty$ (assuming length of lists, employed in training and testing, mathematically grows unlimitedly), the fitted function $A(N_{spikes})$ tends to an asymptotical value

$$A(N_{spikes}) \to \frac{1}{C_3} \approx 75.4\%. \tag{18}$$

Note that (see Fig. 6 and the estimate (17)) it is sufficient to take $N_{spikes} = 100$ for the discussed classifier to approach its theoretical limit accuracy rate 75.4% (provided that the training data pool, defined in Section 2.2 is used) within statistical range of error $\pm 3\%$. It should be observed that this limit accuracy could increase if more training and testing data would be available.

Also, analyzing accuracy of ANN based automatic classification we have made the following observations:

1. The order of spikes seems to have low impact on the accuracy rate of the classifier. Experiments with randomly ordered spikes did not show significant (larger than standard deviation) drop in the accuracy;

2. Noise (falsely positively detected spikes) increases the classification error;
3. If an EEG has low number of spikes, the algorithm still can be used, however with decreased degree of accuracy. For example, classification accuracy of EEG, containing 20 spikes, is about 67% (employing Strategy A, see Fig. 6).

It should be noted that Group I and Group II define unbalanced data set, we have more EEGs (and therefore detected spikes) in Group I than in Group II (see Section 2.1). Therefore, to show real performance of ANN classifier, other metrics (beside accuracy) must be checked. Only after analysis presented in Section 4.2, we have concluded that ANN can be approved for distinguishing between patients diagnosed with different epilepsy types.

4.2. Performance metrics for evaluation of classifiers

Now we discuss metrics, measuring performance of binary (between Group I and Group II) classification algorithms tested – keeping in mind uneven distribution of patients between Group I and Group II (see Section 2.1).

Accuracy (mere proportion of correct classifications) is the most common metric. A good classifier must show high accuracy, however, high accuracy value does not automatically guarantee real quality of classification – if the data set is unbalanced (that is, when sizes of two classification classes are significantly different, as it is in our case). True negative rate (TNR), true positive rate (TPR) and other performance metrics (presented in Table 1) [30] must be analyzed, too. Note that in our case:

- TNR provides classification accuracy within Group I;
- TPR provides classification accuracy within Group II.

Table 1: Performance metrics for ANN and SVM with cubic kernel ($n = 3$).

	ANN	SVM ($n = 3$)
Accuracy$^{\triangle}$	0.72	0.69
TNR$^{\triangle}$	0.73	0.74
TPR$^{\triangle}$	0.71	0.48
F_1 score$^{\triangle}$	0.78	0.57
ROC AUC$^{\triangle}$	0.64	0.49
Cohen's kappa$^{\triangle}$	0.28	0.26
Jaccard index$^{\triangle}$	0.63	0.62
Matthews correlation coefficient$^{\triangle}$	0.38	0.28
Recall$^{\triangle}$	0.78	0.69
Hamming loss$^{\triangledown}$	0.37	0.38
Zero-one loss$^{\triangledown}$	0.25	0.38
Log loss$^{\triangledown}$	12.93	12.96

$^{\triangle}$ The higher the better.
$^{\triangledown}$ The lower the better.

Table 1 compares performance of two different classification techniques (provided lists, containing parameters of spikes, employed in training and testing are defined by Strategy A, see Section 4.1):

1. ANN;
2. SVM with cubic kernel ($n = 3$) and penalty parameter $C = 100$.

To eliminate statistical fluctuations in the dependencies of computed metrics vs N_{spikes}, we were employing fitting, analoguous to that presented by equation (16). The numerical values, presented in Table 1, were obtained by taking $N_{spikes} = 100$ in the fitted trend (*e. g.* ANN accuracy is exactly the value (17)).

Despite competitive (between ANN and SVM) accuracy values (0.72 and 0.69, respectively) and the fact that TNR (classification accuracy within Group I) values are also comparable (0.73 and 0.74, respectively), it is TPR (classification accuracy within Group II) that reveals *failure of SVM.* Only $\approx 48\%$ of Group II cases (from testing pool) were classified correctly by SVM – while ANN yielded $\approx 71\%$ success rate there. Note that ANN dominates SVM in other metrics (we do not provide their interpretation here, see *e. g.* [30]), too.

4.3. Statistical reasoning related to ANN and SVM classifiers

We discuss the importance of some statistical properties of precomputed (before the classification) values of upslopes (k_u) and downslopes (k_d) on the accuracy of ANN and SVM based classifiers.

In order to be efficient, any classification techniques should be based on some features, possessing values that are similar within classification group, but significantly different between the classification groups.

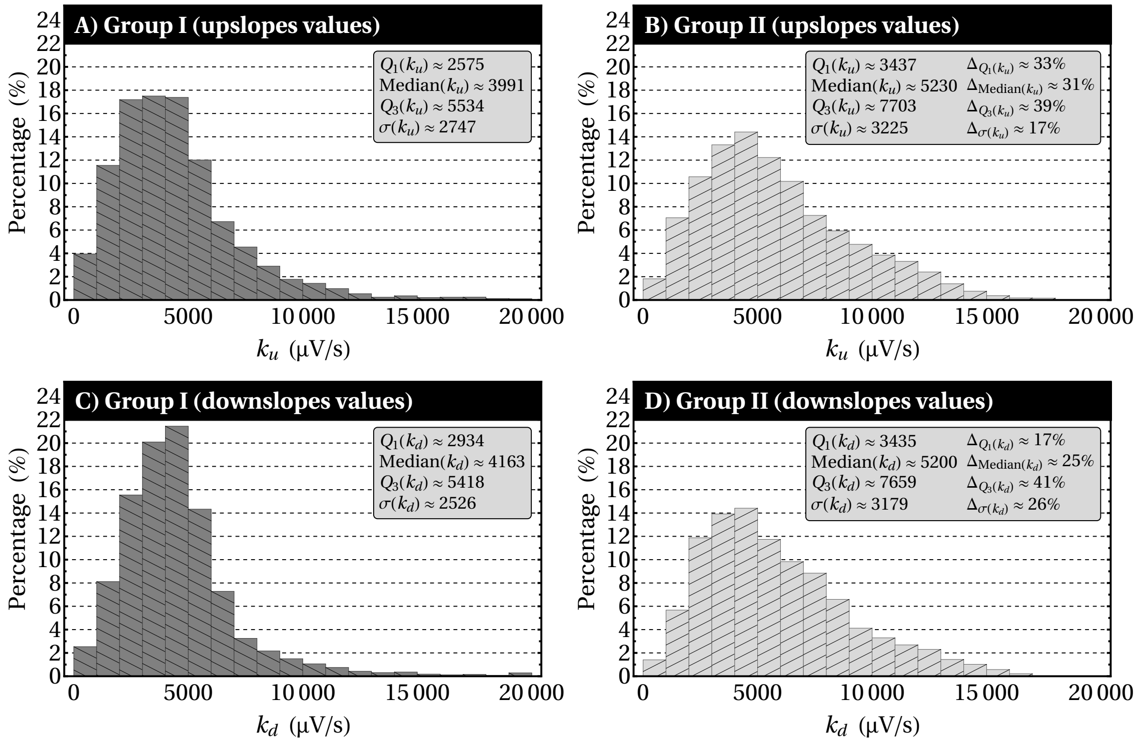


Figure 7: Distribution densities (the same 20 bins, spread uniformly in the interval $[0, 20000]$ μV/s have been employed in all presented cases) of derivative EEG spikes parameters k_u (upslopes values) and k_d (downslopes values) for patients suffering with different forms of epilepsy (Group I and Group II). Percentages Δ_x represent increase (going from Group I to Group II) of statistical quantity x. A) k_u distribution for Group I data; B) k_u distribution for Group II data; C) k_d distribution for Group I data; D) k_d distribution for Group II data.

Quartiles Q_1, Q_2 (median), Q_3 and standard deviation σ (of derivative spike parameters k_u and k_d) show 17 – 41% differences (increments) between Group I and Group II (see Fig. 7). This (also see Fig. 8) agrees with initial hypothesis that spikes of patients in Group II are more varied in their shape than ones in Group I. Even without employing machine learning type algorithms, one could try classification between Group I and Group II – based on estimated values of discussed statistical quantities. However, the accuracy of classification, achieved by ANN, has outmatched (as was expected) the accuracy of less sophisticated methods. Note that reliable estimation of quartiles and σ requires sufficiently long array of derivative EEG spike parameters (this observation is in accordance with results, presented in Fig. 6).

Now we are going to discuss in more detail the importance of the fact that large part of scattering area of points (k_u, k_d) overlaps (in both Group I and Group II cases), see Fig. 8. Preliminary (however, non trivial)

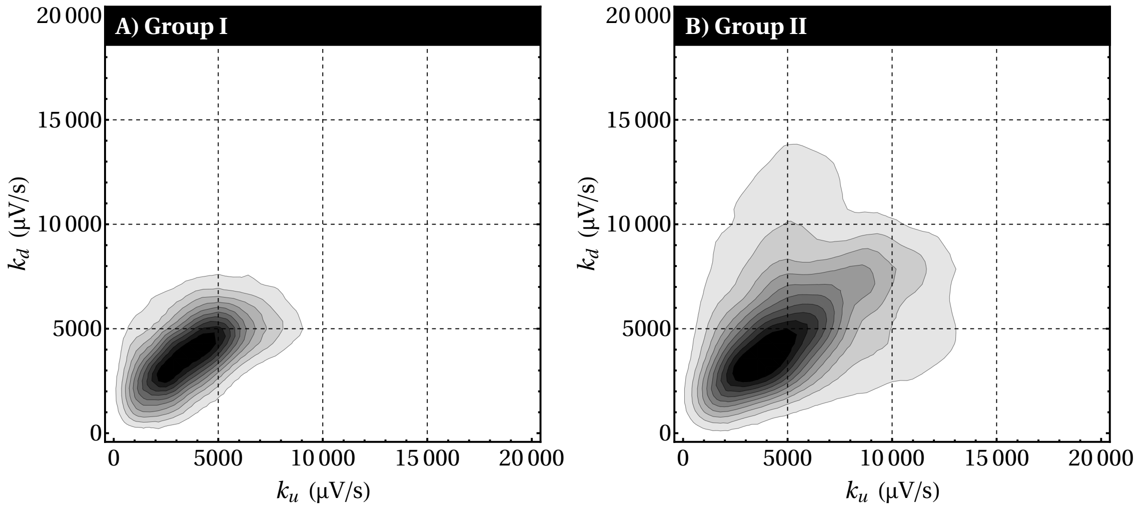


Figure 8: Contour kernel density plot (precise to 10 contours resolution), visually denoting scattering of k_u (upslopes values) and k_d (downslopes values): A) for Group I data; B) for Group II data. Note that large part of scattering area in A) and B) cases overlaps, however, the case B) demonstrates wider scattering (also see Fig. 7).

analysis in high dimensional spaces (defined by array of points $(k_u,\ k_d)$) shows still overlapping clusters. One could argue that this property would present some serious challenge for SVM classification – because all SVM based classifiers depend on the separation (by some hyperplane) of processed data in high dimensional space. This would provide explanation of unsatisfactory SVM classification results, obtained in our study (see Sections 3.3.2 and 4.2).

Due to the results, presented and discussed in Sections 4.1 and 4.2, ANN, as automatic classification technique for distinguishing between patients diagnosed with different epilepsy types, has been selected.

5. Conclusions

In this study we have confirmed our initial hypothesis, that *even visually similar spikes are not really identical*, because of the different etiology and statistical differences between Group I and Group II spike parameters can be detected by machine learning type methods.

The most important result of this study is that the proposed novel algorithm (see Section 3) can be used in order to classify between benign focal childhood epilepsy and structural focal epilepsy patients EEGs. Using this methodology is possible to achieve up to 75% accuracy rate of classification (see Sections 4.1 and 4.2).

To our knowledge, no results, related to the automatic classification between EEGs of patients with different forms of epilepsy (in our case Group I and Group II), have been published.

Also, the proposed algorithm (defined by three consecutive steps: 1) EEG spike detection, 2) EEG spike measurement, 3) ANN based classifier), presents a novel methodology: classification is based on the precomputed (in the second step) values of derivative geometric parameters (upslopes and downslopes) of EEG spikes.

The classification accuracy depends on:

1. Reliability of spike detection, discussed in our previous work [12];
2. Reliability of upslope and downslope values (k_u and k_d) estimation, discussed in our previous work [23];
3. Selection and size of training and testing data pools, discussed in Section 2.2;
4. Choice of strategy for automatic classification, discussed in Section 4.1.

The results presented in Fig. 6 suggest, that, performing ANN based classification, there is no reason to deal with spike lists of size larger than 100 spikes (as statistical saturation within range of error ±3% of accuracy rate is observed for this value of N_{spikes}). Despite this fact, the achieved classification accuracy rate of $72-75\%$ could be even more increased if more training and testing EEG data would be available.

If processed EEG has low number of spikes, the algorithm still can be used, however with decreased degree of accuracy. For example, classification accuracy of EEG, containing 20 spikes, is about 67% (see Fig. 6).

It should be noted that Group I and Group II define unbalanced data set, we have more EEGs (and therefore detected spikes) in Group I than in Group II (see Section 2.1). Therefore, to show real performance of ANN classifier, other metrics (beside accuracy) have been reported. Only after analysis (including ANN and SVM comparison) presented in Section 4.2, we have concluded that ANN (as the 3-rd step of proposed algorithm) can be approved for distinguishing between patients diagnosed with different epilepsy types.

Also, it must be pointed out that not all classification tasks require sophisticated classification algorithms. As we already mentioned in Section 3.3, *linear regression* and *linear discriminant analysis* methods were also given a trial. Poor accuracies achieved by linear regression ($\approx 53\%$) and by linear discriminant analysis ($\approx 59\%$) led us to conclusion that EEGs processed in this study could not be classified trivially.

One could hypothesize that analogous methodology might be used for classification of EEGs of other epilepsy types, which are characterised by spikes. By saying this we hope that this methodology possibly has much more applications yet to be discovered.

Conflicts of interest

The authors declared that they have no conflicts of interest to this work.

[1] M. Li, W. Chen, T. Zhang, Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble, Biomed. Signal Process. Control 31 (2017) 357 – 365. doi:10.1016/j.bspc.2016.09.008.

[2] V. Joshi, R. B. Pachori, A. Vijesh, Classification of ictal and seizure-free EEG signals using fractional linear prediction, Biomed. Signal Process. Control 9 (2014) 1 – 5. doi:10.1016/j.bspc.2013.08.006.

[3] L. M. Patnaik, O. K. Manyam, Epileptic EEG detection using neural networks and post-classification, Comput. Methods Programs Biomed. 91 (2) (2008) 100–109. doi:10.1016/j.cmpb.2008.02.005.

[4] E. Byvatov, U. Fechner, J. Sadowski, G. Schneider, Comparison of support vector machine and artificial neural network systems for drug/nondrug classification, Journal of Chemical Information and Computer Sciences 43 (6) (2003) 1882–1889, pMID: 14632437. doi:10.1021/ci0341161.

[5] H. Lüders, R. P. Lesser, D. S. Dinner, H. H. Morris, Benign Focal Epilepsy of Childhood, Springer London, London, 1987, Ch. 13, pp. 303–346. doi:10.1007/978-1-4471-1401-7_13.

[6] L. L. Medeiros, C. Yasuda, K. M. R. Schmutzler, M. M. Guerreiro, Rolandic discharges: Clinico-neurophysiological correlation, Clinical Neurophysiology 121 (10) (2010) 1740–1743. doi:10.1016/j.clinph.2010.03.050.

[7] T. Morikawa, Rolandic discharges in benign childhood epilepsy with centrotemporal spikes, and in other forms of partial epilepsies, Epileptic Disorders 2 (4) (2000) 23–28.

[8] J. J. Halford, Computerized epileptiform transient detection in the scalp electroencephalogram: Obstacles to progress and the example of computerized ECG interpretation, Clinical Neurophysiology 120 (11) (2009) 1909 – 1915. doi:10.1016/j.clinph.2009.08.007.

[9] S. B. Wilson, R. Emerson, Spike detection: a review and comparison of algorithms, Clinical Neurophysiology 113 (12) (2002) 1873 – 1881. doi:10.1016/S1388-2457(02)00297-3.

[10] S. Nishida, M. Nakamura, A. Ikeda, H. Shibasaki, Signal separation of background EEG and spike by using morphological filter, IFAC Proceedings Volumes of 14th World Congress of IFAC 32 (2) (1999) 4301–4306.

[11] A. Juozapavičius, G. Bacevičius, D. Bugelskis, R. Samaitienė, EEG analysis – automatic spike detection, Nonlinear Analysis: Modelling and Control 16 (4) (2011) 375 – 386.
URL http://www.mii.lt/na/issues/NA_1604/NA16401.pdf

[12] A. V. Misiukas Misiūnas, T. Meškauskas, A. Juozapavičius, On the implementation and improvement of automatic EEG spike detection algorithm, Proc. of the Lithuanian Mathematical Society 56 (Ser. A) (2015) 60–65.

[13] R. Degen, H. Holthausen, T. Pieper, I. Tuxhorn, P. Wolf, Benign epileptic discharges in patients with lesional partial epilepsies, Pediatric Neurology 20 (5) (1999) 354–359. doi:10.1016/S0887-8994(99)00004-1.

[14] J. Aicardi, Atypical semiology of rolandic epilepsy in some related syndromes, Epileptic disorders : international epilepsy journal with videotape 2 Suppl 1 (2000) S5–9.
URL http://europepmc.org/abstract/MED/11231217

[15] D. K. Pal, C. Ferrie, L. Addis, T. Akiyama, G. Capovilla, R. Caraballo, A. de Saint-Martin, N. Fejerman, R. Guerrini, K. Hamandi, I. Helbig, A. A. Ioannides, K. Kobayashi, D. Lal, G. Lesca, H. Muhle, B. A. Neubauer, T. Pisano, G. Rudolf, C. Seegmuller, T. Shibata, A. Smith, P. Striano, L. J. Strug, P. Szepetowski, T. Valeta, H. Yoshinaga, M. Koutroumanidis, Idiopathic focal epilepsies: the "lost tribe", Epileptic Disorders 18 (3) (2016) 252–288. doi:10.1684/epd.2016.0839.

[16] I. E. Scheffer, S. Berkovic, G. Capovilla, M. B. Connolly, J. French, L. Guilhoto, E. Hirsch, S. Jain, G. W. Mathern, S. L. Moshé, D. R. Nordli, E. Perucca, T. Tomson, S. Wiebe, Y.-H. Zhang, S. M. Zuberi, Ilae classification of the epilepsies: Position paper of the ilae commission for classification and terminology, Epilepsia 58 (4) (2017) 512–521. doi:10.1111/epi.13709.

[17] H. Gastaut, "Benign" or "functional" (versus "organic") epilepsies in different stages of life: an analysis of the corresponding age-related variations in the predisposition to epilepsy., Electroencephalogr. Clin. Neurophysiol. Suppl. 35 (1982) 17–44.

[18] R. D. S. Riesgo, P. Jayakar, N. T. Rotta, Benign rolandic epilepsy: clinical and electroencephalographic correlates, Arquivos de Neuro-Psiquiatria 58 (2000) 852 – 861.
URL `http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-282X2000000500010&nrm=iso`

[19] W. Van der Meij, A. C. Van Huffelen, G. H. Wieneke, J. Willemse, Sequential EEG mapping may differentiate "epileptic" from "non-epileptic" rolandic spikes, Electroencephalography and Clinical Neurophysiology 82 (6) (1992) 408–414.

[20] W. van der Meij, G. H. Wieneke, A. C. van Huffelen, Dipole source analysis of rolandic spikes in benign rolandic epilepsy and other clinical syndromes, Brain Topography 5 (3) (1993) 203–213. doi:10.1007/BF01128988.

[21] P. K. H. Wong, Stability of source estimates in rolandic spikes, Brain Topography 2 (1) (1989) 31–36. doi:10.1007/BF01128841.

[22] P. K. Wong, Source modelling of the rolandic focus, Brain Topography 4 (2) (1991) 105–112.

[23] A. V. Misiukas Misiūnas, T. Meškauskas, R. Samaitienė, Derivative parameters of electroencephalograms and their measurement methods, Proc. of the Lithuanian Mathematical Society 57 (Ser. A) (2016) 47–52.

[24] File:21 electrodes of International 10-20 system for EEG.svg, [Under public domain] (2010).
URL `https://commons.wikimedia.org/wiki/File:21_electrodes_of_International_10-20_system_for_EEG.svg`

[25] J. Urigüen, B. na Zapirain, EEG artifact removal – state-of-the-art and guidelines, Journal of neural engineering 12 (2015) 031001. doi:10.1088/1741-2560/12/3/031001.

[26] M. A. Nielsen, Neural Networks and Deep Learning, Determination Press, 2015.

[27] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, 1983.

[28] R. Samaitienė, Associations between EEG data, sleep, behavioural disorders and clinical characteristics in Rolandic epilepsy (Doctoral thesis, in Lithuanian), Vilnius University (2013).

[29] IFSECN, International Federation of Societies for Electroencephalography and Clinical Neurophysiology, A glossary of terms most commonly used by clinical electroencephalographers, Electroencephalography and Clinical Neurophysiology 37 (5) (1974) 538 – 548.

[30] C. Sammut, G. I. Webb, Encyclopedia of Machine Learning and Data Mining, Springer, 2017.

[31] S. Abe, Support Vector Machines for Pattern Classification, Springer, 2005.

B Priedas

Accuracy of different machine learning type methodologies for EEG classification by diagnosis

Andrius Vytautas Misiukas Misiūnas[1⋆], Tadas Meškauskas[1], and Rūta Samaitienė[2,3]

[1] Institute of Computer Science, Faculty of Mathematics and Informatics, Vilnius University, Didlaukio 47, LT-08303 Vilnius, Lithuania

[2] Children's Hospital, Affiliate of Vilnius University Hospital Santaros Klinikos, Santariškių 7, LT-08406, Vilnius, Lithuania

[3] Clinic of Children's Diseases, Faculty of Medicine, Vilnius University, Santariškių 4, LT-08406 Vilnius, Lithuania

Abstract. Electroencephalogram (EEG) classification accuracy of different automatic algorithms (including their setup) is discussed. Two patient groups, characterized by visually similar (to neurologists) EEG rolandic spikes, are under classification. The first group consists of patients with benign focal childhood epilepsy. Patients with structural focal epilepsy define the second group. We analyzed 94 EEGs (with known diagnosis) obtained from Children's Hospital, Affiliate of Vilnius University Hospital Santaros Klinikos.
The EEGs are preprocessed by applying these steps: i) spike detection; ii) extraction of spike parameters. After preprocessing of EEGs we gather parameters of detected spikes into lists of equal length N_{spikes}.
The classification algorithms are trained employing one set of patients (containing patients from both groups) and tested on another non-overlapping set of patients (also from both groups). This prevents artificial accuracy inflation due to overfitting.
We compared eight machine learning type classifiers: 1) random forest, 2) decision tree, 3) extremely randomized tree, 4) adaptive boosting (AdaBoost), 5) artificial neural network (ANN), 6) supported vector machine (SVM), 7) linear discriminant analysis (LDA), 8) logistic regression. To estimate quality of classifiers we discuss a set of metrics. The results are following: I) as expected, for all examined algorithms, the accuracy tends to grow (when N_{spikes} increases), saturating at some asymptotic value; II) ANN has prevailed as best classifier.
Impact of: a) different training strategies and b) spike detection errors on classification accuracy is also discussed.
Novelty and originality of this study comes not only from classifying different types of epilepsy, but also from employed computational methodology (involving parameters of EEG spikes and machine learning type classifier), as well as comparing different methodologies of such type, based on their accuracy and other classifier metrics.

Keywords: Machine learning · EEG · Epilepsy · EEG spikes.

⋆ andrius.misiukas@mif.vu.lt

1 Introduction

Manual analysis of electroencephalograms (EEGs) is a very difficult and time consuming process, thus a lot of automatic algorithms dedicated to help neurologists with this analysis are proposed, *e. g.* [4, 13]. EEG classification analysis may aim at: classifying healthy *vs* ill patient EEGs (*e. g.* ictal and non-ictal EEGs [5], patients diagnosed with epilepsy and healthy [11]), patients with addiction (*e. g.* from alcohol) and addiction-free [1]. Various machine learning (ML) algorithms are employed in solving ever increasing variety of EEG analysis related problems. The choice of classification algorithm is not the only important choice to make. Algorithm parameters and training strategies, EEG preprocessing are significant too.

In this study problem of *classification by diagnosis* is tackled (different from most other works, dealing with healthy *vs* ill type tasks). Previously (in [9]) we proposed the three-stage (EEG spikes detection, evaluation of spikes characteristics, classification) algorithm, based on artificial neural network (ANN) for this problem, as well as some analysis of ANN and SVM classifiers. In this paper we analyze more ML type methods for the same classification problem.

Details and in-depth analysis of data preprocessing (including EEG spikes detection), see *e. g.* [7, 8], are out of scope of this article. Instead, here we try to answer the following questions: which classification algorithm is most accurate and most robust (to low amount of spikes in one EEG, false positive and false negative detections of EEG spikes). By discussing these questions (see Section 4 and Section 5), we challenge ANN based method against other known classifiers: random forest, decision tree, extremely randomized tree [3], adaptive boosting (AdaBoost) [2], supported vector machine (SVM), linear discriminant analysis (LDA) and logistic regression.

In numerical experiments Python 3.6.4 programming language was employed as well as `scikit-learn` library (for implementations of ML algorithms and their performance metrics, see Section 4.1).

2 Data

EEGs analyzed in this study were provided by Children's Hospital, Affiliate of Vilnius University Hospital Santaros Klinikos. EEG recordings span over a 2010 – 2017 period, only 3 – 17 year old patients (all with known exact diagnosis) were included.

94 EEGs (of 86 different patients) were processed – divided into two groups, characterised by visually similar (to neurologists) EEG spikes:

- **RE group:** benign childhood epilepsy (rolandic epilepsy) with centrotemporal spikes (62 EEGs or about 66% of all EEGs), 35 of them boys, $\approx 56.4\%$. 75% of all detected EEG spikes were in this group.
- **CP group:** structural focal epilepsy patients with cerebral palsy, dysplastic brain lesion, gliosis *etc.* (32 EEGs, or about 34% of all EEGs), 18 of them boys, $\approx 56.3\%$. 25% of all detected EEG spikes were in this group.

It should be pointed out that only 30% of all EEGs were manually cleaned by neurologist from artifacts (*e. g.* patient movement) – thus decreasing error of spikes detection (see Section 3.1). Rest of the data were processed uncleaned. Cleaned data were used for classification algorithm training while uncleaned data – for algorithm validation (except for testing robustness of the algorithm, by including uncleaned data in the training data set, see Section 4.2).

3 Preprocessing of data

In this section we briefly discuss the steps and algorithms that are performed on the data before it is been tried to classify. Our algorithm consists of three basic steps:

1. EEG spike detection by morphological filter, discussed in [7, 6, 10], also see Section 3.1;
2. Extraction of EEG parameters, introduced and discussed in [8], also see Section 3.2;
3. EEG parameter validation, for details see [8], also see Section 3.2.

3.1 EEG spike detection by morphological filter

The first step of our algorithm is EEG spike detection by morphological filter based algorithm. The morphological filter is defined using close-opening and open-closing operations which can be defined using morphological erosion and dilation. For exact definitions please see [7, 6].

The idea of morphological filter is to filter out known normal brain activity (brain rhythms, patient movements, *etc.*) leaving abnormal brain activity. Then a detection limit is calculated, all maxima (higher than the detection limit) in filtered signal are treated as potential spikes. However this generates some false positive detections, this problem is addressed in the third step of this algorithm.

3.2 EEG spike parameter extraction and validation

After EEG spikes are detected, some parameters of these spikes are extracted. These parameters include upslope, downslope, width and baseline at half maximum. Our earlier investigations show that upslope and downslope are the most significant for the classification. In this work upslope and downslope value pairs for each spike are employed for classification.

These values are validated against range of known medically possible values of these parameters. This helps us to decrease false positive detections [8] (however some false positive detections remain). All values that do not meet these criteria are excluded from further analysis.

4 Comparison of various classification methodologies by performance

A strict rule of keeping each patient in either testing or training data set was kept during this work. The main reason this was done is that each patients EEG spikes are similar to one another and mixing patients in both training and testing samples would result in artificially inflated accuracy values since it would result in same effect as training and testing the classifiers on same data.

Since all EEGs have different amount of spikes and most classification algorithms can be trained and used with fixed number of inputs, EEGs were cut into non-overlapping lists of length N_{spikes}. This allowed us to treat each list as virtual EEG thus making our data accessible to algorithms and allowing us to measure the performance metrics of algorithms more reliably. For most experiments $N_{spikes} = 100$, except for experiment to test accuracy *vs* length of spike lists (see Section 4.3). As mentioned in Section 2 we had limited amount of artifact-free data as well. All the clean data were used to train or fit algorithms and rest of the data – to test except for testing robustness of the classifiers (see Section 4.2).

In this section we are going to explore quality and robustness of different machine learning type classifiers (presented in the Introduction) in respect to various ML classifier metrics (see Section 4.1), by falsely detected spikes (see Section 4.2) and dependency accuracy spike series length (see Section 4.3).

4.1 Performance metrics of EEG classification algorithms

The main aim of this study is to find the ML classification algorithm best suited for classifying EEGs obtained from patients from RE and CP groups. In order to achieve this task, some quantifiable parameters of algorithm performance are needed. The most obvious metric for this task is accuracy, which is sum of true positives and true negatives divided over all detections. This metric is very useful in detecting poorly performing algorithms.

After measuring the accuracy, LDA and logistic regression algorithms were excluded from further analysis due to their accuracy being poor: 53% and 59% respectively. Multiple supported vector machine (SVM) classifier configurations were tested as well. SVM classifiers with linear and quadratic kernels performed consistently with worse accuracy than SVM with cubic kernel thus were removed from further analysis.

While accuracy is a great tool for finding out some poorly performing algorithms, it does not show all of them. For that reason some true positive rate (TPR) and true negative rate (TNR) analysis was done. Although SVM with both RBF and sigmoid kernels were performing with good accuracy of 75%, they were classifying all the data as RE group. The accuracy was achieved purely due to our data set being biased towards RE group (see Section 2). Due to this reason these algorithms were excluded from further analysis.

Random forest, decision tree, extremely randomized trees, AdaBoost and ANN presented comparable results for both groups and thus were analyzed further. Table 1 presents the commonly used performance metrics [12] for algorithms tested. These tests were performed to evaluate overall quality of the discussed classifiers. In order to minimize statistical error caused by oscillations (*e. g.* see Fig. 1) of these metrics we first apply linear fitting (for large $30 \leqslant N_{spikes} \leqslant 100$) and then take numerical value of the fitted trend.

Table 1. Performance metrics [12] for algorithms selected group of algorithms with $N_{spikes} = 100$. Ideal classifier column represents metric values for theoretical ideal classifier.

Score/ Algorithm	**Random forest**	**Decision tree**	**Extremely randomised tree**	**AdaBoost**	**ANN**	**SVM N=3**	**Ideal classi–fier**
Accuracy	0.78	0.76	0.80	0.81	0.75	0.69	**1.00**
TNR	0.79	0.76	0.83	0.90	0.79	0.79	**1.00**
TPR	0.74	0.77	0.71	0.52	0.74	0.48	**1.00**
F1 score	0.76	0.76	0.75	0.64	0.78	0.57	**1.00**
ROC AUC	0.53	0.49	0.56	0.69	0.64	0.49	**1.00**
Cohen kappa	0.06	-0.01	0.12	0.38	0.28	0.26	**1.00**
Hamming loss	0.48	0.52	0.45	0.32	0.37	0.38	**0.00**
Jaccard simmilarity score	0.52	0.48	0.55	0.68	0.63	0.62	**1.00**
Log loss	16.72	17.92	15.55	11.00	12.93	12.96	**0.00**
Matthews correlation coefficient	0.07	-0.01	0.15	0.42	0.38	0.28	**1.00**
Recall score	0.78	0.76	0.81	0.84	0.78	0.69	**1.00**
Zero one loss	0.48	0.52	0.45	0.32	0.25	0.38	**0.00**

AdaBoost seems to be the best algorithm by most metrics presented in Table 1, except a couple key ones: TPR and F_1 score. This is due to the fact that AdaBoost classifies RE (dominant group) correctly 90% of the time and CP group – only about 52% of the time, SVM with cubic kernel suffers from the same problem. Despite of good performance of AdaBoost across all other metrics, this algorithm is not suited for the task at hand – detecting more rare CP group cases in the pool of CP and RE group data. However AdaBoost could be explored further for potential use in ensemble (voting) type of classifier. This leads to discussion that some classifier quality metrics can be misleading in this case (see Section 5).

Table 1 shows some more interesting results. Although random forest, decision tree and extremely randomized trees show both high TPR and TNR, their ROC AUC, Cohen kappa and Matthews correlation coefficient are poor. This is probably due to the reason that these metrics are designed to take into account chance of classifying a record correctly by guessing, therefore these metrics suggest that these algorithms are getting the correct answer by guessing it. Extremely randomized tree suffers less from this problem, its Cohen kappa and Matthews correlation coefficient scores are still poor. This means that these algorithms are less suited for EEG classification than ANN and are excluded from further analysis.

This leaves us with ANN, SVM (with cubic kernel) and AdaBoost classifiers. Of these three, the ANN classifier is better considering all metrics, thus it is recommended to be used for automatic classification by diagnosis.

4.2 Accuracy dependency on falsely detected spikes in training set

In this section we test robustness of algorithm to both small (by percentage) training sample and falsely positively or falsely negatively detected spikes in the training data set. It should be noted that quality of classification depends not only on choice of classification algorithm, but also on: a) employed training strategy (percentage of training data vs testing data); b) spike detection errors (during EEG preprocessing, falsely positively or falsely negatively detected spikes).

As expected, more false detections exist when analyzing not cleaned EEGs (see Section 2). Our results show that all testing classifiers are sensitive to inclusion of not cleaned data to some extent, however ANN is most sensitive to inclusion of large amounts of artifacts. Similar results are observed when only clean data is used for training, but the percentage of the training sample size is reduced. This result is expected since our analysis (see Section 4.1) show that ANN is the most intelligent classifier while other classifiers are guessing the answer by using uneven distribution across groups of our data sample (see Section 2). However this is not an issue if large and clean enough data sample is available for the training.

4.3 Accuracy dependency on length of spikes lists

It is difficult to tell the diagnosis from a single or a few spikes and their parameters. In order to make more accurate attempt at classification more spike parameters are needed. Thus we divided spikes found in all EEGs into lists (see Section 4) containing up to 100 spikes ($1 \leqslant N_{spikes} \leqslant 100$) and examined the dependency between number of spikes in series and accuracy of classification for all classification algorithms employed in this study (see Fig. 1).

For most classifiers (except SVM with cubic kernel) accuracy increases as N_{Spikes} grows, for higher N_{Spikes} all classification algorithms reach saturation levels. For this reason length of 100 spike parameter series was used in all other experiments in this work. Accuracy of SVM with $N_{spikes} < 30$ is not presented due to very long calculation times.

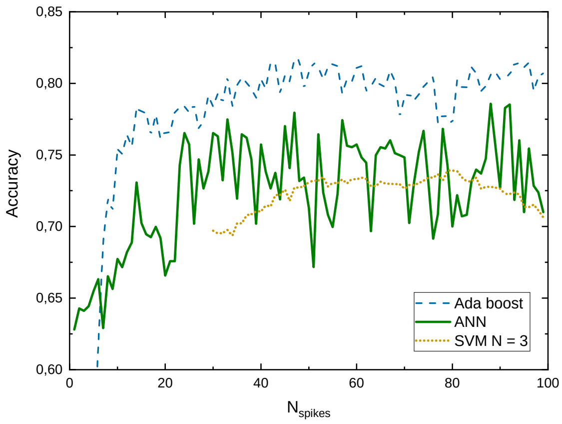


Fig. 1. Accuracy dependency on spike count.

5 Results and discussion

In this paper we compared quality of different machine learning (ML) type methodologies for EEG classification by diagnosis (see Section 4.1 and Table 1).

The analysis of ML classification algorithms present an important point about the metrics (see Table 1) themselves. If unevenly distributed across groups data set is analyzed (see Section 2), and classifier (*e. g.* AdaBoost) is biased towards larger group, most metrics still score high, except for F_1 score and either TPR or TNR. Another important point is that if data set is unevenly distributed across groups – it is much easier to guess the correct answer. Random forest, decision tree and extremely randomized tree seems to be doing that as shown by Cohen kappa and Matthews correlation coefficient. Other metrics (ROC AUC, Hamming loss, Jaccard similarity score, log loss, recall score and zero one loss) do not present any new information compared to plain accuracy. Furthermore one could consider these metrics misleading (*e. g.* with AdaBoost) for binary classification problem with unevenly distributed across groups data set.

For algorithm selection to solve classification problems similar to one presented in this work, we recommend evaluation of F_1 score, TPR, TNR and either Cohen kappa or Matthews correlation coefficient metrics. For purposes of selecting candidate algorithms for further analysis accuracy could be used since low accuracy always means poor performance, while high accuracy does not universally suggest good overall performance of algorithm.

Considering all available data we determined that ANN currently is best tested algorithm for classifying spike series of CP and RE groups.

Robustness of ANN classifier (see Section 4.2) is not an issue if correct training strategy is selected (clean data employed in training) and big enough clean data sample is available. Following that conclusion we employed all clean data available in the training data sample

These results also raise some possibilities for future work: improving spike detection accuracy, using ensemble (voting) or statistical boosting for ANN classifier in order to achieve higher classification accuracy with higher both TPR and TNR.

References

1. Byvatov, E., Fechner, U., Sadowski, J., Schneider, G.: Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. Journal of Chemical Information and Computer Sciences **43**(6), 1882–1889 (2003). https://doi.org/10.1021/ci0341161
2. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences **55**(1), 119 – 139 (1997). https://doi.org/10.1006/jcss.1997.1504
3. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. Machine Learning **63**(1), 3–42 (Apr 2006). https://doi.org/10.1007/s10994-006-6226-1
4. Halford, J.J.: Computerized epileptiform transient detection in the scalp electroencephalogram: Obstacles to progress and the example of computerized ECG interpretation. Clinical Neurophysiology **120**(11), 1909 – 1915 (2009). https://doi.org/10.1016/j.clinph.2009.08.007
5. Joshi, V., Pachori, R.B., Vijesh, A.: Classification of ictal and seizure-free EEG signals using fractional linear prediction. Biomed. Signal Process. Control **9**, 1 – 5 (2014). https://doi.org/10.1016/j.bspc.2013.08.006
6. Juozapavičius, A., Bacevičius, G., Bugelskis, D., Samaitienė, R.: EEG analysis – automatic spike detection. Nonlinear Analysis: Modelling and Control **16**(4), 375 – 386 (2011)
7. Misiukas Misiūnas, A.V., Meškauskas, T., Juozapavičius, A.: On the implementation and improvement of automatic EEG spike detection algorithm. Proc. of the Lithuanian Mathematical Society **56**(Ser. A), 60–65 (2015)
8. Misiukas Misiūnas, A.V., Meškauskas, T., Samaitienė, R.: Derivative parameters of electroencephalograms and their measurement methods. Proc. of the Lithuanian Mathematical Society **57**(Ser. A), 47–52 (2016)
9. Misiukas Misiūnas, A.V., Meškauskas, T., Samaitienė, R.: Algorithm for automatic EEG classification according to the epilepsy type: benign focal childhood epilepsy and structural focal epilepsy. Submitted to Biomedical Signal Processing and Control (2018)
10. Nishida, S., Nakamura, M., Ikeda, A., Shibasaki, H.: Signal separation of background EEG and spike by using morphological filter. IFAC Proceedings Volumes of 14th World Congress of IFAC **32**(2), 4301–4306 (1999)
11. Patnaik, L.M., Manyam, O.K.: Epileptic EEG detection using neural networks and post-classification. Comput. Methods Programs Biomed. **91**(2), 100–109 (2008). https://doi.org/10.1016/j.cmpb.2008.02.005
12. Sammut, C., Webb, G.I.: Encyclopedia of Machine Learning and Data Mining. Springer (2017)
13. Wilson, S.B., Emerson, R.: Spike detection: a review and comparison of algorithms. Clinical Neurophysiology **113**(12), 1873 – 1881 (2002). https://doi.org/10.1016/S1388-2457(02)00297-3