



Diplomová práce

Exploratorní analýza záznamů funkční magnetické rezonance metodami slepé separace

Studijní program:

N0613A140028 Informační technologie

Studijní obor:

Inteligentní systémy

Autor práce:

Bc. Martin Vrátný

Vedoucí práce:

prof. Ing. Zbyněk Koldovský, Ph.D.

Ústav informačních technologií a elektroniky

Liberec 2024



Zadání diplomové práce

Exploratorní analýza záznamů funkční magnetické rezonance metodami slepé separace

<i>Jméno a příjmení:</i>	Bc. Martin Vrátný
<i>Osobní číslo:</i>	M21000165
<i>Studijní program:</i>	N0613A140028 Informační technologie
<i>Specializace:</i>	Inteligentní systémy
<i>Zadávací katedra:</i>	Ústav informačních technologií a elektroniky
<i>Akademický rok:</i>	2022/2023

Zásady pro vypracování:

1. Funkční magnetická rezonance fMRI umožňuje monitorovat mozkovou aktivitu v čase. Pomocí metod slepé separace existuje několik způsobů, jak tato data analyzovat a rozložit je, například na funkční sítě a časové průběhy. Z nich lze dále odvodit digitální biomarkery, které dokáží indikovat různé patologické poruchy, jako je třeba schizofrenie. Úkoly pro vypracování jsou následující.
2. Student se seznámí s problematikou získávání fMRI záznamů, s technickými parametry a postupy pro předzpracování zejména metodou Analýzy hlavních komponent. Dále se seznámí s cíly a postupy, kdy je k analýze dat využívána metoda ICA a se samotnou metodou ICA a jejím rozšířením založeném na mixovacím modelu CSV. Tuto rešeršní část popíše v úvodních kapitolách DP.
3. Student provede repliku experimentu skupinové ICA na vybrané sadě dostupných záznamů a porovná výsledné prostorové mapy s mapami očekávanými, k čemuž zavede vhodné kritérium a vizuální nástroj.
4. Experiment provede pomocí rozšířené metody ICA postavené na mixovacím modelu CSV. Zvolí postupy a nastavení parametrů tak, aby dosáhl co nejlepších výsledků z hlediska očekávaných prostorových map. Například může vyzkoušet různá dělení voxelů do regionů. Výsledky se pokusí ověřit na jiné nezávislé sadě dat.
5. Navrhne doporučení, kterým by se rozšířená metoda mohla používat v praxi tak, aby výpočet trval krátkou dobu a byl zvladatelný na běžném PC.

Rozsah grafických prací: dle potřeby dokumentace
Rozsah pracovní zprávy: 40-50 stran
Forma zpracování práce: tištěná/elektronická
Jazyk práce: čeština

Seznam odborné literatury:

[1] Rangaraj M. Rangayyan. Biomedical Signal Analysis: A Case-Study. John Wiley & Sons Inc (sea) Pte Ltd., 2002. ISBN 981412611X.

Vedoucí práce: prof. Ing. Zbyněk Koldovský, Ph.D.
Ústav informačních technologií a elektroniky

Datum zadání práce: 24. října 2022
Předpokládaný termín odevzdání: 14. května 2024

prof. Ing. Zdeněk Plíva, Ph.D.
děkan

L.S.

prof. Ing. Ondřej Novák, CSc.
vedoucí ústavu

V Liberci dne 24. října 2022

Prohlášení

Prohlašuji, že svou diplomovou práci jsem vypracoval samostatně jako původní dílo s použitím uvedené literatury a na základě konzultací s vedoucím mé diplomové práce a konzultantem.

Jsem si vědom toho, že na mou diplomovou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci nezasahuje do mých autorských práv užitím mé diplomové práce pro vnitřní potřebu Technické univerzity v Liberci.

Užiji-li diplomovou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti Technickou univerzitu v Liberci; v tomto případě má Technická univerzita v Liberci právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Současně čestně prohlašuji, že text elektronické podoby práce vložený do IS/STAG se shoduje s textem tištěné podoby práce.

Beru na vědomí, že má diplomová práce bude zveřejněna Technickou univerzitou v Liberci v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů.

Jsem si vědom následků, které podle zákona o vysokých školách mohou vyplývat z porušení tohoto prohlášení.

Exploratorní analýza záznamů funkční magnetické rezonance metodami slepé separace

Abstrakt

Tato diplomová práce se zabývá exploratorní analýzou využitelnosti rozšířené metody ICA postavené na mixovacím modelu CSV. Konkrétním cílem práce bylo zhodnotit využití ICA CSV na datech funkční magnetické rezonance a porovnat výsledky s klasickou metodou ICA. Dále bylo předmětem experimentálně nalézt vhodnou parametrizaci rozšířené metody. V úvodní části diplomové práce jsou ujasněny klíčové pojmy a teoretický základ pro část praktickou, jež se skládá z pěti experimentů. Nejdříve byly nalezeny vhodné parametry pro dosažení nejlepších možných výsledků podle korelačního koeficientu. Tato parametrizace byla otestována na nezávislé datové sadě, která potvrdila výsledky ze sady původní. Následně bylo vyzkoušeno rozdělení vstupních dat do bloků dle mozkových regionů, což umožňuje rozšířená metoda. Toto rozdělení vedlo k dosažení lepších výsledků v nalezení problémové komponenty. Na závěr práce jsou navržena doporučení pro výpočet rozšířené metody na běžném PC.

Klíčová slova: PCA, ICA, ICA CSV, fMRI, BSS, BSE, IC, ICN, FCN

Exploratory analysis of fMRI recordings using blind source separation

Abstract

This thesis deals with exploratory analysis of the use of the extended ICA method with constant separating vector (CSV). The goal was to evaluate the use of the ICA CSV on fMRI data sets and compare the result with the ICA. Another goal was to experimentally find a suitable parametrization of the extended method. The introductory part of the thesis specifies the key words and theory needed for the experimental part, which consists of five experiments. Firstly, suitable parameters which led to the best results according to the correlation coefficient were found. These parameters were tested on an independent data set and the setting of the parameters was confirmed to be suitable. After that, slicing of the input data based on the anatomy of the brain regions was tested. This was made possible due to the use of the extended method. The slicing of the data set resulted in better results in finding a problematic component. Lastly, author summarizes recommendations for use of the extended method on a PC.

Keywords: PCA, ICA, ICA CSV, fMRI, BSS, BSE, IC, ICN, FCN

Poděkování

Rád bych poděkoval panu profesoru Zbyňku Koldovskému za vedení práce, ochotu při konzultacích a poskytnutí datových souborů. Dále panu inženýru Jaroslavu Čmejlovi za pomoc při práci s Metacentrem. V neposlední řadě bych rád poděkoval rodině a paní inženýrce Vendule Včelkové za podporu při tvorbě této práce.

Obsah

Seznam zkratek	10
1 Úvod	12
2 Vysvětlení pojmů a jejich souvislostí	15
3 Techniky zobrazování mozku a jeho aktivit	18
4 Využití ICA pro analýzu fMRI dat	20
5 Slepá extrakce a slepá separace signálu	22
6 Metody pro řešení BSS a BSE	23
6.1 Analýza hlavních komponent	23
6.2 Analýza nezávislých komponent	25
6.2.1 Lineární model	26
6.2.2 Entropie	27
6.2.3 ICA pomocí minimalizace vzájemné informace	28
6.2.4 Způsoby použití ICA na datových sadách	29
6.2.5 Skupinová ICA	30
6.2.6 Analýza nezávislých vektorů	31
6.3 Analýza nezávislých komponent s konstantním separujícím vektorem	32
6.3.1 Konstantní separující vektor (CSV)	34
6.3.2 One-unit algoritmus	36
6.3.3 Deflation přístup	37
6.3.4 Symmetric přístup	37
6.3.5 Block-Deflation přístup	38
7 Skupinová analýza fMRI dat s využitím metod ICA a ICA CSV	40
7.1 Metodika	40
7.1.1 Data	41
7.1.2 Metody analýzy	41
7.2 Předzpracování dat pomocí PCA	42
7.3 Hledání optimální hodnoty pro redukci dimenze pomocí PCA	43
7.3.1 Průběh experimentu	43
7.4 Závislost nalezených IC na inicializační matici	44
7.4.1 Inicializační matice	44

7.5	Ověření spolehlivosti nalezení výsledků z druhého experimentu	48
7.5.1	PC a Metacentrum	48
7.6	Ověření experimentu na nezávislé sadě dat	49
7.7	Dělení vstupních dat do bloků dle mozkových regionů	53
7.8	Výpočet rozšířené metody na běžném PC	54
8	Závěr	56
	Použitá literatura	58

Seznam zkratek

BSE	slepá extrakce zdroje, Blind Source Extraction
BSS	slepá separace, Blind Source Separation
CSV	konstantní separující vektor, Constant Separating Vector
DMN	Default Mode Network
EEG	elektroencefalografie, Electroencefalography
EFICA	Efficient Fast Independent Component Analysis
EKG	elektrokardiografie, Elektrokardiography
fBIRN	Function Biomedical Informatics Research Network
FCN	funkční síť konektivity, Functional Connectivity Network
fMRI	funkční Magnetická rezonance, Funtional Magnetic Resonance Imaging
GIG-ICA	Group Information Guided ICA
GSP	Genomics Superstruct Project
HCP	Human Connectome Project
IC	nezávislá komponenta, Independent Component
ICA	analýza nezávislých komponent, Independent Component Analysis
ICE	nezávislá extrakce komponent, Independent Component Extraction
ICN	intrinická síť konektivity, Intrinsic Connectivity Network
IRPG	Individual ICA Based Artifacts Removal Plus Group ICA
IVA	analýza nezávislých vektorů, Independent Vector Analysis
IVE	nezávislá extrakce vektorů, Independent Vector Extraction
LMS	metoda nejmenších čtverců, Least Mean Squares
MEG	magnetoencefalografie, Magnetoencephalography
MRI	magnetická rezonance, Magnetic Resonance
NITRC	Neuroimaging Informatics Tools and Resources Clearinghouse
OGC	podmínka ortogonality, Ortogonality Constraint
PCA	analýza hlavních komponent, Principal Component Analysis
PET	pozitronová emisní tomografi, Positron Emission Tomography
RBM	Restricted Boltzman Machines
sICA	prostorová ICA, spatial ICA
SM	prostorová mapa, Spatial Map
SNR	poměr signálu k šumu, Signal to Noise Ratio
SOI	zdroj zájmu, Source of Interest
TC	časová řada, Time Course

Seznam obrázků

2.1	Lineární model rozložení fMRI dat.	17
2.2	Rozdělení intrinsických sítí do čtyřech oblastí podle anatomie a funkčních závislostí. Obrázek převzat z publikace [14].	17
4.1	fMRI data jako mix nezávislých komponent. Vlastní zpracování obrázku 3 z publikace [29].	21
6.1	Druhy ICA vlastní zpracování na základě [7].	30
6.2	Druhy skupinové ICA vlastní zpracování na základě [7].	31
6.3	Ilustrace IVA modelu. Vlastní zpracování na základě [27].	32
7.1	Průměrný korelační koeficient pro jednotlivé dimenze.	44
7.2	Vzorové prostorové mapy a nejlepší nalezené prostorové mapy podle korelačního koeficientu.	47
7.3	Výsledné korelační koeficienty po spuštění algoritmu FastDIVA vzhledem ke vzorovým komponentám.	48
7.4	Nejlepší prostorové mapy po 1 000 nezávislých spuštěních podle korelačního koeficientu	50
7.5	Maximální hodnoty korelačního koeficientu dosažené po 1 000 nezávislých spuštěních.	51
7.6	Vzorové prostorové mapy a nejlepší dosažené výsledky nalezených prostorových map dle korelačního koeficientu. Porovnání výsledků rozšířené metody a metody klasické.	52
7.7	Ilustrace vyplnění bloků metody FastDIVA podle regionů.	53
7.8	Nejlepší výsledky po 1 000 spuštěních s rozdělením do bloků dle regionů podle 7.7b.	55

1 Úvod

V posledních několika letech jsme byli svědky impozantního vzestupu umělé inteligence, což zahrnuje strojové učení inspirované funkcemi lidského mozku. Tato revoluční oblast výzkumu usiluje nejen o modelaci, ale i porozumění způsobu, jakým mozek kombinuje intuici, zkušenosti a učení. Navzdory pokroku ve vědě a technologii stále máme nedostačující poznatky o komplexním fungování mozku. To nás nutí vyhledávat nové přístupy a metody, které budou nápomocné v rozluštění jeho tajemství.

Jednou z klíčových vlastností mozku je jeho schopnost učit se a intuitivně chápat nové koncepty a dovednosti na základě již existujících znalostí a zkušeností. Porozumění této schopnosti by mělo široké využití, od vylepšení algoritmů strojového učení až po zlepšení výuky. Mozek exceluje v efektivním řešení různorodých úkolů a lepší pochopení této vlastnosti by mohlo vést k objasnění mnoha otázek v oblastech zdravotnictví, prevence psychických poruch a léčby neurologických onemocnění.

Nervový systém, složený z bilionů neuronů, je základem myšlenek, emocí a chování. Porozumění tomuto složitému provázání neuronů lze přirovnat k problému slepé separace, při níž je cílem rozdělit složité směsi signálů na jednotlivé nezávislé složky. Tato problematika má široké uplatnění v různých odvětvích a jednou z metod které ji řeší je Independent Component Analysis (ICA).

První nápad použít ICA pro analýzu fMRI dat měl Martin J. McKeown, doktor neurologie, v roce 1998. Vycházel z nedostatků dosavadních metod, které nedokázaly zachytit vztahy mezi jednotlivými regiony mozku. Neumožňovaly tedy studovat propojení mozkové funkcionality, takzvané „functional networks“ [29]. Jeho nápad se ukázal jako správný a rozvinul celé odvětví; aplikaci ICA pro analýzu fMRI dat. V roce 2001 se v literatuře poprvé objevuje skupinová ICA a zároveň je představen rozdíl mezi „Task Networks“ a „Resting networks“, které byly pomocí ICA identifikované [33]. První studie psychologických nemocí pomocí funkcionální konektivity mozku vychází v roce 2002. V tomto roce se zároveň poprvé identifikuje „Default mode network“, což je aktivita mozkových regionů a jejich propojení v klidovém stavu bez vnějších stimulů [16]. Rok 2004 přináší publikaci knihy s názvem „Independent component analysis“ od Aapo Hyvärinena. Středem této knihy je právě metoda ICA a její varianty použité v různých oblastech zájmu, a to od telekomunikace přes finance až po analýzu EEG a dalších neurologických dat [20]. V dalších letech přibývá publikací studujících psychologické nemoci, zejména Alzheimerovu nemoc a schizofrenii. Během roku 2005 je identifikováno několik „resting state networks“ [12]. Množství publikací tohoto zaměření se zvyšuje, jsou zkoumány další nemoci a metody se postupně zpřesňují a vylepšují. Například je možno uvést [28], [24],

[10], [5]. Pomocí ICA není možné přímo vytvářet závěry ohledně skupiny subjektů, a proto byla vytvořena skupinová ICA. V roce 2009 vyšel v žurnálu NeuroImage článek věnující se rozmanitosti předpřípravy dat a následnému použití pěti hlavních kategorií přístupů skupinové ICA, jimiž jsou „Combine Single subject“, „Temporal concatenation“, „Spatial concatenation“, „Pre-averaging“ a „Tensor“, a popisuje rozdíly mezi nimi [8]. Autoři článku také zmiňují software, který je možné využít pro analýzu fMRI dat, jako je GIFT [9], Brain Voyager nebo MELODIC, který je součástí většího balíku pro zpracování dat mozku (FSL) [21].

Kromě softwarových nástrojů pro práci s daty magnetické rezonance se také začínají objevovat volně přístupné datové soubory. Mezi takové zdroje patří například NITRC, blíže představeno v publikaci [23], dále například HCP, k nalezení na <https://www.humanconnectome.org/>, nebo také GSP, k nalezení na <https://www.neuroinfo.org/gsp>. Tyto databáze se stále rozšiřují a dat přibývá. Volně přístupné datové soubory pomáhají posunout vývoj analytických nástrojů vpřed, a to nejen zpřístupněním datových sad, ale také možnostmi porovnání mnoha přístupů k jejich analýze.

V roce 2012 vyšel článek shrnující aktuální poznatky ohledně využití ICA právě pro studium „Default mode network“ a „Intrinsic networks“ [7]. Zde autoři zmiňují stálý růst počtu publikací a jejich historii. Přidávají možnosti využití variací ICA jako například podmíněnou ICA, která dosahuje lepších výsledků než kompletně slepá ICA. Toto omezení je možno využít při analýze fMRI dat konkrétních regionů zájmu. Závěrem přidávají porovnání adaptivních přístupů, kde porovnávají výsledky ICA algoritmů na datech dvaceti subjektů. Žurnál Frontiers in neuroscience vydal publikaci v roce 2013, která potvrzuje sílu aplikace prostorové ICA na fMRI datech. Soustředí se na překryv funkčních sítí, které sICA nachází [37]. V tomto žurnálu se o rok později objeví i validační studie na téma využití hlubokého učení pro klasifikaci fMRI snímků. Autoři této práce se zaměřují na využití takzvaných „Restricted Boltzman machines“ (RBM), což je generativní neuronová síť s jednou skrytou vrstvou. Výsledky RBM jsou porovnatelné s ICA na jimi vybraných datech [30]. Rok 2014 přináší prozkoumání rozmanitosti algoritmů a přístupů ICA a Independent Vector Analysis (IVA) [1], [27].

V následujících letech se prozkoumávají košaté možnosti využití mnoha metodik a algoritmů. Jedním z článků je například [13], který se soustředí na odstranění artefaktů z fMRI dat. Zmiňuje tři přístupy k odstranění artefaktů na simulovaných datech; Individual ICA Based Artifacts Removal Plus Group ICA (IRPG), Group Information Guided ICA (GIG-ICA) a skupinové ICA. Provádí tři experimenty na simulovaných datech pomocí nástroje SimTB, který vytváří simulovaná fMRI data. Na závěr autoři konstatují funkčnost na simulovaných datech, nicméně si jsou vědomi komplexnosti reálných dat a s tím spojených nezodpovězených otázek. Stálý rozvoj metodologií, algoritmů a také rostoucí dostupná datová sada vedly [14] k navržení přístupu, jak využít velké data sety společně pro generaci vzorových intrinsických sítí. Věnují se otázce, jak využít rozdílné data sety ke studiu mozkových onemocnění. Spojují data ze čtyřech studií a nacházejí společné biomarkery napříč těmito daty. Přináší postup, jak z různých datových sad generovat vzorové ICN, které potom lze s vysokou spolehlivostí použít jako vzorové v dalších studiích. Kautský a spol. [22]

vydává v roce 2020 rozšíření ICA, které umožňuje separovat pohybující se zdrojové signály.

Dnes se bádání v oblastech snímání mozku a jeho analýzy těší stále velké pozornosti. Pokroky přesnosti algoritmů a nové přístupy k problematice posouvají poznatky stále vpřed. Dnes, po 25 letech studie, lze považovat tento obor za vyspělý. Nicméně je stále mnoho nezodpovězených otázek. Úspěchy získané pomocí ICA a podobných metod vedou vědce k jejich využití v hlubokém učení [32]. V posledních letech jsou v něm výsledky ICA používány za účelem vytvoření automatických nástrojů pro brzkou detekci neurologických onemocnění [31], [15].

Vzhledem k rostoucímu významu neurovědy a rozvoji technologií je pochopení mozku stále důležitější. Tato práce usiluje o přispění k tomuto cíli poskytnutím pevného základu (přesné identifikace IC) pro další hlubší analýzy.

fMRI umožňuje zachycení detailního záznamu aktivity mozku, a přinesla tak revoluční pokrok v jeho sledování. Data získaná pomocí fMRI jsou však mohutná a jejich zpracování vyžaduje pokročilé techniky.

Tato práce se zabývá zpracováním fMRI dat za pomoci metod PCA, ICA a rozšířením klasické ICA s mixujícím modelem konstantního separujícího vektoru (ICA CSV). Na začátku práce jsou vysvětleny klíčové pojmy, krátce popsány techniky k zobrazení mozku a jeho aktivity. Následuje úvaha o tom, proč používáme právě ICA pro studium fMRI dat. Dále je představen problém slepé separace a slepé extrakce, který úzce souvisí s ICA. Kapitola 6 obsahuje teoretické základy a související koncepty k metodám použitých v praktické části. Praktickou část práce pokrývá kapitola 7. Zde jsou jednotlivé experimenty rozděleny do samostatných podkapitol. První experiment 7.3 se zabývá optimální redukcí dat pomocí PCA. Druhý experiment 7.4 zkoumá závislost nalezených výsledků na inicializační matici vybraného algoritmu. Třetí experiment 7.5 slouží k ověření výsledků a sleduje spolehlivost nalezení. Čtvrtý experiment 7.6 využívá poznatků dosažených v předchozích experimentech a aplikuje znalosti na nezávislou datovou sadu. Pátý experiment 7.7 se zaměřuje na možný způsob vložení vstupních dat pro metodu ICA CSV.

2 Vysvětlení pojmů a jejich souvislostí

Než budou popisovány metody a postupy, je zpracována terminologie z oblasti analýzy dat z magnetické rezonance pomocí ICA, aby se předešlo záměně termínů, zejména při jejich překladu.

Nezávislá komponenta, Independent Component (IC)

Nezávislá komponenta v tomto případě představuje separovaný signál, mozkovou aktivitu, která je maximálně nezávislá od dalších signálů.

Časová řada, Time Course (TC)

Časová řada je sloupec mixující matice **A** nebo demixující matice **W**. Obsahuje váhy pro jednotlivé IC. Říká jaké IC jsou v daný čas aktivní a s jakou vahou.

Prostorová mapa, Spatial Map (SM)

Prostorová mapa je 3D obrázek mozku daného snímku.

Funkční konektivita je definována jako časová závislost mezi prostorově vzdálenými neurofyziologickými událostmi. V kontextu funkčního neuroimagingu se navrhuje funkční konektivita k popisu vztahu mezi neuronálními aktivačními vzory anatomicky oddělených oblastí mozku, což odráží úroveň funkční komunikace mezi oblastmi. Publikace [17], konkrétně obrázek 1, ilustruje průběh klasického fMRI měření a obrázek 3 ilustruje model funkční konektivity mozku.

Funkční síť konektivity (FCN) a intrinsická síť konektivity (ICN) jsou související koncepty v oblasti neurovědy, konkrétně ve studiu propojenosti mozku. Ačkoli sdílejí podobnosti, odkazují na odlišné aspekty mozkové aktivity a propojenosti.

Funkční síť konektivity, Functional connectivity network (FCN):

- FCN odkazuje na časovou korelaci neuronální aktivity mezi různými mozkovými oblastmi.
- Představuje míru, do jaké je aktivita v jedné mozkové oblasti synchronizována s aktivitou v jiné oblasti, bez ohledu na konkrétní kognitivní stav nebo úkol, který se provádí.

- FCN lze odvodit z fMRI, analýzou korelace vzorů mezi různými mozkovými oblastmi, aby byly odhaleny sítě propojených oblastí.
- FCN lze také studovat během konkrétních úkolů nebo kognitivních stavů, aby bylo možné porozumět tomu, jak různé mozkové oblasti spolupracují a komunikují.

Intrinsická síť konektivity, Intrinsic connectivity network (ICN):

- ICN jsou sítě mozkových oblastí, které vykazují synchronizovanou a korelovanou aktivitu v nepřítomnosti konkrétních úkolů nebo vnějších podnětů.
- Tyto sítě představují organizaci mozku do funkčně souvisejících oblastí, které mají tendenci aktivovat společně a odrážet podkladové funkční systémy nebo kognitivní procesy.
- ICN jsou často identifikovány pomocí funkční magnetické rezonance v klidovém stavu (fMRI), při němž jsou účastníci v klidu a nevykonávají žádné konkrétní úkoly.
- Mezi známou ICN patří síť výchozího módu (Default Mode Network, DMN).

Intrinsické sítě konektivity jsou konkrétním typem funkční sítě konektivity. ICN odkazují na sítě mozkových oblastí, které pravidelně ukazují korelované vzory aktivity během klidu, což odráží vnitřní funkční organizaci mozku. FCN na druhé straně zahrnuje širší koncept studia korelací v neuronální aktivitě mezi mozkovými oblastmi, ať už během klidového stavu nebo během provádění úkolů.

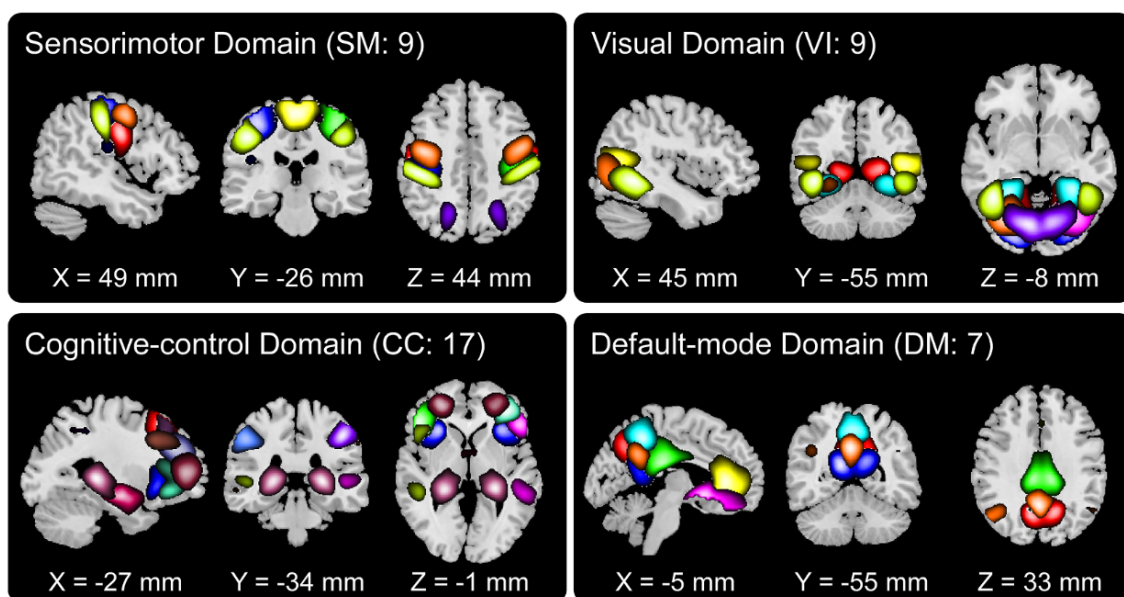
Neurální propojení mezi mozkovými regiony a jejich funkční závislosti existují i když nejsou zrovna aktivně využívány. Existují anatomická propojení, která představují uspořádání mozku a také funkční propojení, která jsou měřena pomocí úrovně okysličení krve mozku, což je považováno za aktivaci určité části. Funkční propojení odpovídá korelaci mezi aktivními částmi ve stejný čas. Čím vyšší korelace, tím silnější propojení je uvažováno. Funkční propojení mohou mít i části mozku, které nemají propojení anatomické. Pomocí fMRI je snímána úroveň okysličení mozku za určitý čas [36]. Data naměřená pomocí fMRI lze modelovat pomocí následujícího obrázku 2.1, kde \mathbf{X} představuje fMRI data, tedy směs naměřených signálů; \mathbf{A} je mixující matice; \mathbf{s} je matice původních signálů. Vybráním jednoho sloupce matice \mathbf{A} a řádku z matice \mathbf{s} vzniká nezávislá komponenta. fMRI tedy poskytuje data, která je možné separovat na nezávislé komponenty představující aktivní části mozku a jejich časové řady, které udávají kdy a jak moc je komponenta aktivní. Je předpokládáno, že jednotlivé oblasti mozku jsou nezávislé od jiných. Jednoduše řečeno, zdravý člověk zvládne pohybovat rukou a zároveň se na ruku dívat. Jednotlivé sloupce matice \mathbf{A} obsahují časové řady, tedy čas a úroveň aktivace nezávislých komponent. Funkční sítě se mohou skládat z několika nezávislých komponent najednou. Intrinsické sítě jsou využívány při studiu neurologických onemocnění, jako je například Parkinsonova choroba nebo Alzheimerova choroba. Pacienti trpící těmito nemocemi vykazují rozdílnou funkční aktivitu od zdravých pacientů. Na obrázku 2.2 jsou zobrazeny

$$\boxed{X} = \boxed{A} \cdot \boxed{s}$$

Komponenta

Obrázek 2.1: Lineární model rozložení fMRI dat.

čtyři ICN. Každá ICN se skládá z několika IC, které jsou barevně odlišeny. V závorkách je uveden konkrétní počet IC pro každou ICN. Tento obrázek byl převzat z publikace [14]. Nalezení IC záleží na mnoha faktorech ICA. Předzpracování dat a pa-



Obrázek 2.2: Rozdělení intrinsických sítí do čtyřech oblastí podle anatomie a funkčních závislostí. Obrázek převzat z publikace [14].

rametry ICA jsou jedny z hlavních. Stephen Smith a spol. ukázali rozdíly nalezení IC v závislosti na nastavení parametrů ICA [35].

3 Techniky zobrazování mozku a jeho aktivit

Zobrazovací techniky mozku revolučně změnilly poznání mozku a jeho složitých funkcí. Jednou z nejpoužívanějších technik je magnetická rezonance (MRI), která používá silné magnetické pole a rádiové vlny k vytváření detailních obrazů anatomie mozku. MRI může poskytnout obrazy mozku s vysokým prostorovým rozlišením, což umožňuje výzkumníkům identifikovat strukturální abnormality nebo poškození. Technika je neinvazivní a nevystavuje pacienty ionizujícímu záření, což ji činí bezpečnou pro rutinní klinické použití.

Funkční magnetická rezonance je specifický typ MRI, který měří změny v průtoku krve a hladin kyslíku v odpovědi na neuronální aktivitu. fMRI se stala nezbytným nástrojem v kognitivní neurovědě a vedla k významným pokrokům v poznání funkce mozku. fMRI má však relativně nízké temporální rozlišení a je náchylná k artefaktům, jako je pohyb hlavy nebo fyziologický šum.

Další populární technikou je pozitronová emisní tomografie (PET), která používá radioaktivní tracer k měření metabolické aktivity v mozku. PET může poskytnout informace o funkci neurotransmiterových systémů a byla použita k výzkumu široké škály stavů, jako je závislost a Alzheimerova choroba. PET však vyžaduje injekci radioaktivního traceru a vystavuje pacienty ionizujícímu záření.

Elektroencefalografie (EEG) a magnetoencefalografie (MEG) jsou neinvazivní techniky, které měří elektrická nebo magnetická pole generovaná neuronální aktivitou. EEG má vysoké temporální rozlišení, což umožňuje výzkumníkům studovat rychlé změny spojené s kognitivními procesy, jako je pozornost nebo vnímání. MEG má vyšší prostorové rozlišení než EEG a je schopná detekovat hluboké struktury mozku. Obě techniky jsou relativně levné a snadno použitelné, ale mají omezené prostorové rozlišení. Pro detailnější vlastnosti a techniky pořizování biologických signálů lze nahlédnout do [34].

Každá technika má své silné stránky a omezení. Možná je i kombinace několika technik, kdy jsou využity výhody jednotlivých technik. S neustálým pokrokem technologií budou techniky zobrazování mozku i nadále zůstat klíčovými nástroji pro výzkumníky i klinickou praxi. [6]

Formát dat pořízených z fMRI

Data z funkční magnetické rezonance jsou obvykle ukládána v souborovém formátu nazývaném NIfTI (Neuroimaging Informatics Technology Initiative), který je stan-

dardním formátem pro zobrazovací medicínská data. Formát NIfTI je 3D nebo 4D obrazový formát, který obsahuje anatomická i funkční data z fMRI skenů. V běžné fMRI studii jsou po dobu několika minut snímány opakovaně série 3D obrazů mozku, tzv. „časové body“ nebo „časové rámy“. Výsledná datová sada je tedy 4D datová sada se třemi prostorovými rozměry (x, y, z) a jedním časovým rozměrem. Formát NIfTI ukládá fMRI data v komprimované formě, obvykle jako jeden soubor (.nii nebo .nii.gz), a obsahuje i informace o parametrech skeneru, rozměrech obrazu a dalších metadatech.

Příprava fMRI dat k další analýze

Pro potřeby skupinové analýzy bylo nezbytné provést předzpracování dat získaných z funkční magnetické rezonance. Předzpracování dat bylo provedeno za účelem zajistit konzistentní a spolehlivé výsledky při aplikaci metody ICA a jejího rozšíření ICA CSV. Následující kroky jsou běžně využívány před další analýzou:

Zarovnání obrazů (Realignment): V prvním kroku bývá provedeno zarovnání všech fMRI snímků jednoho subjektu tak, aby byly koregistrovány na stejný referenční snímek. To zajišťuje, že všechny snímky jsou ve stejném prostorovém uspořádání a jsou minimalizovány možné pohyby nebo posuny, které by mohly ovlivnit výslednou analýzu.

Prostorová normalizace (Spatial normalization): Následně bývá provedena prostorová normalizace, která transformuje zarovnané fMRI snímky do standardního anatomického prostorového referenčního systému. Tím je dosaženo porovnatelnosti mezi různými subjekty, což je klíčové pro skupinovou analýzu, kde je cílem porovnat aktivace mezi různými jedinci nebo skupinami.

Vyhlazení (Smoothing): Dalším krokem bývá vyhlazení dat. To je provedeno za účelem snížení šumu a zvýšení poměru signálu k šumu, což usnadňuje identifikaci relevantních aktivací během analýzy. Vyhlazení se provádí aplikací prostorového filtru na každý voxel fMRI obrazu, čímž se redukuje prostorová rozlišovací schopnost dat.

Všechny výše uvedené kroky předzpracování dat jsou zásadní pro správnou funkci modelu 5.1. Předzpracování zajišťuje, že data jsou připravena pro aplikaci ICA a umožňuje identifikovat vzájemně nezávislé prostorové mapy (aktivace) a časové křivky, což umožňuje rozlišit různé funkční složky mozku. Celkově je předzpracování dat klíčovým krokem, který umožňuje dosáhnout spolehlivých a interpretovatelných výsledků analýzy fMRI dat pomocí metod ICA. Data využitá v této práci byla již předzpracována výše popsanými kroky.

4 Využití ICA pro analýzu fMRI dat

Data snímaná pomocí magnetické rezonance mají dvě hlavní složky, časovou a prostorovou. Výstupem z fMRI je tedy soubor 3D prostorových map za časový úsek. Tyto prostorové mapy mají vysokou přesnost v řádů milimetrů.

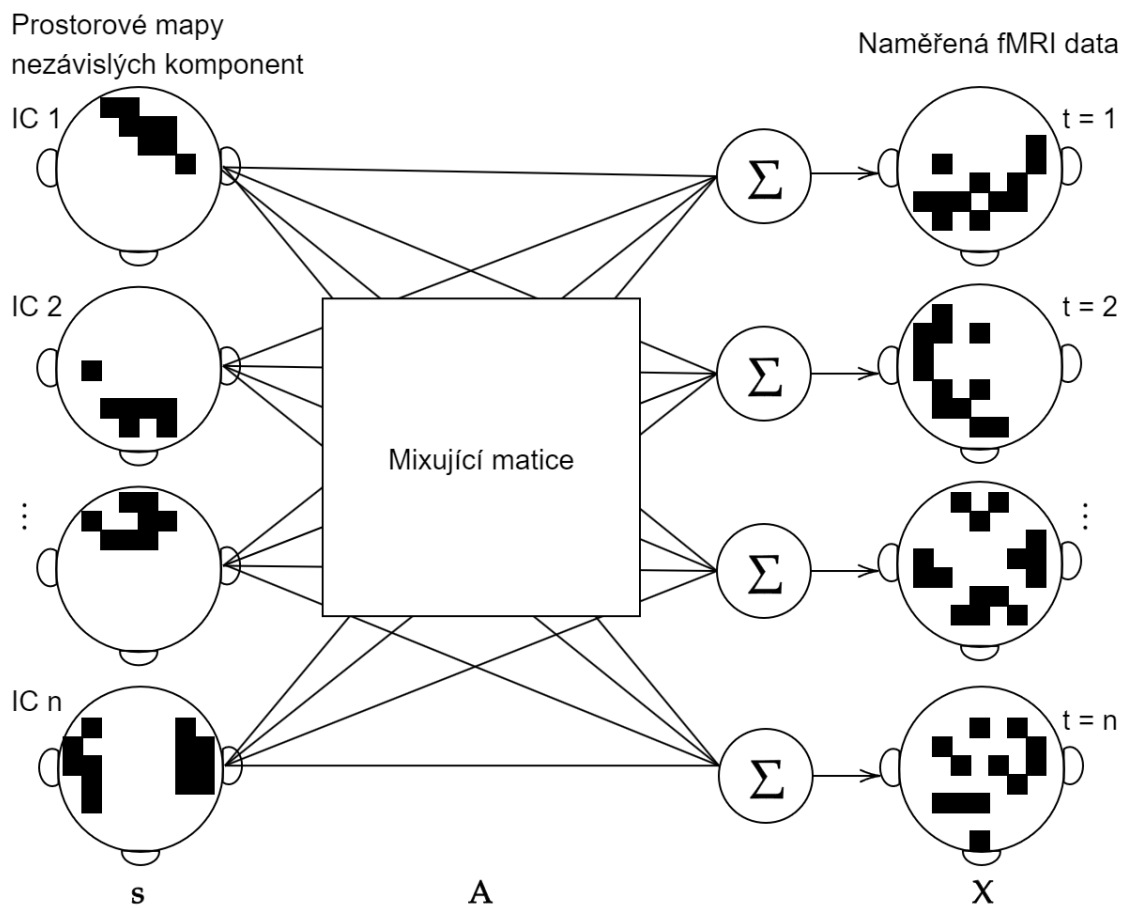
Pořízení fMRI dat však není bezproblémové. Řeší se problémy jako korekce pohybu pacienta během snímání, což vytváří nežádoucí artefakty snižující rozlišení skenu. Data z fMRI jsou vytvářena postupně pomocí řezů a může nastat situace, že mezi pořízením jednotlivých řezů nastane prodleva, což naruší časovou složku dat. Data fMRI obvykle obsahují pro jeden subjekt kolem 100 časových řezů, a proto se typicky využívá PCA pro redukci dimenze. Předzpracování pomocí PCA bude věnována samostatná sekce.

Na takto připravená data se následně aplikuje metoda ICA. ICA byla navržena pro řešení problému slepé separace zdrojů. Výstupem z fMRI jsou data s informací o mozkové aktivitě, která si lze představit jako směsici signálů měnící se v čase. Tento mix signálů se ICA snaží modelovat jako součet vážených nezávislých komponent v čase. ICA modelu, vlastnostem a využití je věnována samostatná kapitola.

Obrázek 4.1 znázorňuje rozklad fMRI dat pomocí metody ICA. \mathbf{X} představuje naměřená fMRI data, \mathbf{A} reprezentuje mixující matici a \mathbf{s} jsou nezávislé komponenty nalezené metodou ICA.

Tento rozklad si lze představit jako aktivní regiony mozku v určitém čase, kde aktivní regiony jsou právě nezávislé komponenty. Každá ICN představuje určitý neurální proces, například motorickou kontrolu, vstupy smyslů nebo kognitivní kontrolu. ICA má výhodu oproti předchozím technikám, a tou je potřeba minimální, často vůbec žádné, předchozí informace o snímaných datech. Díky tomu je ICA schopna identifikovat i neočekávané mozkové aktivity v závislosti na stimulaci. Toto z ICA dělá vhodný nástroj pro exploratorní výzkum mozkové aktivity [28].

Pomocí ICA byla také poprvé identifikována ICN s názvem Default mode network, někdy také jen Default mode, která označuje aktivitu mozku za klidového stavu [33]. Default mode network se úspěšně využívá ke studiu psychologických nemocí [7].



Obrázek 4.1: fMRI data jako mix nezávislých komponent. Vlastní zpracování obrázku 3 z publikace [29].

5 Slepá extrakce a slepá separace signálu

Slepá extrakce a separace signálu (BSE/BSS) představují v oblasti zpracování signálů a analýzy dat komplexní problém, který se zabývá extrakcí určitého původního zdrojového signálu nebo separací všech původních signálů ze smíšeného mixu, přičemž způsob jejich vzájemného mísení je neznámý. Ve většině metod se uvažuje statický lineární mixující model, který je možné zapsat pomocí

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (5.1)$$

kde \mathbf{x} je $d \times 1$ vektor označující smíchané signály; \mathbf{A} je $d \times d$ časově nezávislá mixující matice mající inverzi; \mathbf{s} je $d \times 1$ vektor reprezentující původní signály. V případě BSE jsou odhadovány pouze některé sloupce z matice \mathbf{A} a jejich odpovídající elementy z \mathbf{s} . Často je preferovaný případ, kdy je matice \mathbf{A} čtvercová a dimenze \mathbf{x} a \mathbf{s} jsou totožné, protože následně lze problém formulovat jako nalezení čtvercové demixující matice \mathbf{W} . V případě BSE jsou odhadovány pouze některé řádky matice \mathbf{W} . U ICA je hlavním předpokladem vzájemná nezávislost prvků \mathbf{s} . Matice \mathbf{W} je počítána tak, aby platilo 5.2, kde $\mathbf{W}\mathbf{x}$ jsou co nejvíce nezávislé.

$$\mathbf{s} = \mathbf{W}\mathbf{x} \quad (5.2)$$

Slepá extrakce a separace signálu jsou koncepty, které se objevují v různých oblastech, v nichž se vyskytuje mix signálů, jež je nutné rozdělit nebo zrekonstruovat bez znalosti jejich původních složek. Tento problém se vyskytuje v širokém spektru aplikací, včetně akustického zpracování, elektroencefalografie (EEG), telekomunikace a dalších multidimenzionálních sensorových systémů. Problém slepé separace signálu je založen na optimalizaci kritérií, která by měla mít určité základní vlastnosti, aby se předešlo nejednoznačné separaci. To znamená, že řešení by mělo být jednoznačné s výjimkou možných permutací či škálování zdrojových signálů. Tato práce se soustředí na aplikaci metody Analýzy nezávislých komponent s konkrétním důrazem na rozšíření této metody pomocí konstantního separujícího vektoru pro separaci signálů neuronální aktivity mozku měřené pomocí funkční magnetické rezonance.

6 Metody pro řešení BSS a BSE

Mezi nejpobulárnější metody, které řeší BSS a BSE patří analýza nezávislých komponent a její zobecnění analýza nezávislých vektorů. Často je využívána také metoda analýzy hlavních komponent (PCA), která samostatně neřeší problém BSS, ale pomáhá předpřipravít data právě pro metodu ICA. V následujících sekcích jsou popsány jednotlivé metody; PCA, ICA a IVA. Jelikož je metoda ICA hlavním zájmem práce, je jí věnováno nejvíce pozornosti. Jsou zmíněny nutné předpoklady metody, datový model a nejednoznačnosti, které jsou s ICA spojeny. Následuje definice optimalizačního kritéria a algoritmy, které kritérium odhadují. Rozšíření ICA CSV a přístupy k odhadu jsou definovány na závěr kapitoly.

6.1 Analýza hlavních komponent

Analýza hlavních komponent je jednou ze statistických technik pro extrakci vlastností (feature extraction) a redukci dimenze. Na sadě vícerozměrných měření je cílem PCA najít menší sadu proměnných s menší redundancí, která by byla co nejvěrnější reprezentací původní sady. Redundance je u PCA počítána na základě korelací mezi elementy datové sady. Využití korelace v PCA je výhodné pro výpočet, jelikož stačí pouze statistiky druhého řádu. [19]

Výchozím bodem pro PCA je náhodný vektor \mathbf{x} obsahující n prvků, kde jednotlivé prvky jsou značeny $\mathbf{x}_1, \dots, \mathbf{x}_T$. Pro PCA nejsou potřeba žádné další předpoklady, nutné je znát pouze statistiky prvního a druhého řádu, které lze odvodit z dostupných dat. Podmínkou pro využití PCA je vzájemná korelace, díky které \mathbf{x} obsahuje jistou redundanci a umožňuje tedy kompresi. Pokud jsou prvky nezávislé, není důvod používat PCA. Pro PCA transformaci je nejprve vycentrováno \mathbf{x} tak, že se odečte jeho průměr.

$$\mathbf{x} \leftarrow \mathbf{x} - E\{\mathbf{x}\} \quad (6.1)$$

Je předpokládáno, že vycentrování bylo provedeno a tudíž $E\{\mathbf{x}\} = 0$. Dále je \mathbf{x} transformován na vektor \mathbf{y} s m prvky, $m < n$, takže redundance související s korelacemi je odstraněna. Toho je dosaženo nalezením rotace ortogonálního souřadného systému tak, že prvky \mathbf{x} jsou v nových koordinátech nekorelované. Zároveň je maximalizován rozptyl projekce \mathbf{x} na nových koordinátech tak, že maximální rozptyl leží na první ose. Druhá osa odpovídá maximálnímu rozptylu osy, která je ortogonální s první osou a stejným postupem lze pokračovat dále [19].

PCA pomocí maximálního rozptylu

Nejdříve je nutné definovat, co je to komponenta:

$$y_1 = \sum_{k=1}^n w_{k1} x_k = \mathbf{w}_1^T \mathbf{x} \quad (6.2)$$

Jedná se o lineární kombinaci elementů x_1, \dots, x_n vektoru \mathbf{x} . Skalární koeficienty w_{11}, \dots, w_{n1} , které lze nazvat také jako váhy, jsou elementy n -rozměrného vektoru \mathbf{w}_1 a \mathbf{w}_1^T značí transpozici \mathbf{w}_1 . Faktor y_1 je nazýván první hlavní komponentou \mathbf{x} , pokud je rozptyl y_1 maximálně velký. Velikost rozptylu y_1 závisí na normě a orientaci vektoru \mathbf{w}_1 . Protože variance roste zároveň s velikostí normy, je v praxi zavedeno omezení velikosti normy vektoru \mathbf{w}_1 na jednotkový vektor. Je hledán váhový vektor \mathbf{w}_1 , který maximalizuje PCA kritérium 6.3

$$J_1^{PCA}(w_1) = E\{y_1^2\} = E\{(\mathbf{w}_1^T \mathbf{x})^2\} = \mathbf{w}_1^T E\{\mathbf{x}\mathbf{x}^T\} \mathbf{w}_1 = \mathbf{w}_1^T \mathbf{C}_x \mathbf{w}_1 \quad (6.3)$$

$$\text{tak, že } \|\mathbf{w}_1\| = 1. \quad (6.4)$$

Zde $E\{\cdot\}$ značí operátor střední hodnoty, která je v praxi nahrazena průměrem. Norma vektoru \mathbf{w}_1 je definována následovně:

$$\|\mathbf{w}_1\| = (\mathbf{w}_1^T \mathbf{w}_1)^{1/2} = \left[\sum_{k=1}^n w_{k1}^2 \right]^{1/2} \quad (6.5)$$

\mathbf{C}_x značí $n \times n$ kovariační matici počítanou z \mathbf{x} a lze jí zapsat jako

$$\mathbf{C}_x = E\{\mathbf{x}\mathbf{x}^T\}.$$

V praxi není teoretická hodnota kovariance známa, a proto je její odhad počítán z průměru dat. Z lineární algebry je známo, že řešení PCA je ve formě vlastních vektorů $\mathbf{e}_1, \dots, \mathbf{e}_n$ o jednotkové délce matice \mathbf{C}_x . Uspořádání vlastních vektorů koreponduje s vlastními čísly d_1, \dots, d_n a je seřazeno od největších po nejmenší. Maximální řešení 6.3 ve smyslu 6.4 je dáno

$$\mathbf{w}_1 = \mathbf{e}_1.$$

a hlavní komponentou je tedy $y_1 = \mathbf{e}_1^T \mathbf{x}$. Kritérium J_1^{PCA} popsané rovnicí 6.3 lze generalizovat na m hlavních komponent, kde m je číslo od 1 do n . Obecně pro m komponentu platí

$$y_m = \mathbf{w}_m^T \mathbf{x}$$

kde \mathbf{w}_m odpovídá jednotkovému váhovému vektoru. Rozptyl y_m je maximalizován pod podmínkou, že y_m je nekorelovan se všemi předchozími komponentami.

$$E\{y_m y_k\} = 0, k < m$$

z toho vyplývá, že hlavní komponenty y_m mají nulovou korelaci, protože

$$E\{y_m y_k\} = E\{(\mathbf{w}_m^T \mathbf{x})(\mathbf{w}_k^T \mathbf{x})\} = \mathbf{w}_m^T \mathbf{C}_x \mathbf{w}_k = 0. \quad (6.6)$$

Pro druhou hlavní komponentu platí

$$\mathbf{w}_2^T \mathbf{C} \mathbf{w}_1 = d_1 \mathbf{w}_2^T \mathbf{e}_1 = 0 \quad (6.7)$$

již víme, že $\mathbf{w}_1 = \mathbf{e}_1$. Je tedy hledána maximální variance $E\{y_2^2\} = E\{(\mathbf{w}_2^T \mathbf{x})^2\}$ v ortogonálním podprostoru prvního vlastního čísla z matice \mathbf{C}_x . Výsledkem je

$$w_2 = e_2$$

Obdobně jsou získány všechny další komponenty. Tedy platí

$$y_k = e_k^T x.$$

PCA pro předpřípravu dat ICA

Data pořízená magnetickou rezonancí mají obvykle vysokou dimenzionalitu. Jelikož fMRI snímá data ve vysokém rozlišení, není neobvyklé setkat se s datovými soubory v řádu desítek megabajtů (MB) pro jedno měření jednoho pacienta. Data s vysokou dimenzí se velmi těžko modelují. Čím víc dat je k dispozici, tím složitější je jejich estimace, a navíc často exponenciálně rostou nároky na výpočetní kapacitu a čas pro výpočet. PCA se často provádí na fMRI datech za účelem redukce dimenzionality a také redukci šumu. Pokud jsou komponenty po rozkladu seřazeny podle velikosti variance, je možné ve většině případů odebrat poslední komponenty. Poslední komponenty ve většině případů obsahují šum nebo artefakty, které jsou nežádoucí. Jejich odebráním je snížena dimenze dat a zároveň zlepšena SNR, aniž by byla ztracena užitečná informace.

6.2 Analýza nezávislých komponent

Analýza nezávislých komponent je matematická metoda, která se používá k řešení problémů slepé separace. Jejím cílem je separovat mix signálů na jednotlivé nezávislé komponenty, které odpovídají jednotlivým původním signálům, ze kterých se mix skládá. Tato separace je možná pouze pokud jsou splněny následující předpoklady:

- Původní signály jsou statisticky nezávislé.
- Původní signály jsou negausovské.
- Signál je chápán jako posloupnost nezávislých realizací jedné náhodné proměnné.

Modelů pro ICA je několik, jako příklad lze uvést lineární, nelineární, konvolutorní a další [20]. Tato práce se zabývá pouze základním lineárním modelem.

6.2.1 Lineární model

Lineární model je jedním z nejčastěji používaných. Tento model předpokládá, že pozorované signály jsou lineární kombinací skrytých komponent. Lineární model je definován jako

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (6.8)$$

kde $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ označuje n měření signálů v čase t a $\mathbf{s}(t) = [s(t)_1, s(t)_2, \dots, s(t)_m]^T$ jsou komponenty. Regulární matice \mathbf{A} je o velikosti $n \times n$ a reprezentuje váhy daných nezávislých komponent ve směsi. Matici \mathbf{A} je potřeba odhadnout a její odhad je proveden skrze odhad inverze \mathbf{A} , který je označen jako \mathbf{W} a nezávislé komponenty jsou dopočítány pomocí následujícího vztahu:

$$\hat{\mathbf{s}}(t) = \mathbf{W}\mathbf{x}(t) \quad (6.9)$$

ICA tedy spočívá v nalezení \mathbf{W} , aby jednotlivá $\hat{\mathbf{s}}$ měla co nejnížší míru závislosti.

Nejednoznačnosti ICA

Ze vzorce 6.8 je možné vidět několik nejednoznačností:

- Nejednoznačnost ve škále.
- Nejednoznačnost znaménka.
- Výsledné pořadí komponent.

Důvodem nejednoznačnosti škály je fakt, že \mathbf{A} a \mathbf{s} jsou neznámé. Jakýkoliv skalární násobek s_i lze vždy pokrátit korespondujícím sloupcem \mathbf{a}_i z matice \mathbf{A} . Pro jednodušost se proto rozptyl komponent považuje za jednotkový. Obdobný problém je se znaménkem. Vynásobením nezávislých komponent -1 neovlivní výsledný model, je však vhodné dbát na tuto nejednoznačnost. Ve většině aplikací je tento problém zanedbatelný. Poslední nejednoznačností je výsledné pořadí komponent. Důvodem je znovu neznámost proměnných \mathbf{A} a \mathbf{s} . Pokud je do modelu 6.8 přidána permutační matice \mathbf{P} a její inverze, bude model vypadat takto

$$\mathbf{x} = \mathbf{A}\mathbf{P}^{-1}\mathbf{P}\mathbf{s}, \quad (6.10)$$

kde prvky $\mathbf{P}\mathbf{s}$ jsou originální komponenty, ale v jiném pořadí. Potom $\mathbf{A}\mathbf{P}^{-1}$ je pouze nová neznámá mixující matice, kterou je potřeba najít.

Kovariance

Signály jsou nezávislé, pokud je jejich kovariance rovna nule. Kovariance dvou náhodných proměnných \mathbf{x} a \mathbf{y} je definována následovně:

$$\mathbf{C}_{\mathbf{xy}} = E[(\mathbf{x} - \eta_{\mathbf{x}})(\mathbf{y} - \eta_{\mathbf{y}})], \quad (6.11)$$

kde $\eta_{\mathbf{x}} = E[x]$ a $\eta_{\mathbf{y}} = E[y]$. Pokud jsou uvažovány signály s nulovou střední hodnotou, lze 6.11 zjednodušit pouze na

$$\mathbf{C}_{\mathbf{xy}} = E[\mathbf{x}]\mathbf{E}[\mathbf{y}] = 0, \quad (6.12)$$

Statistická nezávislost

Náhodné proměnné X_1, \dots, X_n jsou statisticky nezávislé, pokud je sdružená pravděpodobnostní hustota rovna součinu marginálních pravděpodobnostních hustot, tedy

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i) \quad (6.13)$$

Jsou-li X_i a X_j nekorelované, platí

$$\text{Cov}(X_i, X_j) = \int \int xy f_{X_1 X_2}(x, y) dx dy = \int \int xy f_{X_1}(x) f_{X_2}(y) dx dy = 0. \quad (6.14)$$

Rovnice 6.14 říká, že pokud jsou náhodné proměnné X_1 a X_2 nekorelované, znamená to, že jejich společná pravděpodobnostní distribuce se dá faktorizovat do součinu jejich marginálních distribucí. Potom kovariance mezi dvěma libovolnými instancemi X_i a X_j těchto proměnných bude nula. Pokud jsou X_i a X_j nezávislé, je jejich kovariance rovna nule. Opačná implikace ale platit nemusí.

6.2.2 Entropie

Entropie je míra neuspořádanosti nebo nejistoty systému. Jedná se o fyzikální a informatickou míru, která se vyskytuje v mnoha oborech, jako je statistika, termodynamika, informační teorie a další. V informační teorii se entropie vztahuje k množství informace obsažené v náhodné veličině. V tomto případě je předpokládáno, že jednotlivé signály jsou nezávislé a identicky rozložené (i.i.d). Pokud je systém uspořádaný a má nízkou nejistotu, obsahuje málo informací a jeho entropie je nízká. Naopak, pokud je systém neuspořádaný a má vysokou nejistotu, obsahuje více informací a jeho entropie je vysoká.

Entropie H jedné náhodné proměnné X_i s hustotou pravděpodobnosti $p_{X_i}(\cdot)$ je definována jako

$$H(X_i) = - \int p_{X_i}(x) \log p_{X_i}(x) dx \quad (6.15)$$

a entropie náhodných proměnných X_1, \dots, X_n je rovna součtu podmíněných entropií

$$H(X_1, \dots, X_n) = \sum_{i=1}^n H(X_i | X_{i-1}, \dots, X_1), \quad (6.16)$$

kde podmíněnou entropií je definována jako

$$H(X, Y) = H(X) + H(Y|X) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) \quad (6.17)$$

Entropii lze generalizovat na multidimenzionální případ, který má stejný tvar jako 6.15 s rozdílem, že X_i je náhodný vektor.

Pro množinu náhodných proměnných se stejnou střední hodnotou a rozptylem má entropie maximální hodnoty pro proměnné s gaussovým rozložením. Tento fakt

brání separaci více než jednoho gaussovského signálu. Entropii lze využít k definici vzájemné informace náhodných proměnných X_1, \dots, X_j následovně

$$I(X_1, \dots, X_n) = \sum_{i=1}^n H(X_i) - H(X_1, \dots, X_n) \quad (6.18)$$

což je přirozená míra závislosti mezi náhodnými veličinami. Vzorec 6.18 je zároveň ekvivalentní se známou Kullback-Liebler divergencí.

Kullback-Lieblerova divergence

Kullback-Lieblerova divergence je měřítkem rozdílu mezi pravděpodobnostními rozděleními. Konkrétně se jedná o měřítko „vzdálenosti“ mezi rozděleními pravděpodobnosti X_1, \dots, X_n . Kullback-Lieblerovu divergenci lze definovat následovně

$$I(X_1, \dots, X_n) = \int f_{X_1, \dots, X_n} \ln \frac{f_{X_1, \dots, X_n}}{f_{X_1} \dots f_{X_n}} \quad (6.19)$$

Vlastnosti a podrobnější informace lze nalézt v [11].

Pokud je využito vztahu 6.18, lze přenést problém nalezení \mathbf{W} ze vztahu 6.9 jako hledání takové transformace signálů, aby součet entropií transformovaných signálů byl co nejmenší. Tato myšlenka se opírá o centrální limitní větu, která tvrdí, že lineární kombinace náhodných proměnných konverguje k proměnné s gaussovským rozložením. Jak již bylo výše zmíněno, je známo, že největší entropii má právě gaussovo rozložení, tudíž přidáváním signálů do mixu jejich entropie roste.

6.2.3 ICA pomocí minimalizace vzájemné informace

Motivace pro odhad metody ICA pomocí vzájemné informace mohou vzniknout v případech, kdy data přesně nepodléhají ICA modelu, jak je popsán v kapitole 6.8. V takových situacích se k datům přistupuje bez předpokladů a ICA se používá jako nástroj pro měření závislosti mezi složkami náhodného vektoru. Tento přístup umožňuje využít vzájemnou informaci, což je důkladně odůvodněné informační měření statistické závislosti mezi proměnnými. Vzájemná informace poskytuje nástroj pro hodnocení míry závislosti mezi složkami, aniž by byly předem stanoveny předpoklady o přesném ICA modelu. Tím se ICA stává užitečnou metodou pro analýzu komplexních datových souborů, kde závislosti mohou být složité a nepřesně definované.

Vzájemná informace 6.18 bere v úvahu celou strukturu závislosti mezi proměnnými, a proto lze vzájemnou informaci použít jako kritérium pro nalezení rozkladu ICA. ICA náhodného vektoru \mathbf{x} je definována jako invertibilní transformace 6.9, kde matice \mathbf{W} je určena tak, aby vzájemná informace transformovaných komponent $\hat{\mathbf{s}}$ byla minimalizována. Minimalizace vzájemné informace může být interpretována jako nalezení maximálně nezávislých komponent, jelikož 6.19 je nulová pouze když jsou proměnné nezávislé.

Pro definici vzájemné informace lze využít vzorce 6.18. Nalezení maximálně nezávislých komponent $\hat{\mathbf{s}}$ z 6.9 je odhadováno pomocí

$$\begin{aligned}
I(\hat{s}_1, \dots, \hat{s}_n) &= \int \log \frac{f(\hat{s}_1, \dots, \hat{s}_n)}{\prod_i f(\hat{s}_i)} f(\hat{s}_1, \dots, \hat{s}_n) \\
&= - \int \log f(\hat{s}_1, \dots, \hat{s}_n) f(\hat{s}_1, \dots, \hat{s}_n) - \sum_i \log f(\hat{s}_i) f(\hat{s}_1, \dots, \hat{s}_n) \\
&= -H(\hat{s}_1, \dots, \hat{s}_n) + \sum_i (-\log f(\hat{s}_i) f(\hat{s}_i)) \\
&= \sum_i H(\hat{s}_i) - H(\hat{s}_1, \dots, \hat{s}_n)
\end{aligned} \tag{6.20}$$

Což lze zkráceně zapsat jako

$$I(\hat{s}_1, \dots, \hat{s}_n) = I(\mathbf{W}\mathbf{x}) = \sum_i H(\hat{s}_i) - H(\hat{s}_1, \dots, \hat{s}_n) \tag{6.21}$$

Pokud je \mathbf{W} ortogonální funkcí, je člen $H(\hat{s}_1, \dots, \hat{s}_n)$ konstantní, a jelikož jsou hledány maximálně nezávislé komponenty lze konstantu vytknout. Vzájemná informace je zapsána následovně:

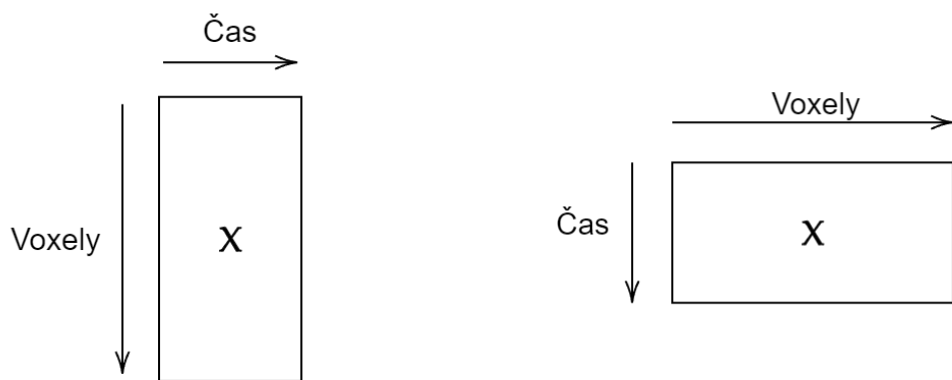
$$I(\hat{s}_1, \dots, \hat{s}_n) = \text{const.} - \sum_i H(\hat{s}_i) \tag{6.22}$$

Zde konstanta nezávisí na \mathbf{W} . Rovnice 6.22 ukazuje, že odhad ICA prostřednictvím minimalizace vzájemné informace je ekvivalentní maximalizaci sumy nelineárních odhadů nezávislých komponent, pokud jsou odhady omezeny tak, aby byly nekorelované.

6.2.4 Způsoby použití ICA na datových sadách

ICA je statistická metoda, kterou lze aplikovat na různé sady dat, avšak interpretace výsledků ICA může být náročná a závisí na mnoha faktorech, včetně poměru mezi počtem vzorků N a počtem senzorů nebo měření d . Pro fMRI data jednoho subjektu je počet vzorků N obvykle v řádu desítek až stovek tisíců, zatímco počet senzorů nebo měření d bývá v řádu desítek až stovek. Použití časové ICA na fMRI datech je možné, ale může vést k výsledkům, které jsou obtížně interpretovatelné. To je způsobeno nedostatečným množstvím dat pro statisticky významné odhady. Jedním z přístupů k vylepšení interpretovatelnosti výsledků je využití prostorové dimenze dat, která obvykle poskytuje dostatek dat pro statisticky významné odhady. V případě fMRI lze transponovat datovou matici a provádět ICA analýzu v prostorovém směru, což umožňuje identifikaci nezávislých obrazů. Tyto nezávislé obrazy mohou reprezentovat funkční síť v mozku. Obrázek 6.1 ilustruje oba přístupy na datové matici fMRI.

Časová ICA se zaměřuje na rozklad časových řad do nezávislých časových komponent. Tento případ ilustruje obrázek 6.1a Tento typ ICA se používá v oblasti časové řady a signálové analýzy, například pro oddělení různých signálů, jako jsou biologické signály, EEG nebo EKG, které jsou nahrány v průběhu času.



(a) Časová ICA.

(b) Prostorová ICA.

Obrázek 6.1: Druhy ICA vlastní zpracování na základě [7].

Na druhé straně Prostorová ICA se zaměřuje na rozklad vstupních dat do nezávislých prostorových komponent. Tento případ ilustruje obrázek 6.1b Tento typ ICA je často používán v oblasti zpracování obrazu, kde se jsou oddělovány různé signály, které jsou smíšeny do jednoho vstupního signálu. V případě zpracování obrazu každá prostorová komponenta může odpovídat jednomu objektu nebo části scény. Cílem je oddělit tyto komponenty, aby bylo možné jednotlivé objekty nebo části scény zobrazit zvlášť [7].

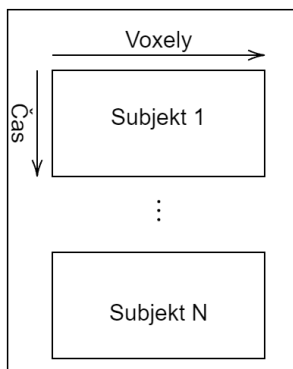
Pro fMRI data se obvykle používá prostorová ICA. Jedním z důvodů je to, že fMRI data jsou získávána jako obrazová data, tedy jako 3D nebo 4D objemy, které obsahují informace o aktivitě mozku v různých oblastech. Proto se často používají metody zpracování obrazu a prostorové analýzy pro získání informací o aktivitě mozku [7].

Oba typy ICA mají společné rysy, jako je potřeba lineární transformace dat a předpoklad nezávislosti komponent. Zatímco prostorová ICA se zaměřuje na rozklad dat do prostorových komponent, časová ICA se zaměřuje na rozklad dat do časových komponent.

6.2.5 Skupinová ICA

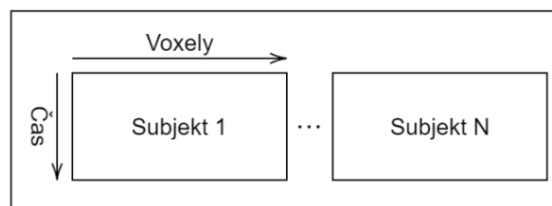
Skupinová ICA je využívána k analýze skupinových dat, tedy v případě, kdy jsou k dispozici fMRI nahrávky od více jedinců. Obecně jde o samostatná měření, která mají společné rysy. V případě fMRI se většinou jedná o stejný pokus naměřený u více jedinců. Cílem skupinové ICA je identifikovat společné nezávislé komponenty, které jsou přítomny v datových sadách od různých jedinců. Tento přístup umožňuje získat detailní informace o mozku a jeho funkci v různých skupinách pacientů.

Časová ICA pro skupinovou analýzu



(a) Časová ICA pro skupinovou analýzu.

Prostorová ICA pro skupinovou analýzu



(b) Prostorová ICA pro skupinovou analýzu.

Obrázek 6.2: Druhy skupinové ICA vlastní zpracování na základě [7].

Druhy skupinové ICA

Existuje několik možností jak zpracovávat skupinová data. Jednou z možností je provést na každém měření ICA a výsledky spojit do skupin po výpočtu. Následně je porovnávat například pomocí klastrování nebo korelací. Další možností je vytvořit skupinová data pomocí spojení jednotlivých měření. Vytvoření skupinové matice lze provést spojením dat přes časovou osu (temporal concatenation) nebo přes prostorovou osu (spatial concatenation). V případě skupinové analýzy fMRI dat je možné použít oba přístupy. Časová ICA umožňuje sledovat unikátní TC jednotlivých subjektů, ale předpokládá společné SM. Oproti tomu prostorová ICA umožňuje sledovat unikátní SM za předpokladu společných TC. Výběr ICA tedy záleží na datovém souboru, a také na otázce, na kterou chceme pomocí ICA odpovědět. Spojení dat přes časovou složku je při analýze fMRI dat využíváno nejčastěji, a to z toho důvodu, že časové variace jsou značně větší než variace prostorové. Možností jak spojit data do skupinové matice či tensoru je mnoho a jedná se o důležitý krok skupinové ICA [8].

6.2.6 Analýza nezávislých vektorů

V mnoha aplikacích není potřeba analyzovat pouze jednotlivá data, ale také více souborů dat s mezi nimi existující závislostí. IVA zobecňuje problém ICA na více datových sad. IVA může být formulována pomocí minimalizace míry vzájemné informace stejně jako ICA, avšak nyní s přidáním další informace o závislosti mezi signály napříč datovými sadami. Výsledek ICA se stane zvláštním případem, když je počet sad dat roven jedné nebo když závislosti nejsou vzaty v potaz.

Kritérium optimalizace

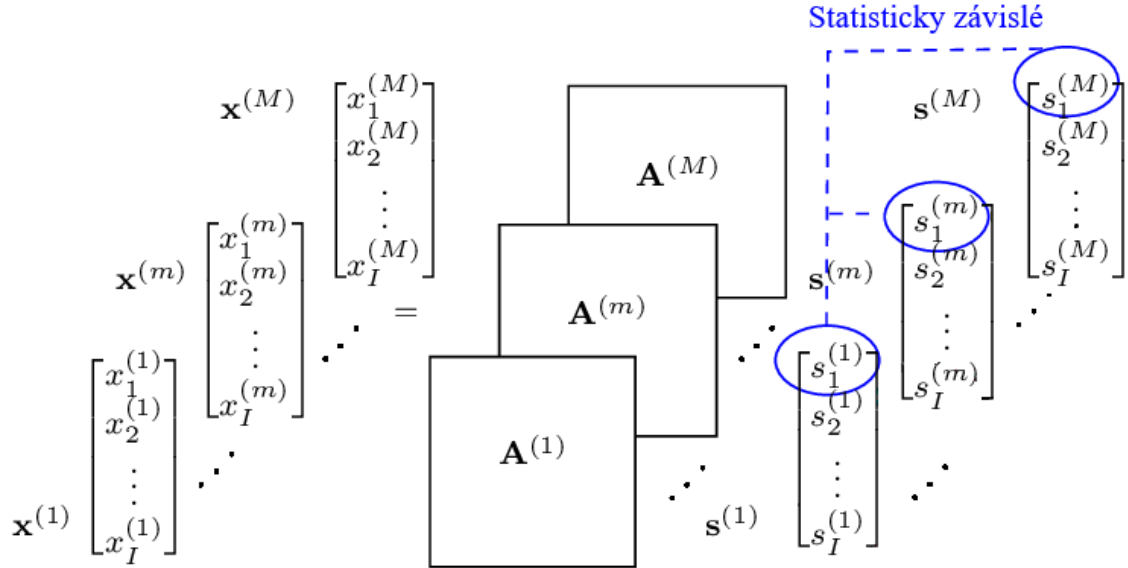
IVA je podobná formulaci ICA, jen s tím rozdílem, že je k dispozici K datových sad. Každá sada obsahuje T vzorků a je vytvořena pomocí lineárního míchání N nezávislých zdrojů.

$$\mathbf{X}^k(t) = \mathbf{A}^k \mathbf{s}^k(t), \quad 1 \leq k \leq K, \quad 1 \leq t \leq T, \quad (6.23)$$

kde $\mathbf{A}^k \in \mathbb{R}^{N \times N}$, $k = 1, \dots, K$ jsou matice mající inverzi a k je index sady. V tomto modelu jsou komponenty jednotlivých \mathbf{s}^k považovány za nezávislé, zatímco je umožněna závislost napříč odpovídajícími komponentami \mathbf{s}^k v datových sadách. Tato závislost mezi sadami je další informace, ze které IVA těží. Problém je nalézt K demixujících matic \mathbf{W}^k takových, aby zdroje pro každý datový soubor mohly být odhadnuty pomocí

$$\hat{\mathbf{s}}^k(t) = \mathbf{W}^k \mathbf{x}^k(t) \quad (6.24)$$

pro $k = 1, \dots, K$. IVA nespojuje datové sady dohromady jako skupinová ICA, ale využívá informace ze všech datových sad pro odhad jednotlivých nezávislých komponent.



Obrázek 6.3: Ilustrace IVA modelu. Vlastní zpracování na základě [27].

6.3 Analýza nezávislých komponent s konstantním separujícím vektorem

Je pozorováno N vzorků signálu pomocí d senzorů v K datových sadách. Vzorky lze rozdělit do $T \geq 1$ časových intervalů, které jsou označovány jako bloky. Všechny bloky N_b mají stejnou délku $N = T \cdot N_b$.

$$\mathbf{x}^{k,t} = \mathbf{A}^{k,t} \mathbf{s}^{k,t} \quad (6.25)$$

6.25 značí blokově se měnící model míchání, kde $k = 1, \dots, K$ je index datové sady; $t = 1, \dots, T$ je index bloku; $\mathbf{A}^{k,t}$, $\mathbf{s}^{k,t}$ a $\mathbf{x}^{k,t}$ reprezentují mixující matici, původní signály a pozorované signály v k směsi a v t bloku. Bez jakékoliv ztráty na obecnosti, jsou všechny signály uvažovány s nulovým průměrem. Vzorky signálu v každém bloku jsou považovány za nezávislé a identicky distribuované (i.i.d). Nyní s ohledem na nejasnosti 6.2.1 BSS lze definovat jako: „hledání demixující matice $\mathbf{W}^{k,t}$, takové aby $\mathbf{W}^{k,t}\mathbf{x}^{k,t}$ byly rovny $\mathbf{s}^{k,t}$ až na jejich škálu a fázi. Pořadí separovaných signálů se může lišit od původního, je však žádoucí, aby bylo stejné ve všech datových sadách a blocích. [26]“ V případě, kdy $T = 1$ se jedná o statické ICA 6.8 nebo IVA 6.24. Pokud je $T > 1$, jedná se o dynamický případ, kde se mixující matice může měnit blok od bloku. Když je použita ICA či IVA na každý blok samostatně, v případě $T > 1$ není zaručeno stejné pořadí separovaných signálů přes všechny bloky. Problém plyne z 6.10. Dále se sníží přesnost separace, jelikož je kontext omezen z celého signálu pouze na daný blok.

Separaci více než jednoho signálu lze provést jako několik BSE za sebou 6.3.3 nebo paralelně 6.3.4.

Extrakce jednoho zdroje

V případě extrakce jednoho zdroje je potřeba vzít v potaz pouze odpovídající sloupec matice $\mathbf{A}^{k,t}$ a odpovídající řádek matice $(\mathbf{A}^{k,t})^{-1}$. Zbylé sloupce matice $\mathbf{A}^{k,t}$ nemusí být odhadovány, je nutné pouze určit jejich podprostor. Matici $\mathbf{A}^{k,t}$ ze vzorce 6.25 lze parametrizovat stejně jako 6.26. Pro odvození parametrizace $\mathbf{A}^{k,t}$ lze nahlédnout do [26], [2]

$$\mathbf{A}_{\text{BSE}}^{k,t} = \begin{pmatrix} \mathbf{a}^{k,t} & \mathbf{Q}^{k,t} \end{pmatrix} = \begin{pmatrix} \gamma^{k,t} & (\mathbf{h}^{k,t})^H \\ \mathbf{g}^{k,t} & \frac{1}{\gamma^{k,t}} \left(\mathbf{g}^{k,t} (\mathbf{h}^{k,t})^H - \mathbf{I}_{d-1} \right) \end{pmatrix}, \quad (6.26)$$

kde $\mathbf{a}^{k,t} = [\gamma^{k,t}; \mathbf{g}^{k,t}]$ je mixující vektor odpovídající prvnímu sloupci matice $\mathbf{A}^{k,t}$; \mathbf{I}_d značí jednotkovou matici o velikosti $d \times d$; $(\mathbf{w}^{k,t})^H$ značí první řádek matice $(\mathbf{A}^{k,t})^{-1} = \mathbf{W}_{\text{BSE}}^{k,t}$; $\mathbf{w}^{k,t}$ značí separující vektor. Dále platí

$$\mathbf{W}_{\text{BSE}}^{k,t} = \begin{pmatrix} (\mathbf{w}^{k,t})^H \\ \mathbf{B}^{k,t} \end{pmatrix} = \begin{pmatrix} (\beta^{k,t})^* & (\mathbf{h}^{k,t})^H \\ \mathbf{g}^{k,t} & -\gamma^{k,t} \mathbf{I}_{d-1} \end{pmatrix}, \quad (6.27)$$

kde $\mathbf{w}^{k,t} = [\beta^{k,t}; \mathbf{h}^{k,t}]$ a $\mathbf{B}^{k,t} = [\mathbf{g}^{k,t}, -\gamma^{k,t} \mathbf{I}_{d-1}]$, splňující podmínku $\mathbf{B}^{k,t} \mathbf{a}^{k,t} = 0$ [26]. Jelikož $\mathbf{W}_{\text{BSE}}^{k,t} \mathbf{A}_{\text{BSE}}^{k,t} = \mathbf{I}_d$, jsou $\mathbf{a}^{k,t}$ a $\mathbf{w}^{k,t}$ propojeny pomocí takzvané „distortionless constraint“ $(\mathbf{w}^{k,t})^H \mathbf{a}^{k,t} = 1$. Lze zapsat také pomocí

$$(\beta^{k,t})^* \gamma^{k,t} = 1 - (\mathbf{h}^{k,t})^H \mathbf{g}^{k,t}. \quad (6.28)$$

Podprostor signálů, které je značen jako „pozadí“, je generován pomocí $\mathbf{z}^{k,t} = \mathbf{B}^{k,t} \mathbf{x}^{k,t}$. Výše zmíněná podmínka $\mathbf{B}^{k,t} \mathbf{a}^{k,t} = 0$ zaručuje, že $\mathbf{z}^{k,t}$ pokrývá stejný podprostor jako $s_2^{k,t}, \dots, s_d^{k,t}$. [26]

6.3.1 Konstantní separující vektor (CSV)

V případě ICA s modelem CSV je žádoucí separovat signály jeden za druhým nebo paralelně. Hlavní částí výpočtu je tedy vyřešení problému slepé extrakce jedné komponenty. ICE či IVE lze opakovat, dokud není extrahován požadovaný počet komponent. Pokud je použita extrakce pro $T > 1$ a pro každý blok separátně, znovu nastává problém s pořadím komponent. Problém pořadí lze vyřešit pomocí následujících modelů separace. Model s konstantním mixujícím vektorem, zapsaný v 6.29

$$\mathbf{a}^{k,1} = \mathbf{a}^{k,2} = \dots = \mathbf{a}^{k,T} = \mathbf{a}^k \quad (6.29)$$

a model s konstantním separujícím vektorem, zapsaným v 6.30

$$\mathbf{w}^{k,1} = \mathbf{w}^{k,2} = \dots = \mathbf{w}^{k,T} = \mathbf{w}^k. \quad (6.30)$$

Oba modely 6.29 i 6.30 zaručují správné pořadí komponent v jednotlivých blocích.

Zároveň je 6.30 vhodný pro separaci signálů, které se pohybují [26]. Aby bylo možné využít 6.30 pro separaci více než jednoho zdroje, je potřeba definovat následující podmínky:

- (P1) Všechny zdrojové signály značeny r se musí řídit 6.30, což znamená, že prvních r řádků matice $(\mathbf{A}^{k,1})^{-1}, \dots, (\mathbf{A}^{k,T})^{-1}$ ve vzorci 6.25 jsou konstantní přes t .
- (P2) Pro každé $i = 1, \dots, r$ se i zdroj řídí podle 6.30 v redukované směsi, kde zdroje $1, \dots, i - 1$ byly odečteny.

Platnost těchto podmínek je nutno předpokládat. Důkaz jejich užitečnosti lze nalézt v [26] v sekcích 1 a 5.

Statistický model

Je dána funkce hustoty pravděpodobnosti \mathbf{s}^t značená jako $p(\mathbf{s}^t)$. Autor poznamenává, že $p(\mathbf{s}^t)$ je závislá na t . Dále je dána $p_{\mathbf{z}^{k,t}}(\mathbf{z}^{k,t})$ označující hustotu pravděpodobnosti $\mathbf{z}^{k,t}$ signálů považovaných za pozadí.

Pokud je uvažována strukturu demixovací matice 6.27 s předpokladem CSV 6.30 a je využita nezávislosti mezi extrahovaným signálem a pozadím a zároveň jsou vzorky uvnitř bloku i.i.d, je možné získat společnou hustotu pravděpodobnosti pro jeden vzorek pozorovaných signálů v t bloku ve tvaru

$$p_{\mathbf{x}^{k,t}}(\{\mathbf{x}^{k,t}\}_k) = p\left(\left\{\left(\mathbf{w}^k\right)^H \mathbf{x}^{k,t}\right\}_k\right) \times \prod_{k=1}^K p_{\mathbf{z}^{k,t}}(\mathbf{B}^{k,t} \mathbf{x}^{k,t}) \left| \det \mathbf{W}_{\text{BSE}}^{k,t} \right|^2. \quad (6.31)$$

Druhá mocnina absolutní hodnoty determinantu je nutná kvůli transformaci hustot komplexních náhodných veličin. V případě, kdy je počítáno pouze s reálnými hodnotami je exponent roven jedné. Hustotu pravděpodobnosti pro všechny vzorky je

možné získat pomocí $\prod_{t=1}^T p_{\mathbf{x}^{k,t}} (\{\mathbf{x}^{k,t}\}_k)^{N_b}$, z čehož lze odvodit log-věrohodnostní funkci dělenou N jako

$$\begin{aligned} \mathcal{L} \left(\{\mathbf{w}^k, \mathbf{a}^{k,t}\}_{k,t} \right) = & \left\langle \hat{\mathbb{E}} \left[\log p \left(\left\{ (\mathbf{w}^k)^H \mathbf{x}^{k,t} \right\}_k \right) \right] + \right. \\ & \left. \sum_{k=1}^K \hat{\mathbb{E}} \left[p(\mathbf{B}^{k,t} \mathbf{x}^{k,t}) \right] + (d-2) \sum_{k=1}^K \log |\gamma^{k,t}|^2 \right\rangle_t \end{aligned} \quad (6.32)$$

Pro odvození funkcí zmíněných v této části lze nahlédnout do [25] sekce II. a [26] sekce III.

Kontrastní funkce

V předchozí sekci byla uvedena funkce 6.32. Nalezení maxima této funkce poskytuje maximálně věrohodný odhad vektorů parametrů. Jelikož nejsou známy hustoty pravděpodobnosti $p(\mathbf{s}^t)$ a $p_{\mathbf{z}^{k,t}}(\mathbf{z}^{k,t})$, je nezbytné tyto pravděpodobnostní hustoty nahradit vhodnými modelovými hustotami. Ve statickém případě lze modelové hustoty škálovat na jednotkový rozptyl, tímto je využito nejednoznačnosti ve škále 6.2.1. Pro dynamický případ je ovšem nutné zohlednit fakt, že rozptyl signálů se může měnit blok od bloku. Dle [26] lze $p(\cdot)$ nahradit pomocí

$$p(\mathbf{s}^t) \approx f \left(\left\{ \frac{s^{k,t}}{\hat{\sigma}^{k,t}} \right\}_k \right) \left(\prod_{k=1}^K \hat{\sigma}^{k,t} \right)^{-2}, \quad (6.33)$$

kde $f(\cdot)$ by mělo být vhodné normalizované ne-gaussovské pravděpodobnostní rozdělení a $(\hat{\sigma}^{k,t})^2$ značí rozptyl založený na odhadu $s^{k,t}$. Platí

$$\hat{\sigma}^{k,t} = \sqrt{(\mathbf{w}^k)^H \hat{\mathbf{C}}^{k,t} \mathbf{w}^k}, \quad (6.34)$$

kde $\hat{\mathbf{C}}^{k,t} = \hat{\mathbb{E}} \left[\mathbf{x}^{k,t} (\mathbf{x}^{k,t})^H \right]$ značí odhad kovarianční matice, ze vzorků signálů $\mathbf{x}^{k,t}$; $\hat{\sigma}^{k,t}$ je funkcí \mathbf{w}^k . Je důležité poznamenat, že $f(\cdot)$ může záviset na t . Dle [26] často neexistuje dostatek informací o skutečné hustotě pravděpodobnosti, a proto lze předpokládat, že $f(\cdot)$ je nezávislá na t a tím model 6.34 zjednodušit.

Neznámá $p_{\mathbf{z}^{k,t}}(\mathbf{z}^{k,t})$ byla v případě [26] nahrazena cirkulární hustotou pravděpodobnosti $\mathcal{CN}(0, \mathbf{C}_{\mathbf{z}}^{k,t})$ představující komplexní gaussovo rozložení s nulovou střední hodnotou a kovarianční maticí signálů pozadí. Pro modelování $p_{\mathbf{z}^{k,t}}(\mathbf{z}^{k,t})$ bylo $\mathcal{CN}(0, \mathbf{C}_{\mathbf{z}}^{k,t})$ vybráno, protože komponenty $\mathbf{z}^{k,t}$ jsou smíšené a korelované. Z centrální limitní věty potom plyne, že distribuce těchto veličin se přibližuje k normálnímu rozdělení. $\mathbf{C}_{\mathbf{z}}^{k,t}$ je parametr, který později bude nahrazen odhadem konkrétních vzorků. Po dosazení výše uvedených parametrů do rovnice 6.32 dostává kontrastní funkce

následující tvar

$$\begin{aligned} \mathcal{C} \left(\{ \mathbf{w}^k, \mathbf{a}^{k,t} \}_{k,t} \right) &= \left\langle \hat{\mathbb{E}} \left[\log f \left(\left\{ \frac{\hat{s}^{k,t}}{\hat{\sigma}^{k,t}} \right\}_k \right) \right] \right. \\ &\quad - \sum_{k=1}^K \log (\hat{\sigma}^{k,t})^2 - \sum_{k=1}^K \hat{\mathbb{E}} \left[(\hat{\mathbf{z}}^{k,t})^H (\mathbf{C}_z^{k,t})^{-1} \hat{\mathbf{z}}^{k,t} \right] \\ &\quad \left. + (d-2) \sum_{k=1}^K \log |\gamma^{k,t}|^2 \right\rangle_t + \text{const.}, \end{aligned} \quad (6.35)$$

kde $\hat{s}^{k,t} = (\mathbf{w}^k)^H \mathbf{x}^{k,t}$ a $\hat{\mathbf{z}}^{k,t} = \mathbf{B}^{k,t} \mathbf{x}^{k,t}$. Konstanta na konci výrazu je nezávislá na mixujících parametrech modelu. Pro případ $K = 1$ a $T = 1$ lze 6.35 zjednodušit na

$$\begin{aligned} \mathcal{C}^{1,1}(\mathbf{w}, \mathbf{a}) &= \hat{\mathbb{E}} \left[\log f \left(\frac{\hat{s}}{\hat{\sigma}} \right) \right] - \log \hat{\sigma}^2 - \hat{\mathbb{E}} \left[\hat{\mathbf{z}}^H \mathbf{C}_z^{-1} \hat{\mathbf{z}} \right] \\ &\quad + (d-2) \log |\gamma|^2 + \text{const.} \end{aligned} \quad (6.36)$$

Podmínka ortogonality

Kontrastní funkce 6.35 může mít mnoho falešných extrémů [26]. Může se stát že, $\mathbf{a}^{k,t}, t = 1, \dots, T$ a \mathbf{w}^k neodpovídají stejnému signálu, proto je nutné zavést spolehlivé propojení mezi mixujícími a separujícími vektory. Takovým vhodným propojením je podmínka ortogonality (OGC) [26]. Jelikož jsou $s^{k,t}$ a $\mathbf{z}^{k,t}$ nezávislé, tudíž zároveň i nekorelované, vyžaduje OGC, aby podprostor generovaný vzorky $\hat{s}^{k,t}$ byl ortogonální vůči podprostoru $\hat{\mathbf{z}}^{k,t}$. Zároveň musí platit 6.28. Za těchto podmínek platí

$$\mathbf{a}^{k,t} = \frac{\hat{\mathbf{C}}^{k,t} \mathbf{w}^k}{(\mathbf{w}^k)^H \hat{\mathbf{C}}^{k,t} \mathbf{w}^k}, \quad (6.37)$$

kde je závislá proměnná \mathbf{w} nebo jako

$$\mathbf{w}^k = \frac{\left(\hat{\mathbf{C}}^{k,t} \right)^{-1} \mathbf{a}^{k,t}}{(\mathbf{a}^{k,t})^H \left(\hat{\mathbf{C}}^{k,t} \right)^{-1} \mathbf{a}^{k,t}}, \quad (6.38)$$

kde je závislou proměnnou \mathbf{a} . Odvození rovnic 6.37 a 6.38 lze nalézt v [25], konkrétně se jedná o přílohu A.

6.3.2 One-unit algoritmus

One-unit FastICA je navržena pro extrakci jednoho zdrojového signálu ze směsi v případě, kdy $K = 1$ a $T = 1$, optimalizací funkce 6.36 za podmínky 6.37. Pokud byla data předzpracována a byla odstraněna korelace, tedy $\hat{\mathbf{C}} = \mathbf{I}_d$, vypadá vzorec pro výpočet nových vah \mathbf{w} následovně:

$$\mathbf{w} \leftarrow E \left\{ \mathbf{x} (\mathbf{w}^H \mathbf{x})^* g \left(|\mathbf{w}^H \mathbf{x}|^2 \right) \right\} - E \left\{ g \left(|\mathbf{w}^H \mathbf{x}|^2 \right) + |\mathbf{w}^H \mathbf{x}|^2 g' \left(|\mathbf{w}^H \mathbf{x}|^2 \right) \right\} \mathbf{w}, \quad (6.39)$$

pro případ komplexních proměnných, kde $g(\cdot)$ je vhodná nelineární funkce. Odvození 6.39 a příklady vhodných nelineárních funkcí lze nalézt v [4] nebo [26]. Po každé aktualizaci je \mathbf{w} normalizováno, jelikož jsou prvky vektoru nezávislé, je $\hat{\sigma} = 1$. [20]

One-unit FastDIVA je rozšířením One-unit FastICA pro $T > 1$ a $K > 1$ pod podmínkou CSV 6.30. One-unit FastDIVA je zároveň rozšířením a zjednodušením One-unit FastICA v případě signálů s neznámou fází s komplexními hodnotami [26].

6.3.3 Deflation přístup

Deflation přístup odhaduje nezávislé komponenty jednu po druhé a využívá k tomu One-unit algoritmus. Pokud je odhadováno t nezávislých komponent nebo t vektorů $\mathbf{w}_1, \dots, \mathbf{w}_t$, bude potřeba $t + 1$ iterací One-unit přístupu, kde po každé iteraci bude odebrána z \mathbf{w}_{t+1} projekce $(\mathbf{w}_{t+1}^T \mathbf{w}_j) \mathbf{w}_j, j = 1, \dots, t$ z dříve odhadnutých t vektorů a následně je renormalizován \mathbf{w}_{t+1} . Tento postup se nazývá Gram-Schmidt ortogonalizace [20]. Algoritmus lze shrnout následovně:

1. Je vybráno, kolik IC je odhadováno. Je nastaveno $t = 1$.
2. Je inicializováno \mathbf{w}_t .
3. Je provedena iterace One-unit algoritmu na \mathbf{w}_t .
4. Je provedena ortogonalizace podle:

$$\mathbf{w}_t \leftarrow \mathbf{w}_t - \sum_{j=1}^{t-1} (\mathbf{w}_t^T \mathbf{w}_j) \mathbf{w}_j \quad (6.40)$$

5. Je normalizován \mathbf{w}_t tak, že je vydělen jeho normou.
6. Pokud \mathbf{w}_t nekonverguje, vrací se algoritmus na krok 3.
7. Je nastaveno $t \leftarrow t + 1$. pokud t není větší než počet odhadovaných IC, pokračuje algoritmus od kroku číslo 2.

Postupné odečítání nalezených IC má přímý dopad na výsledek, jelikož záleží na tom, v jakém pořadí jsou extrahovány. Pokud je počet extrahovaných IC menší než celkový počet možných IC, je možné získat pokaždé jiné výsledky. Například mix o třech zdrojových signálech. Jsou extrahovány dva ze tří možných signálů pomocí deflation přístupu. Počet možných výsledků extrahovaných signálů je roven variaci bez opakování tedy $V_2(3) = 6$ Jaký signál bude extrahován a v jakém pořadí značně záleží na inicializaci \mathbf{w}_i .

6.3.4 Symmetric přístup

Symmetric přístup je vhodný pro případy, kdy není upřednostněn žádný ze signálů. To znamená, že vektory \mathbf{w}_i nejsou odhadovány jeden po druhém, ale zároveň.

Tento přístup umožňuje odhadovat IC paralelně. Zároveň je oproti deflation přístupu rozdílný v tom, že nekumuluje chybu odhadu, která se může zvyšovat při odečtení každého signálu a při následné ortogonalizaci. Symmetric přístup funguje následovně:

1. Je vybráno kolik IC bude odhadnuto a tento počet je označen jako r .
2. Je inicializováno $\mathbf{w}_i, i = 1, \dots, r$.
3. Paralelně je provedena iterace One-unit algoritmu na každém \mathbf{w}_i .
4. Je provedena symetrická ortogonalizace matice \mathbf{W} , kde $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_n]$.
5. Pokud nedošlo ke konvergenci, algoritmus je opakován od kroku 3.

V kroku 4 je zmíněna symetrická ortogonalizace, kterou je možno provést následujícím způsobem [4].

$$\mathbf{W} \leftarrow \mathbf{W}(\mathbf{W}^H \mathbf{W})^{-\frac{1}{2}}$$

FastDIVA s přístupem symmetric tak separuje r nezávislých signálů tím, že alternuje mezi odhadem separujících vektorů a jejich ortogonalizací, dokud nezkonverguje.

6.3.5 Block-Deflation přístup

Tento přístup byl vytvořen pro směsi 6.25 splňující podmínku 6.30. Block-Deflation přístup klade silnější podmínku na extrahované signály tím, že je činí ortogonálními samostatně v každém bloku. Nejdříve je z původního mixu extrahován jeden signál pomocí One-unit algoritmu. Extrahovaný signál je následně odečten z původního mixu z každého bloku. Následně je One-unit algoritmus aplikován na nově vzniklá data, která vznikla odečtením prvního signálu, čímž lze extrahovat druhý signál. Tento postup je opakován dokud není extrahováno r požadovaných signálů [26].

Je dán vstupní signál z t bloku v i fázi Block-Deflation přístupu a je označen jako \mathbf{x}_i^t . Dále je dán také separující vektor \mathbf{w}_i , který byl získán po jedné iteraci One-unit algoritmu aplikovaného na \mathbf{x}_i^t . Nová data \mathbf{x}_{i+1}^t jsou získána odečtením $\hat{s}_i^t = \mathbf{w}_i^H \mathbf{x}_i^t$ z \mathbf{x}_i^t metodou nejmenších čtverců. Vzhledem k OGC 6.37 jsou nová data získána pomocí

$$\mathbf{x}_{i+1}^t = \mathbf{\Pi}_i^t \mathbf{x}_i^t, \quad (6.41)$$

kde $\mathbf{\Pi}_i^t = \mathbf{E}_i (\mathbf{I}_{d-i+1} - \tilde{\mathbf{a}}_i^t \mathbf{w}_i^H)$ a $\tilde{\mathbf{a}}_i^t$ je odhadovaný mixující vektor t bloku odpovídající i extrahovanému signálu z dat \mathbf{x}_i^t . \mathbf{E}_i je matice, která redukuje dimenzi \mathbf{x}_{i+1}^t o jedna, na \mathbf{x}_i^t , jehož dimenze je $d - i + 1$. \mathbf{E}_i neobsahuje žádné závislé řádky. Vektory \mathbf{w}_i a $\tilde{\mathbf{a}}_i^T$ jsou odhadovány z \mathbf{x}_i^t pomocí One-unit algoritmu, což znamená, že jsou provázány pomocí OGC 6.37. Postup Block-Deflation přístupu lze v krocích popsat následovně:

1. Je vybráno, kolik IC bude odhadováno. Je nastaveno $r = 1$.
2. Je inicializován \mathbf{w}_i .

3. Pomocí One-unit algoritmu je extrahován signál.
4. Extrahovaný signál je odečten z každého bloku.
5. Postup je rekurzivně opakován od kroku 3. dokud není extrahováno r IC.

7 Skupinová analýza fMRI dat s využitím metod ICA a ICA CSV

V současné době hrají metody analýzy obrazových dat získaných z magnetické rezonance klíčovou roli v diagnostice a výzkumu různých neurologických a neurodegenerativních onemocnění.

Praktická část této diplomové práce provádí exploratorní analýzu fMRI dat pomocí metod ICA a ICA CSV s cílem identifikovat a izolovat různé zdroje signálu v fMRI obrazech mozku. Hlavní důraz je kladen na nalezení nezávislých komponent, které jsou podle vybraného kritéria nejvěrnější vzorovým nezávislým komponentám.

Pro účely analýzy budou data předzpracována pomocí PCA a následně budou použity metody ICA a ICA CSV pro separaci zdrojů signálu.

Důležitou součástí této práce je porovnání výsledků mezi metodami ICA a ICA CSV s cílem zhodnotit přínos rozšíření CSV pro analýzu MRI dat. V neposlední řadě bude také provedena validace navržených postupů analýzy na kontrolních datech, aby byla ověřena robustnost a spolehlivost těchto metod.

Autor očekává, že tato diplomová práce přinese nové poznatky a přispěje k efektivnějšímu použití ICA CSV na fMRI datech. Výsledky by mohly mít potenciální využití v klinické praxi pro podporu diagnostiky neurologických onemocnění a také v neurovědních studiích, které se zabývají komplexním zkoumáním struktur a funkcí mozku.

7.1 Metodika

V této sekci jsou popsány postupy a metody, které jsou použity v praktické části diplomové práce k provádění exploratorní analýzy dat z magnetické rezonance za pomoci skupinové ICA a ICA CSV namísto ICA jednoho subjektu z několika důvodů.

Prvním důvodem je zvýšení statistické síly. Skupinová analýza kombinuje data z více subjektů do jednoho souboru, což zvyšuje statistickou sílu analýzy. Vzhledem k tomu, že skupinová ICA a ICA CSV pracují s vyšším počtem vzorků, je možné dosáhnout spolehlivějších a robustnějších výsledků než při analýze jednotlivých subjektů.

Dalším důvodem je, že skupinová ICA a ICA CSV umožňují identifikovat globální vzory aktivity, které jsou společné pro celou skupinu. Tato schopnost je zvláště užitečná, když je cílem studium obecných funkcí mozku nebo když jsou porovnávány aktivační vzory mezi různými skupinami subjektů (např. zdraví a nemocní).

Posledním důvodem je snížení vlivu šumu. Skupinová analýza umožňuje statisticky zvýšit poměr signálu k šumu, což může vylepšit identifikaci relevantních vzorů aktivity. Jednotlivé subjekty mohou mít různé úrovně šumu, a proto agregace dat ve skupinové analýze může zvýšit poměr signálu k šumu a zlepšit detekci aktivních oblastí mozku.

7.1.1 Data

První datová sada obsahovala fMRI data dvaceti subjektů. Tato data prošla náležitým předzpracováním podle sekce 3 pro skupinové zpracování. Navíc byly řádky jednotlivých matic normalizovány na nulovou střední hodnotu a směrodatnou odchylku rovnou jedné. Matice jednoho subjektu obsahuje 152 snímků a každý snímek obsahuje 71 523 voxelů. K této sadě také patří maska pro reprezentaci dat v 3D prostoru. Tato maska měla rozměry $53 \times 63 \times 52$ a byla nedílnou součástí při tvorbě výsledných obrázků.

Pro nezávislou kontrolu výsledků byla vybrána data z databáze fBIRN. Tato data byla převzata z článku [14], kde jsou k dispozici kompletní informace o jejich původu a vlastnostech. Zde budou uvedeny pouze vlastnosti, které jsou nezbytné pro další práci. Na datech jednotlivých subjektů byla provedena PCA, jež redukovala data na 110 hlavních komponent. Data jednotlivců byla spojena dohromady a byla provedena skupinová PCA, která je zredukovala na 100 hlavních komponent. Dále byl použit Infomax algoritmus [3], který dekomponoval hlavní komponenty na nezávislé komponenty. Tento postup byl opakován stokrát s využitím techniky ICASSO [18] a ICA s nejlepšími výsledky byla vybrána pro vygenerování 100 IC. Výsledná data jsou rozdělena na IC a TC. U každého subjektu je IC rovno matici o rozměrech 100×58425 , tedy 100 IC o 58425 voxelích. TC je matice o rozměrech 157×100 , tedy 157 snímků. Datová matice pro každý subjekt je potom pouze jejich násobek podle vzorce 5.1. Pro tuto práci bylo vybráno pouze dvacet subjektů, aby bylo možné porovnat výsledky s první datovou sadou. K těmto SM byla také dodána maska, která měla stejné rozměry jako maska z první datové sady.

Z Neuromark článku [14] byly také převzaty vzory, ke kterým byly porovnávány nalezené výsledky.

7.1.2 Metody analýzy

Tato analýza bude zaměřena na identifikaci vzájemně nezávislých IC a možných variací mezi různými subjekty v rámci celé skupiny. Prvním krokem bylo vytvoření skupinové datové matice z fMRI dat jednotlivých subjektů. Matice subjektů byly spojeny přes časovou dimenzi (Temporal concatenation) 6.2a. Výhodou tohoto přístupu je potřeba pouze jednoho provedení ICA. Toto spojení také umožňuje nalézt jedinečné TC pro každý subjekt avšak předpokládá společné SM. Časové spojení je široce používáno pro skupinovou ICA na datech z fMRI [7], s největší pravděpodobností proto, že časové variace v signálu fMRI jsou mnohem větší než prostorové variace. Nechť \mathbf{Y}_i je $C \times V$ předzpracovaná a normalizovaná datová matice i -tého subjektu, kde C představuje časové body (jednotlivé snímky fMRI) a V značí namě-

řené voxely přes M subjektů. Potom $\mathbf{Y} \equiv [\mathbf{Y}_1^C, \dots, \mathbf{Y}_M^C]$, $\mathbf{Y}_i \in \mathbf{R}^{C \times V}$, $\mathbf{Y} \in \mathbf{R}^{MC \times V}$ jsou prostorově spojená data jednotlivých subjektů. Jedná se o případ z obrázku 6.2b.

Tato skupinová matice je následně zpracována pomocí PCA, které je věnována samostatná sekce. Výstupní redukováná matice z PCA je použita jako vstup pro metody ICA a ICA CSV. Jednotlivým metodám je věnována samostatná sekce. Pro výpočet ICA a ICA CSV byly vybrány algoritmy FastDIVA a EFICA. Algoritmus FastDIVA byl vybrán, protože se jedná o rozšíření populárního algoritmu FastICA. Pod hlavičkou FastDIVA lze po nastavení správných parametrů počítat rychle a efektivně klasickou ICA, ale také rozšíření ICA CSV. Algoritmus EFICA byl vybrán pro porovnání výsledků s klasickou ICA. Tento algoritmus byl zvolen, protože se jedná o efektivní a rychlý výpočet ICA v zarušeném prostředí [24].

Korelační koeficient

Nalezené IC mají náhodné pořadí a pro vyhodnocení podobnosti je nutné jejich spárování se vzorovými IC, k čemuž byl využit korelační koeficient. Nejdříve byla spočtena korelace mezi nalezenými IC, značeno jako $\hat{\mathbf{X}}^{C \times V}$ a vzorovými IC, značeno jako $\mathbf{X}^{14 \times V}$. Pro každý sloupec $1 \leq n \leq 14$ ze vzorové matice \mathbf{X} byl nalezen vektor \mathbf{c}_n podle

$$\mathbf{c}_n = \frac{\hat{\mathbf{X}} \mathbf{X}_n^T}{V}$$

kde V je počet voxelů a \mathbf{c}_n má velikost 90×1 a obsahuje hodnoty korelace mezi nalezenými IC a n tou vzorovou IC. První dimenze vektoru \mathbf{c}_n je rovna 90, což odpovídá velikosti výstupní matice z PCA, které je věnována samostatná sekce. Abychom předešli problému se znaménkem, je z \mathbf{c}_n vybraná maximální absolutní hodnota $\max|\mathbf{c}_n|$, kde index této hodnoty představuje číslo nalezené IC co nejvíce podobné vzorové IC. Pro čtrnáct vybraných IC z $\hat{\mathbf{X}}$ je spočítán Pearsonův korelační koeficient s odpovídající vzorovou IC. Na základě tohoto koeficientu byla vyhodnocena úspěšnost a přesnost nalezených IC.

7.2 Předzpracování dat pomocí PCA

PCA je běžným krokem předzpracování skupinové matice pro ICA analýzu [1] [8]. Výsledkem provedení PCA na matici \mathbf{Y} je nová matice \mathbf{Z} obsahující hlavní komponenty.

Tyto hlavní komponenty mohou být použity jako nová skupinová matice pro ICA. ICA následně rozkládá tuto novou matici do vzájemně nezávislých IC a TC. Výhodou použití PCA jako předpřípravy je snížení dimenzionality dat a eliminace šumu. To může vést k lepší identifikaci relevantních vzorců aktivity a zvýšení spolehlivosti analýzy.

Celkově kombinace PCA a ICA umožňuje snížit náročnost analýzy dat a získat kompaktnější a interpretovatelnější reprezentaci funkční aktivity mozku v rámci celé

skupiny subjektů. Tímto způsobem lze lépe porozumět obecným vzorcům aktivace a variabilitě v mozku mezi různými subjekty.

7.3 Hledání optimální hodnoty pro redukcí dimenze pomocí PCA

Otázkou je, jaká je optimální hodnota pro redukcí dimenze? Pro nalezení odpovědi byl vytvořen následující experiment. Tento experiment měl za úkol najít optimální hodnotu R představující velikost matice po redukcí.

Matice $\mathbf{Y}^{MC \times V}$, která obsahuje data jednotlivých subjektů a je spojena přes časovou dimenzi. Na $\mathbf{Y}^{MC \times V}$ bylo provedeno několik nezávislých redukcí pomocí PCA. Dle publikace [1] se dimenze u jednotlivých subjektů redukcí ze stovek přibližně na 30 až 60 a záleží vždy na konkrétních datech. Z tohoto důvodu byla zvolena počáteční hodnota parametru $R = 14$, jelikož vzorových IC je právě 14. Jako maximální možnou velikost parametru R byla zvolena hodnota, která dle výše zmíněné publikace překročila horní hranici s rezervou. Horní limit byl nastaven na $R = 200$.

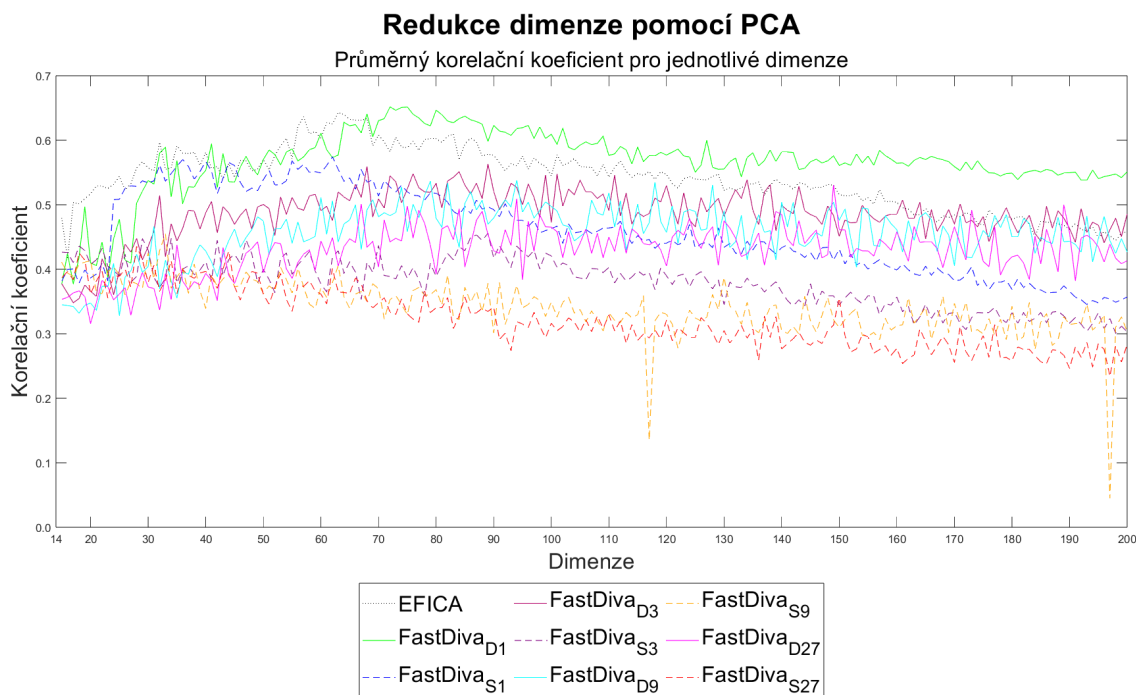
7.3.1 Průběh experimentu

Nejprve byla z matice \mathbf{Y} vypočítána kovarianční matice, která pak byla rozložena na vlastní vektory a vlastní čísla tak, aby platilo $\mathbf{Y}\mathbf{V} = \mathbf{V}\mathbf{D}$, kde \mathbf{V} je matice pravých vlastních vektorů a \mathbf{D} je matice vlastních čísel. Následně bylo z matice \mathbf{V} vybráno R vlastních vektorů, které odpovídají prvním R hlavním komponentám. Posledním krokem bylo vytvoření matice \mathbf{Z} pomocí vzorce 7.1

$$\mathbf{Z} = \mathbf{V}^T \mathbf{Y} \quad (7.1)$$

a vzniká nová matice $\mathbf{Z}^{R \times V}$. Zde bylo R nastavováno na hodnoty $14 \leq R \leq 200$, a bylo tedy provedeno 186 nezávislých redukcí dimenze a následné nalezení prostorových map v každém spuštění. Po provedení redukcí dimenze byly spočítány dvě varianty ICA a šest variant ICA CSV. Pro zjištění optimální hodnoty R byly pro každou redukcí nalezené IC porovnány se vzorovými IC dle korelačního koeficientu 7.1.2. Pro každou vzorovou IC je přes absolutní hodnotu korelačního koeficientu vybrána nalezená IC s nejvyšší hodnotou. Stalo se, že k více vzorovým IC byla přiřazena jako nejlepší shoda jedna nalezená IC. Pokud se tak stane, nebyl výsledek nijak upravován. Ve většině případů se totiž jednalo o komponenty plné šumu nebo nepoužitelné výsledky, tedy případy, kdy se ICA nepodařilo v datech najít odpovídající vzorové IC. Ilustrací takového případu je obrázek 7.6, konkrétně IC 4 pro algoritmus ICA.

Obrázek 7.1 shrnuje výsledky průměrného korelačního koeficientu přes 14 nalezených IC pro každé provedení redukcí dimenze. Výsledky nalezené pomocí EFICA a klasické ICA (značeno jako $FastDIVA_{D1}$ a $FastDIVA_{S1}$) nachází nejlepší výsledky při nižší hodnotě R než výsledky nalezené pomocí ICA CSV. U block-deflation ICA CSV metod byly nejlepší výsledky nalezeny kolem $R = 90$ a růst parametru



Obrázek 7.1: Průměrný korelační koeficient pro jednotlivé dimenze.

R již nepřinášel zlepšení výsledků. Parametr R byl tedy na základě těchto výsledků zvolen jako optimální na hodnotě 90.

7.4 Závislost nalezených IC na inicializační matici

V ideálním případě by měl algoritmus nalézt komponenty bez jakékoliv předchozí informace. Je však nezbytné vybrat správný algoritmus a jeho parametry, které přímo ovlivňují výsledek. Z nejednoznačnosti v pořadí extrahovaných komponent plyne i nejednoznačnost v pořadí extrakce. Toto je neodmyslitelná součást problému BSS/BSE. Ovlivnit extrakci lze zvolenou inicializací a přístupem algoritmu. V sekci 6.3.3 a 6.3.5 je pořadí extrahovaných komponent určeno zvoleným přístupem. Zároveň záleží na počtu extrahovaných komponent.

Tento experiment se zaměřuje na inicializaci algoritmů, které jsou podle [26] důležitým krokem pro spolehlivost separace. Zároveň bere v potaz rozdíly přístupů popsané v sekcích 6.3.3, 6.3.4 a 6.3.5. Inicializace algoritmů umožňuje algoritmus nasměrovat určitým směrem k požadovanému výsledku. V tomto experimentu inicializační matice také přímo udává, kolik komponent bude separováno.

7.4.1 Inicializační matice

Bylo vybráno 6 typů inicializační matice \mathbf{W} . Matice \mathbf{W}_A , \mathbf{W}_C a \mathbf{W}_E jsou čtvercové a byly u nich využity oba zmíněné algoritmy EFICA a FastDIVA. Matice \mathbf{W}_B , \mathbf{W}_D a \mathbf{W}_F jsou obdélníkové. V tomto případě bylo tedy možné využít jenom FastDIVA,

jelikož EFICA počítá pouze se čtvercovou maticí. Popis a vytvoření inicializačních matic popisuje následující sekce.

Vytvoření inicializačních matic

Matrice $\mathbf{W}_A^{90 \times 90}$ obsahuje náhodné hodnoty normálního rozložení s nulovou střední hodnotou a jednotkovým rozptylem. Představuje inicializaci v duchu BSS, při níž $\mathbf{W}_A^{90 \times 90}$ algoritmu nepředává žádnou informaci o hledaných signálech. V případě $\mathbf{W}_A^{90 \times 90}$ je prohledáván celý prostor, tedy je hledáno 90 IC z možných 90. Pokud data obsahují vzorové IC, tato inicializace by je měla nalézt. Matice $\mathbf{W}_B^{90 \times 14}$ je náhodná matice. U $\mathbf{W}_B^{90 \times 14}$ je hledáno pouze 14 IC z možných 90, je zde tedy šance, že nebude nalezena vzorová IC. Matice $\mathbf{W}_C^{90 \times 90}$ byla vytvořena pomocí informace vzorových komponent. Nejdříve byla vygenerována náhodná matice o rozměrech 90×90 a bylo nahrazeno prvních 14 sloupců hodnotami, které jsou podle metody nejmenších čtverců (LMS) nejpodobnější hledaným vzorům. Tato matice byla označena jako $\tilde{\mathbf{W}}_C^{90 \times 90}$. Následně byla přičtena náhodná matice o stejné velikosti, která reprezentovala šum.

$$\mathbf{W}_C^{90 \times 90} = \tilde{\mathbf{W}}_C^{90 \times 90} + 0,5 * \mathbf{W}^{90 \times 90} \quad (7.2)$$

Matrice $\mathbf{W}_D^{90 \times 14}$ byla vytvořena podobně jako matice $\mathbf{W}_C^{90 \times 90}$. Nejprve byla vygenerována $\mathbf{W}_B^{90 \times 14}$, u které byly podle LMS nahrazeny všechny sloupce hodnotami nejpodobnějšími vzorovým komponentám. Následně byla přičtena matice představující šum. Matice $\mathbf{W}_E^{90 \times 90}$ je rovna matici $\mathbf{W}_C^{90 \times 90}$ bez přičtení matice šumu. Obsahuje tedy prvních 14 sloupců podobných vzorům a zbytek matice je náhodný. Matice $\mathbf{W}_F^{90 \times 14}$ je rovna matici $\mathbf{W}_D^{90 \times 14}$ bez přičtení matice šumu. Parametr pro počet hledaných nezávislých signálů pomocí FastDIVA byl nastaven u všech podle počtu sloupců inicializační matice. U matic $\mathbf{W}_A, \mathbf{W}_C$ a \mathbf{W}_E bylo hledáno 90 signálů a u $\mathbf{W}_B, \mathbf{W}_D$ a \mathbf{W}_F bylo hledáno 14 signálů. V případě EFICA bylo hledáno 90 signálů a byly využity inicializace $\mathbf{W}_A, \mathbf{W}_C$ a \mathbf{W}_E . Inicializace pomocí matic \mathbf{W}_E a \mathbf{W}_F byla zvolena uměle, jelikož se v inicializační matici se již nachází věrná data ohledně hledaných komponent. Tato inicializace byla vyzkoušena záměrně, cílem bylo algoritmu pomoci dosáhnout co nejlepšího výsledku. Matice \mathbf{W}_C a \mathbf{W}_D představují mezikrok v inicializaci algoritmu pomocí vzorů a náhody. \mathbf{W}_C a \mathbf{W}_D jsou vytvářeny s pomocí informace o hledaných komponentách, ke kterým je přidán náhodný šum. Důvodem pro využití inicializace pomocí \mathbf{W}_A a \mathbf{W}_B byla simulace reálného problému BSS. V něm totiž není definováno, co má algoritmus najít, ani přesný počet komponent které lze najít a obsah informace v těchto maticích nijak nenavádí k nalezení vzorových komponent. Rozdělení matic na čtvercové $\mathbf{W}_A, \mathbf{W}_C, \mathbf{W}_E$ a obdélníkové $\mathbf{W}_B, \mathbf{W}_D, \mathbf{W}_F$ vedlo k vyzkoušení chování algoritmu FastDIVA, který lze inicializovat i obdélníkovou maticí. Pokud FastDIVA hledá v prostoru o dimenzi 90 právě 90 nezávislých komponent, najde jiné komponenty, než když hledá pouze 14 nezávislých komponent.

Problémem obdélníkové inicializační matice je skutečnost, že nezaručuje nalezení právě 14 nejlepších komponent, ale pouze nalezení 14 komponent. V tomto případě je také obtížné kontrolovat nebo validovat výsledky, jelikož při každém spuštění

může algoritmus najít 14 různých komponent. I kdyby našel 14 stejných komponent je velmi malá pravděpodobnost, že budou ve stejném pořadí a se stejnou škálou a znaménkem. Není tedy zřejmé, zda se jedná o již dříve nalezenou komponentu.

Výsledky a poznatky experimentu

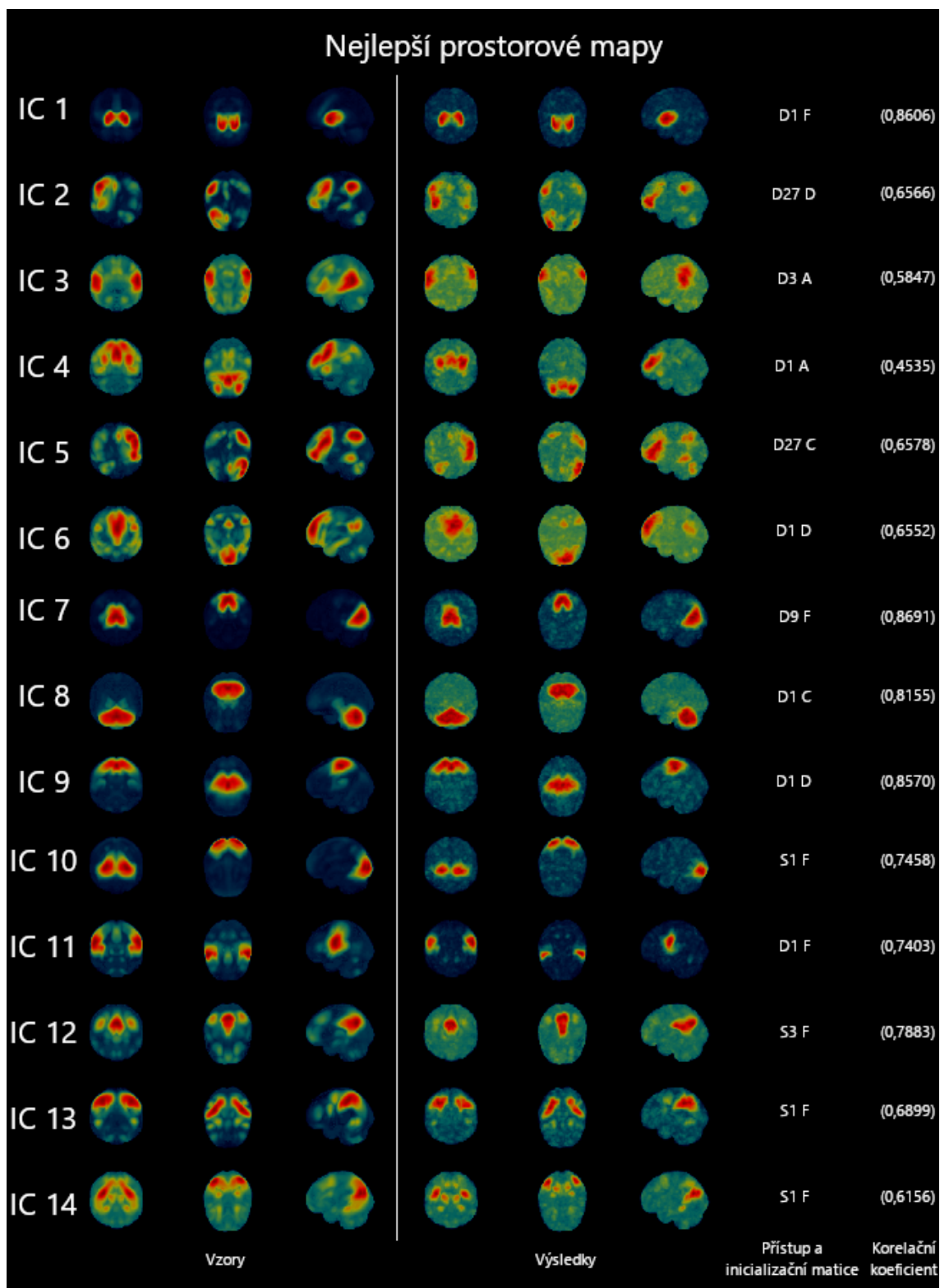
Pro první datový soubor ze sekce 7.1.1 byl spuštěn jeden cyklus následujícího pokusu. Nejdříve proběhla předpříprava dat podle sekce 7.4, kde byla spojena data jednotlivých subjektů do jedné společné matice. Byla provedena PCA, která redukovala dimenzi dat na $\mathbf{Z}^{90 \times N}$. Následně byly načteny vzorové komponenty a připraveny inicializační matice. Vstupem pro algoritmus EFICA byly pouze inicializační matice \mathbf{W}_A , \mathbf{W}_C , \mathbf{W}_E , tedy všechny čtvercové matice a také datová matice \mathbf{Z} . Výstupem je $\mathbf{W}_{EFICA}^{90 \times 90}$, demixující matice, ze které jsou dopočítány prostorové mapy pomocí vzorce 7.3

$$\mathbf{P}_{mapy} = \mathbf{W}_{EFICA}^T \cdot \mathbf{Z} \quad (7.3)$$

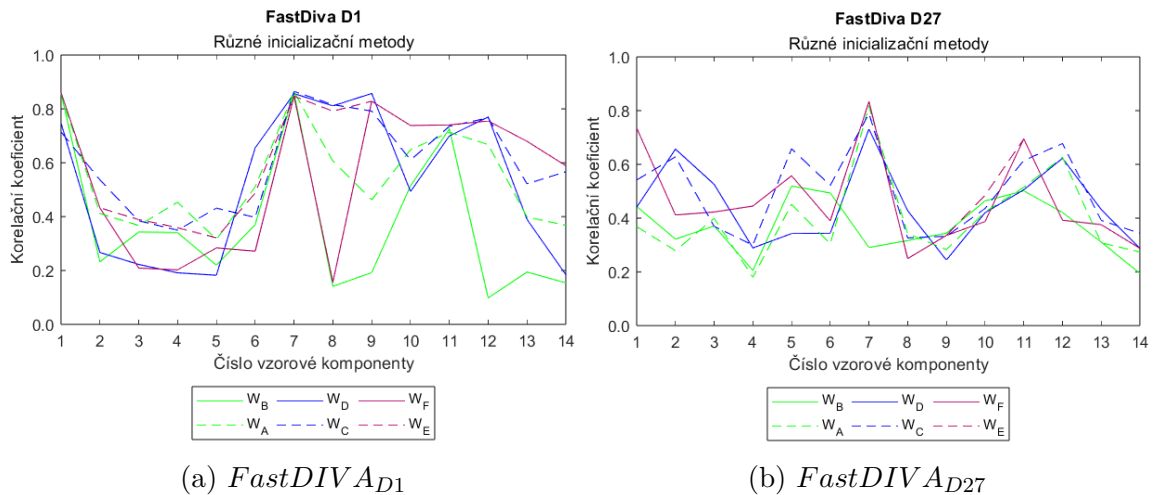
Následně jsou \mathbf{P}_{mapy} normalizovány pomocí euklidovské normy. Tyto mapy jsou porovnány podle korelačního koeficientu 7.1.2 s mapami vzorovými. Ke každé vzorové IC je přiřazena nalezená IC s nejvyšší hodnotou korelačního koeficientu bez ohledu na znaménko.

Tento postup řeší problém pořadí nalezených komponent. Nejlepší nalezené prostorové mapy podle korelačního koeficientu byly uloženy jako nejlepší dosažené výsledky. Na obrázku 7.2 jsou vyobrazeny dosažené výsledky. V nejlepších prostorových mapách vychází 9 ze 14 pomocí klasické ICA a zbylých 5 pomocí ICA CSV s různou parametrizací. Také je možné si povšimnout, že žádný z nejlepších výsledků nebyl spočítán pomocí algoritmu EFICA. Dále výsledky obsahují pouze některé z inicializačních matic, konkrétně \mathbf{W}_A , \mathbf{W}_C , \mathbf{W}_D a \mathbf{W}_F . Bez ohledu na hodnotu korelačního koeficientu je ve většině nalezených komponent vidět zašumění a nepřesné identifikace aktivních částí. Zašumění je nejlépe vidět na IC 9 a nepřesnosti v určení vzorové komponenty si lze povšimnout na několika výsledcích, konkrétně IC 2, IC 6, IC 12 a IC 14. Zašumění je méně závažný problém než nepřesné nalezení komponenty, protože šum lze odstranit další filtrací. Nenalezenou komponentu však nelze napravit. Obecně je možné říci, že je pro algoritmy problematické nalézt IC 2 až IC 6 a také IC 13 a IC 14. Tyto IC mají oproti ostatním výsledkům korelační koeficient až o čtyři desetiny menší než úspěšněji nalezené IC. Příkladem je zejména IC 4, která má nejmenší korelační koeficient.

Obrázek 7.3a zachycuje výsledky pro všech 14 vzorových komponent při různých inicializacích pro algoritmus FastDIVA s přístupem block-deflation a parametrem $T = 1$ (klasická ICA), který byl nejlepší při nalezení 6 ze 14 hledaných komponent (viz obrázek 7.2). Na této ilustraci je vidět již zmiňovaný problém s komponentami IC 2 až 6, ale také nenalezení IC 8, 9, 12 až 14. Obrázek 7.3b popisuje výsledky algoritmu FastDIVA s block-deflation přístupem a parametrem $T = 27$ (ICA CSV 27 bloků), jež našel nejlépe dvě IC. Jedná se o problematické IC 2 a IC 5, které jsou pomocí klasické ICA (viz obrázek 7.3a) obtížně nalezené. Na obrázku 7.2 si



Obrázek 7.2: Vzorové prostorové mapy a nejlepší nalezené prostorové mapy podle korelačního koeficientu.



Obrázek 7.3: Výsledné korelační koeficienty po spuštění algoritmu FastDIVA vzhledem ke vzorovým komponentám.

lze v obrázcích prvních zleva povšimnout, že IC 2 a 5 jsou nesymetrické podle osy YZ (Coronal view), což by mohl být důvod k obtížnému nalezení.

Tyto výsledky jsou považovány pouze za orientační, protože experiment nebyl robustní. Oba algoritmy byly jednou spuštěny na šesti inicializačních maticích. Tím bylo otestováno nastavení algoritmů a dosažitelnost výsledků. Robustnost je předmětem následujícího experimentu.

7.5 Ověření spolehlivosti nalezení výsledků z druhého experimentu

Tento experiment měl za cíl ověřit spolehlivost nalezení předchozích výsledků. Cílem bylo spuštění 1 000 nezávislých pokusů ze sekce 7.4 a uložení pouze těch nejlepších výsledků dle korelačního koeficientu.

7.5.1 PC a Metacentrum

Experiment 7.4 bylo možné spočítat na PC běžného uživatele. Výsledky byly dosažitelné v rozmezí minut až desítek minut v závislosti na parametrizaci metod. Ověření prvního experimentu ovšem bude trvat 1 000 krát déle, a proto se autor rozhodl využít výpočetních kapacit Metacentra. Celkově se jednalo o 51 způsobů, jak provést 1 000 nezávislých spuštění. Pro EFICA byly zvoleny tři inicializační matice, tedy tři způsoby. FastDIVA s parametry $D1$ a $S1$, tedy klasické ICA, byla provedena pro všech šest inicializačních matic. V neposlední řadě bylo provedeno 36 kombinací pro přístup FastDIVA s CSV. Nejprve s block-deflation přístupem pro všech šest inicializačních matic a pro každou inicializaci byly provedeny tři varianty s parametry $T = 3$, $T = 9$ a $T = 27$. Poté byl stejný postup opakován pro přístup

symmetric. Z těchto spuštění byly ukládány, jen ty nejlepší dosažené výsledky na základě korelačního koeficientu 7.1.2.

Výsledky

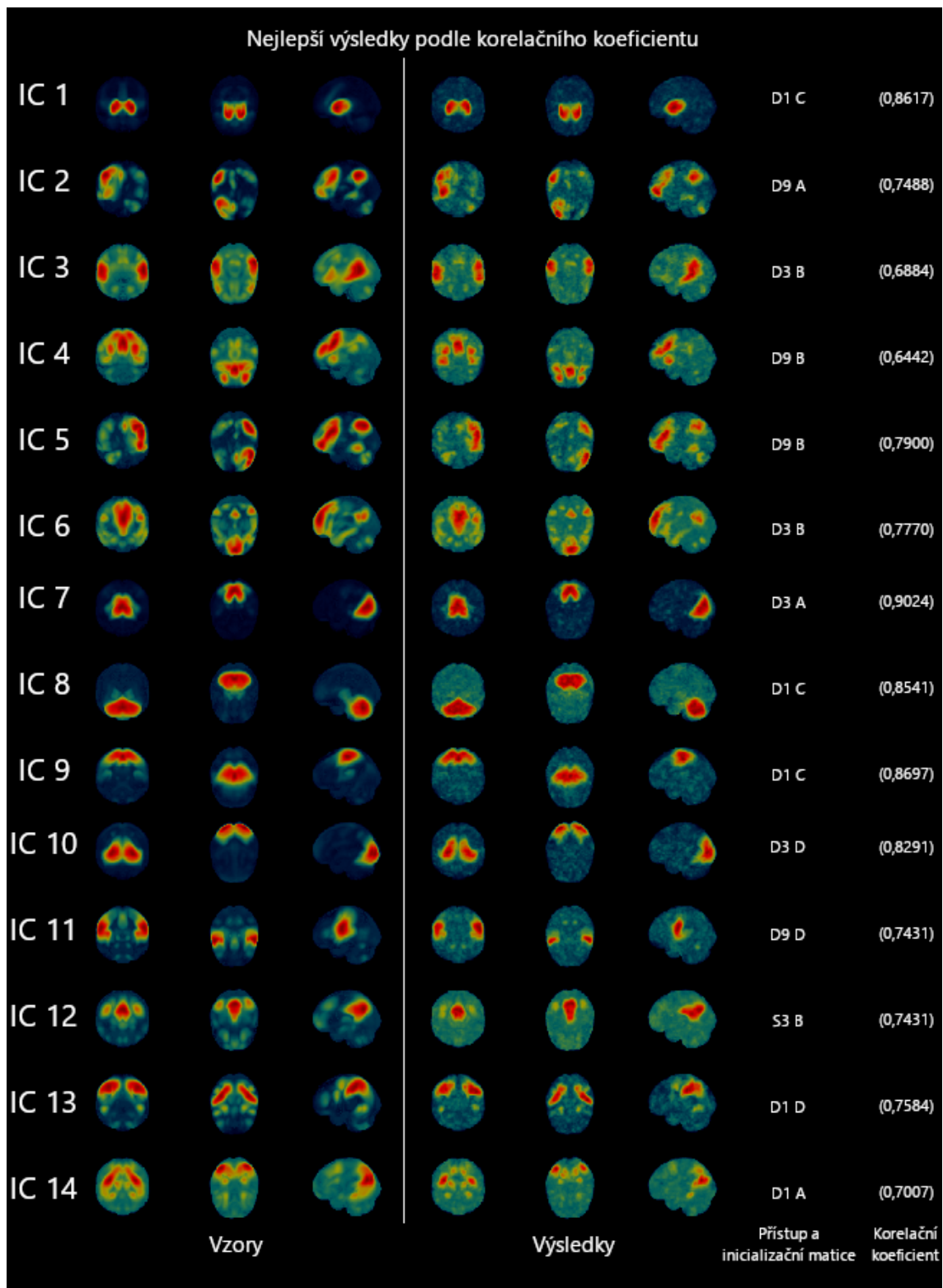
Obrázek 7.4 shrnuje nejlepší dosažené výsledky po 1 000 spuštěních. Hlavním výsledkem byla změna poměru nejlepších výsledků vzhledem k algoritmu. Při prvním pokusu 7.4 bylo 9 ze 14 komponent nalezeno klasickou ICA (výsledky z obrázku 7.2), oproti tomu v tomto pokusu bylo nalezeno 9 ze 14 komponent pomocí ICA CSV (výsledky z obrázku 7.4). U všech IC byl nalezen lepší výsledek než u jednoho spuštění, až na IC 12, která měla korelační koeficient 0,7883 a po 1 000 spuštěních vyšel 0,7431. Tento výsledek lze považovat za náhodu spojenou s inicializací pokusu, což však autora navedlo k přidání další sledované metriky do kontrolního pokusu na datech druhé sady. Tato metrika bude zmíněna dále při kontrolním pokusu. Další zajímavostí tohoto experimentu jsou výsledky v závislosti na inicializačních maticích.

Obrázek 7.5a zobrazuje maximální dosažené hodnoty korelačních koeficientů pro jednotlivé vzorové komponenty. Je zde vidět, že jedna inicializační matice \mathbf{W}_F dokonce zabránila nalezení IC 8. Naopak kompletně náhodná inicializace a vzory zarušené šumem jsou v rámci maximálních výsledků velmi podobné. Dále i po 1 000 spuštěních přetrvávají nižší výsledné hodnoty pro IC 2 až 6. Obrázek 7.5b popisuje nejlepší dosažené výsledky pro jednotlivé vzorové komponenty dle korelačního koeficientu za pomoci ICA CSV. Na první pohled je zde jasný rozdíl v inicializaci pomocí matic \mathbf{W}_E a \mathbf{W}_F . Naopak matice \mathbf{W}_A , \mathbf{W}_B , \mathbf{W}_C a \mathbf{W}_D dosahují srovnatelných výsledků. To znamená, že náhodná inicializační matice nemá podstatně rozdílné výsledky od matice, která obsahuje hledané komponenty s šumem navíc.

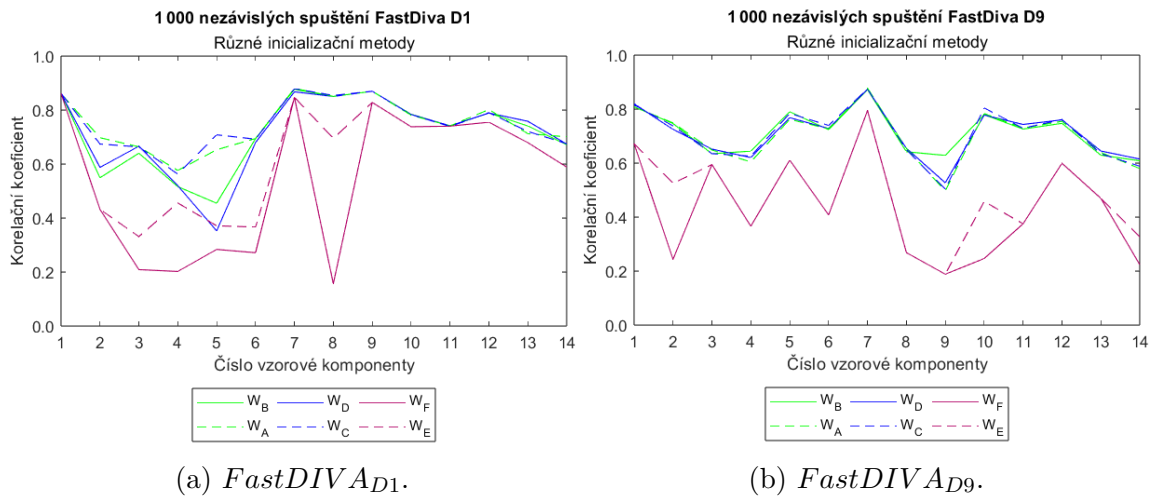
7.6 Ověření experimentu na nezávislé sadě dat

Ověření výsledků předchozího pokusu o 1 000 spuštěních probíhalo znovu na serverech Metacentra. Průběh byl identický až na načtení dat, která byla nahrazena druhou datovou sadou, znovu o 20 subjektech. Zároveň byly u kontrolních dat zaznamenávány četnosti nalezení korelačních koeficientů pro všech 14 vzorových komponent. Jako hranice pro zaznamenání četnosti byly zvoleny nejlepší dosažené výsledky z pokusu 7.4, snižené o dvě desetiny. Snižování hranice bylo provedeno, protože předchozí experiment měl za úkol nalézt nejlepší možné výsledky. Pokud nalezený korelační koeficient dané komponenty přesáhl tuto hodnotu, byl považován za kladný výsledek. Cílem bylo na základě četnosti výsledků potvrdit validitu nalezení daných komponent.

Přidaná informace četnosti přinesla další náhled na výsledná data. Ukázalo se, že EFICA nachází IC stabilněji než FastDIVA. V deseti případech EFICA našla IC, jejichž korelační koeficient přesáhl určenou hranici. Výjimkou jsou IC 2, 8 a 9, kdy IC 2 byla nalezena pouze v pětině spuštění, IC 9 byla nalezena pouze v polovině spuštění a IC 8 nebyla pro určený korelační koeficient nalezena ani v jednom spuštění bez ohledu na inicializační matici. Neznamena to ovšem, že EFICA není schopna



Obrázek 7.4: Nejlepší prostorové mapy po 1 000 nezávislých spuštěních podle korelačního koeficientu

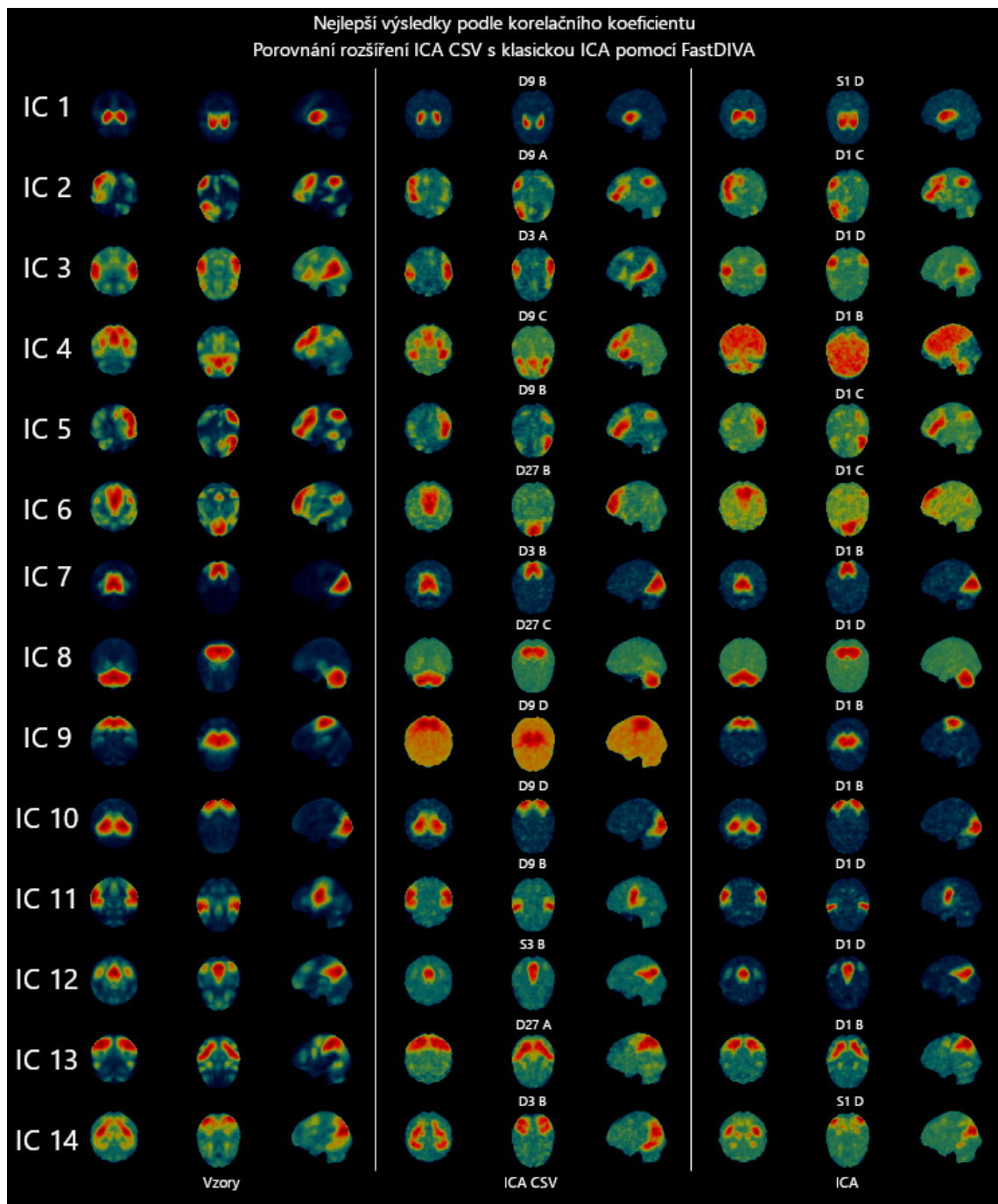


Obrázek 7.5: Maximální hodnoty korelačního koeficientu dosažené po 1 000 nezávislých spuštěních.

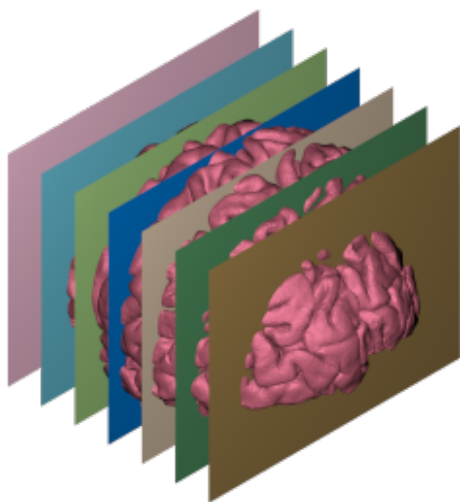
IC 8 nalézt, její výsledek je pouze v porovnání s FastDIVA zřetelně horší. Dále bylo zjištěno, že rozdělení na 27 bloků s přístupem symmetric je chybné. Takto zpracovávaná data bez ohledu na inicializaci nepřinesla tak dobré výsledky, jako při rozdělení na méně bloků. Výjimkou byla pouze IC 4, jejíž nalezení ovšem lze označit jako problematické u všech zkoušených algoritmů a inicializací. U komponent 8 a 9 FastDIVA s $T = 1$ a EFICA nacházejí tyto komponenty alespoň v polovině případů, naopak při rozdělení do více než jednoho bloku se nedaří nalézt IC s takovou přesností, a jejich četnosti jsou pouze v řádu desítek. Zajímavé bylo také porovnání četností mezi inicializacemi \mathbf{W}_A , \mathbf{W}_C , \mathbf{W}_E oproti \mathbf{W}_B , \mathbf{W}_D , \mathbf{W}_F . Autor očekával, že výsledky z inicializací B , D a F budou dosahovat lepších výsledků, protože je prohledáván celý prostor. Po zhodnocení ovšem nelze tvrdit, že inicializace B nebo D nebo F vždy vedou k přesnějším výsledkům. Na závěr data ohledně četnosti nalezení ukazují na častější nalezení IC s vysokým korelačním koeficientem u block-deflation přístupu.

Výsledky

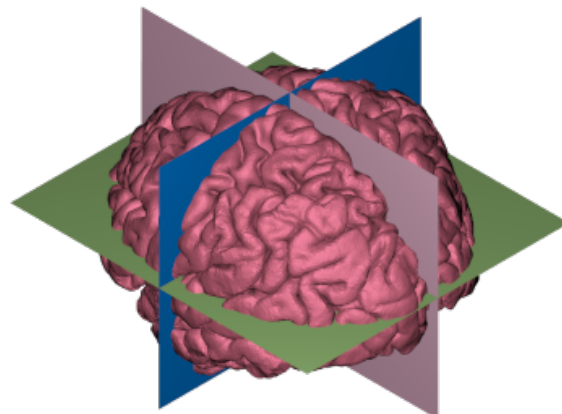
Obrázek 7.6 porovnává nejlepší výsledky ICA CSV s nejlepšími výsledky ICA podle korelačního koeficientu. Náhodná inicializace přinesla většinu nejlepších výsledků v podání ICA CSV. Klasická ICA získala většinu nejlepších výsledků za pomoci inicializační matice s dodatečnou informací, tedy matice \mathbf{W}_C a \mathbf{W}_D . Z předchozího pokusu a obrázku 7.4 autor očekával, že matice \mathbf{W}_E a \mathbf{W}_F nepřinesou žádné lepší výsledky než matice ostatní, což se potvrdilo. Nejrozdílnější výsledky jsou u IC 4 a IC 14. V případě IC 4 nebyla ICA schopna nalézt vzorovou komponentu ani po 1 000 spuštěních, naopak ICA CSV nachází IC 4, ovšem zaměřuje se na části komponenty, které ve vzoru nejsou výrazné a snaží se IC 4 rozdělit do více částí. Pro výsledky u IC 14 klasická ICA nachází komponentu vzdáleně podobnou vzorové, nicméně výsledek z ICA CSV je věrnější vzoru. V některých případech, například



Obrázek 7.6: Vzorové prostorové mapy a nejlepší dosažené výsledky nalezených prostorových map dle korelačního koeficientu. Porovnání výsledků rozšířené metody a metody klasické.



(a) Rozdělení dat bez úpravy na osm regionů.



(b) Upravené rozdělení dat na osm regionů.

Obrázek 7.7: Ilustrace vyplnění bloků metody FastDIVA podle regionů.

IC 11 a IC 12, je vidět šum u metody ICA CSV a naopak u klasické ICA tento šum není. Další výsledek pro IC 9 je v podání ICA CSV na pohled rozdílný od vzorového a klasická ICA je v tomto případě vhodnější volbou.

7.7 Dělení vstupních dat do bloků dle mozkových regionů

Myšlenka rozdělení vstupních dat do regionů vychází z algoritmu 6.3.5, který vstupní data rozdělí na bloky o stejné délce. Toto rozdělení ilustruje obrázek 7.7a. Cílem tohoto experimentu bylo vyplnit bloky daty tak, aby hledané IC byly rozděleny do co nejmenšího počtu bloků, ideálně tak, aby každá IC zasahovala pouze do jednoho bloku. Vybrané rozdělení, jež bylo předmětem experimentu, ilustruje obrázek 7.7b. Rozdělení z obrázku 7.7b bylo vybráno, protože je symetrické a lépe dělí prostor podle anatomie mozkových center. Dále se také zdálo jako vhodné vzhledem ke vzorovým IC, které ve většině případů byly rozděleny lépe než v případě 7.7a.

Pro tento experiment bylo znovu využito všech šest inicializačních matic 7.4.1. Nejprve byl algoritmus FastDIVA spuštěn pouze jednou, stejně jako v 7.4.1. Výsledky po jednom spuštění nepřinesly očekávané zlepšení rozpoznání. Většina komponent nebyla nalezena korektně a v některých případech nebylo možné rozeznat IC od šumu. Tento problém byl viditelný zejména u matic \mathbf{W}_B a \mathbf{W}_D . V případě \mathbf{W}_F byly výsledky silně zarušené šumem. Některé nalezené IC obsahovaly aktivní části, které dle vzorů neměly být aktivní. Zlepšení výsledků přineslo využití matic \mathbf{W}_A a \mathbf{W}_C , které našly IC se znatelně menším šumem a bez přidaných nežádoucích aktivních částí. Na základě toho bylo znovu provedeno 1 000 spuštění algoritmu. Postup byl stejný jako při experimentu v sekci 7.5.

Výsledky

Výsledky lze vidět na obrázku 7.8. Dle porovnání pouze korelačním koeficientem jsou výsledky z obrázku 7.8 velmi podobné výsledkům z obrázku 7.4. Při vizuální kontrole však lze odhalit, že ačkoliv mají výsledky velmi podobný korelační koeficient, jsou rozdílné. Rozdělení dat do bloků podle regionů nepřineslo očekávané zlepšení výsledků, naopak poukázalo na nedostatky porovnávání pomocí korelačního koeficientu. Avšak výhodou tohoto rozdělení bylo nalezení všech hledaných IC, i když se v některých případech jednalo o méně přesné identifikace. Navíc rozdělení pomohlo k nalezení problémové IC 4. Možnost vytvoření rozdělení prostoru na míru hledaným IC je reálnou možností s použitelnými výsledky. Nicméně bez vytvoření podpůrného softwaru který by umožnil přizpůsobení výběru dat do bloků se jedná o časově náročnou přípravu.

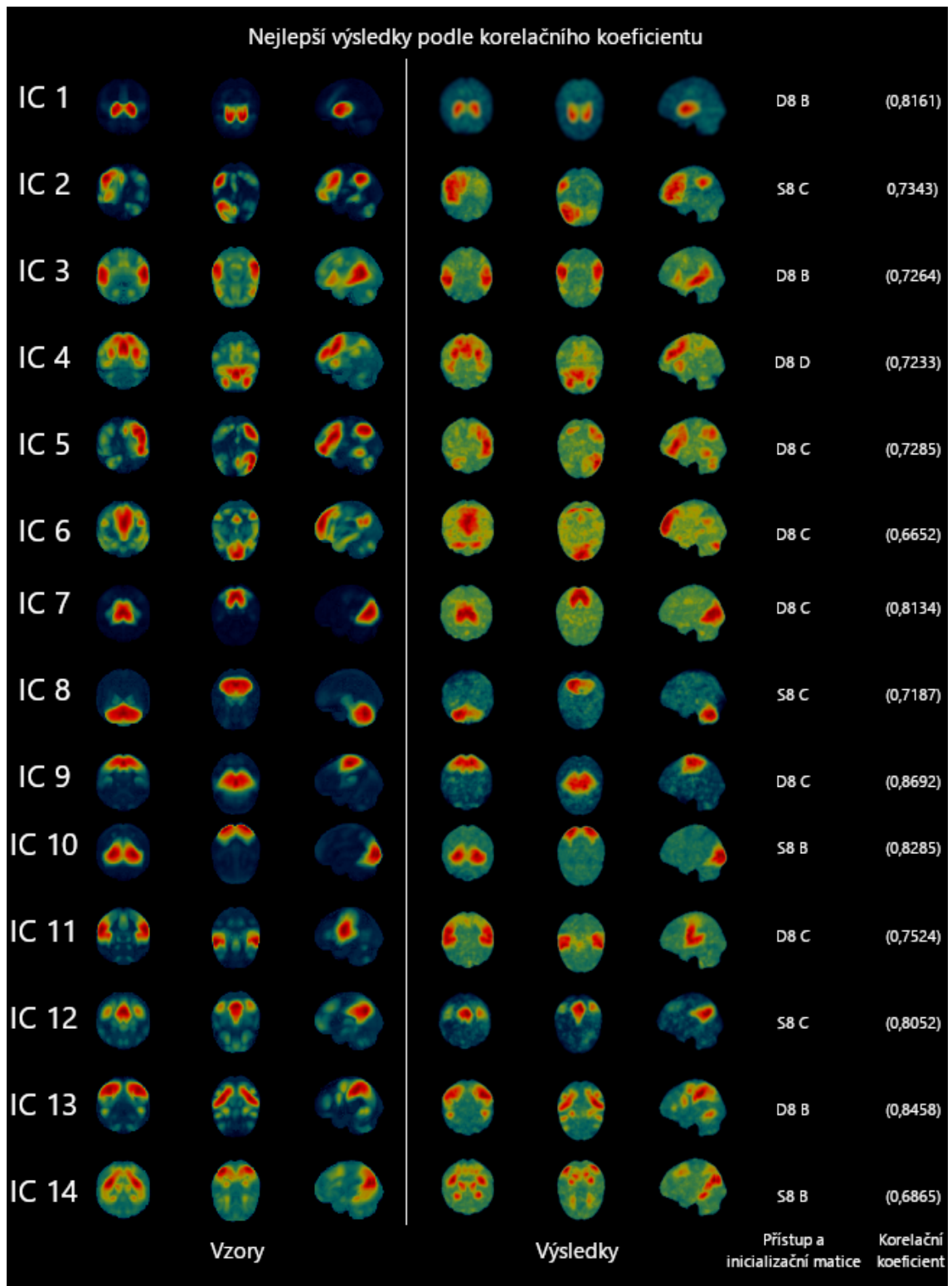
7.8 Výpočet rozšířené metody na běžném PC

Výpočty provedené v experimentech 7.3 a 7.4 byly provedeny na běžném PC. Jednalo se o HP EliteBook 845 G8, který disponoval procesorem Ryzen 3 PRO 5450U od AMD s integrovaným grafickým čipem. Operační paměť dostupná pro výpočet byla 22 GB. Výpočty obou výše zmíněných experimentů probíhaly s využitím jednoho vlákna. Jeden výpočet z experimentu 7.4 trval přibližně dvě až tři minuty. Rozdílné inicializační matice měly na trvání výpočtu vliv pouze v jednotkách sekund. S velikostí inicializační matice roste i čas výpočtu

Rozdílné inicializační matice o stejné velikosti, které obsahovaly různá data měly na trvání výpočtu vliv pouze v jednotkách sekund. Avšak s velikostí inicializační matice úměrně roste čas potřebný k výpočtu. Doba trvání výpočtu roste úměrně i s počtem bloků. Při spuštění dvou identických pokusů, které jsou rozdílné pouze v přístupu metody bylo spuštění s block-deflation přístupem nejen věrnější hledaným vzorům, zároveň bylo průměrně o patnáct sekund kratší.

Spolehlivé výsledky na běžném PC jsou reálně dosažitelné do jednotek hodin. Důvodem je provedení několika na sobě nezávislých výpočtů za účelem nalezení nejlepších možných výsledků. Způsob rozdělení do bloků podle tvaru a umístění hledané IC se podle experimentu 7.7 jeví jako vhodná volba. Správné rozdělení má potenciál separovat hledané IC spolehlivěji, avšak je zde nutné brát ohled na menší přesnost separovaných IC. Rozdělení také přináší další krok v přípravě dat, který v závislosti na složitosti přizpůsobení rozdělení navýší čas provedení výpočtu.

Aby bylo možné metodu využít v praxi na běžném PC, a zároveň její výpočet trval krátkou dobu, bylo by vhodné vytvořit předpřipravená rozdělení do bloků v závislosti na hledaných IC. Pro snížení časové náročnosti výpočtu bude nutné parametry upravovat na základě individuálních vstupních dat a na cílech experimentu. Výstupním krokem by mohly být výsledky získané v experimentech 7.5, 7.6 a 7.7. V těchto experimentech byly dosaženy nejlepší výsledky po 1 000 spuštění, nicméně rychlého a přesto objektivního výpočtu lze dosáhnout již při 100 opakováních.



Obrázek 7.8: Nejlepší výsledky po 1 000 spuštěních s rozdělením do bloků dle regionů podle 7.7b.

8 Závěr

V této diplomové práci bylo exploratorně zkoumáno využití metody ICA CSV na fMRI datech. První část práce vymezuje používané termíny, protože jsou v odborných publikacích často zaměňovány. Poté je představena a popsána teorie potřebná k navrženým experimentům. Autor práce se zaměřil na přesnost nalezení IC a porovnání výsledků klasické ICA metody a rozšiřující metody ICA CSV. Pro výpočet výsledků byly vybrány algoritmy EFICA a FastDIVA.

První z experimentů se zabýval optimální velikostí vstupních dat. Pro datové sady využití v této práci byla nalezena optimální míra redukce dimenze. Druhý experiment zkoumal závislost výsledků na inicializační matici zvoleného algoritmu. V tomto experimentu bylo vytvořeno šest inicializačních matic, které byly mezi sebou porovnány. Výsledné prostorové mapy byly porovnávány s vzorovými prostorovými mapami pomocí Pearsonova korelačního koeficientu.

Na základě šesti inicializačních matic bylo odhaleno, že inicializace maticí obsahující hledané vzory a šum dosahuje srovnatelných výsledků jako inicializace maticí náhodnou. Ve třetím pokusu bylo spuštěno tisíc nezávislých průběhů pokusu druhého a uloženy pouze nejlepší výsledky napříč všemi průběhy. Cílem bylo zajištění robustnosti nalezení výsledku. Nejlepší výsledky při použití ICA CSV byly nalezeny s náhodnou inicializační maticí, zatímco klasická ICA nacházela nejlepší výsledky pomocí inicializace s dodatečnou informací.

Čtvrtý pokus byl proveden pro ověření výsledků předcházejícího pokusu na nezávislé sadě dat. Zároveň byla přidána informace o spolehlivosti nalezení požadovaných IC. Zde byly potvrzeny výsledky třetího pokusu. Důležité je poznamenat, že ICA CSV s vhodně nastavenými parametry vždy nalezla vzorové IC, i když v některých případech s vysokou mírou šumu. Tento problém by však bylo možné řešit přeškálováním nebo další filtrací výsledků. Naopak klasická ICA má větší problém s nalezením některých částí IC, a v případě problémové IC 4 nedokáže vzor nalézt vůbec. Četnosti nalezení výsledků také poukazují na stabilnější nalezení IC pomocí block-deflation přístupu.

Pátý experiment měl za cíl lépe využít možnosti rozdělení signálu do bloků. Výsledky navrženého rozdělení měly velmi podobné hodnoty jako výsledky z experimentu tři a čtyři. Bylo zjištěno, že výsledná data není vhodné porovnávat pouze korelačním koeficientem, jelikož dva vizuálně odlišné výsledky mohou mít velmi podobný korelační koeficient. Nicméně je třeba zmínit, že zvolené rozdělení pomohlo k nalezení problémové IC 4. Při hledání více IC najednou je velmi obtížné rozdělit data do regionů tak, aby vyhovovaly najednou všem IC. Tento přístup si žádá podrobnější studie.

Pro data využitá v této práci nachází ICA CSV jednotlivé IC podle korelačního koeficientu stejně dobře jako klasická ICA, a v některých případech i lépe. U IC, které obsahují více oddělených aktivních částí, byly výsledky z ICA CSV ve většině případů přesnější než ty získané pomocí klasické ICA. Se správným rozdělením do regionů ICA CSV dokázala nalézt i IC, které bylo bez rozdělení problematické nalézt.

V závěru práce byla navržena doporučení pro rychlý výpočet na běžném PC. Je nutné brát ohled na možné chyby v nalezení nebo nenalezení IC při jednotlivých spuštěních. Z četnosti nalezení ze čtvrtého experimentu vyplývá, že pro nalezení nejlepších možných výsledků je nutné spuštění opakovat. Nelze tedy po jednom spuštění prohlásit výsledek za nejlepší možný. Pro zvýšení spolehlivosti nalezení výsledků se autor práce přiklání k využití dělení vstupních dat do bloků podle hledaných IC. Toto doporučení vychází z výsledků experimentu pět v němž bylo zkoumáno rozdělení do bloků. Podle provedených experimentů je vhodné rozdělení do tří až devíti bloků a je preferován block-deflation přístup.

Rozšířená metoda ICA s mixovacím modelem CSV byla v této práci poprvé aplikována na fMRI data. Její výsledky jsou srovnatelné s klasickou ICA a v určitých případech dosahuje rozšíření lepších výsledků. V této práci nezbyl prostor pro zkoumání vlastnosti metody, kterou je separace pohybujících se zdrojů. Tato vlastnost by mohla lépe identifikovat postup a vývoj mozkové aktivity. Pro tuto analýzu je základem přesná identifikace IC, kterou ICA CSV v této práci prokázala.

Použitá literatura

- [1] Tulay Adali, Matthew Anderson a Geng-Shen Fu. „Diversity in Independent Component and Vector Analyses: Identifiability, algorithms, and applications in medical imaging“. In: *IEEE Signal Processing Magazine* 31.3 (květ. 2014), s. 18–33. ISSN: 1053-5888. DOI: [10.1109/MSP.2014.2300511](https://doi.org/10.1109/MSP.2014.2300511). URL: <http://ieeexplore.ieee.org/document/6784026/> (cit. 29. 06. 2023).
- [2] N. Amor et al. „Blind Extraction of Moving Sources via Independent Component and Vector Analysis: Examples“. In: *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). ISSN: 2379-190X. Čvn. 2021, s. 3725–3729. DOI: [10.1109/ICASSP39728.2021.9413422](https://doi.org/10.1109/ICASSP39728.2021.9413422).
- [3] Anthony J. Bell a Terrence J. Sejnowski. „An Information-Maximization Approach to Blind Separation and Blind Deconvolution“. In: *Neural Computation* 7.6 (lis. 1995), s. 1129–1159. ISSN: 0899-7667, 1530-888X. DOI: [10.1162/neco.1995.7.6.1129](https://doi.org/10.1162/neco.1995.7.6.1129). URL: <https://direct.mit.edu/neco/article/7/6/1129-1159/5909> (cit. 10. 12. 2023).
- [4] E. Bingham a A. Hyvarinen. „ICA of complex valued signals: a fast and robust deflationary algorithm“. In: *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium. Sv. 3. ISSN: 1098-7576. Čvc. 2000, 357–362 vol.3. DOI: [10.1109/IJCNN.2000.861330](https://doi.org/10.1109/IJCNN.2000.861330). URL: <https://ieeexplore.ieee.org/abstract/document/861330> (cit. 05. 11. 2023).
- [5] Randy L. Buckner, Jessica R. Andrews-Hanna a Daniel L. Schacter. „The Brain’s Default Network“. In: *Annals of the New York Academy of Sciences* 1124.1 (2008). _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1196/annals.1440.011>, s. 1–38. ISSN: 1749-6632. DOI: [10.1196/annals.1440.011](https://doi.org/10.1196/annals.1440.011). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1196/annals.1440.011> (cit. 09. 07. 2023).

- [6] Richard B. Buxton. *Introduction to functional magnetic resonance imaging: principles and techniques*. 2nd ed. OCLC: 476252577. Cambridge: Cambridge University Press, 2009. ISBN: 978-0-511-60520-8.
- [7] Vince D. Calhoun a Tülay Adalı. „Multisubject Independent Component Analysis of fMRI: A Decade of Intrinsic Networks, Default Mode, and Neurodiagnostic Discovery“. In: *IEEE Reviews in Biomedical Engineering* 5 (2012). Conference Name: IEEE Reviews in Biomedical Engineering, s. 60–73. ISSN: 1941-1189. DOI: [10.1109/RBME.2012.2211076](https://doi.org/10.1109/RBME.2012.2211076).
- [8] Vince D. Calhoun, Jingyu Liu a Tülay Adalı. „A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data“. In: *NeuroImage*. Mathematics in Brain Imaging 45.1 (1. břez. 2009), S163–S172. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2008.10.057](https://doi.org/10.1016/j.neuroimage.2008.10.057). URL: <https://www.sciencedirect.com/science/article/pii/S1053811908012032> (cit. 29. 06. 2023).
- [9] N. Correa et al. „Comparison of blind source separation algorithms for FMRI using a new Matlab toolbox: GIFT“. In: *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005*. Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Sv. 5. ISSN: 2379-190X. Břez. 2005, v/401–v/404 Vol. 5. DOI: [10.1109/ICASSP.2005.1416325](https://doi.org/10.1109/ICASSP.2005.1416325).
- [10] Nicolle Correa, Tülay Adalı a Vince D. Calhoun. „Performance of blind source separation algorithms for fMRI analysis using a group ICA method“. In: *Magnetic Resonance Imaging* 25.5 (čvn. 2007), s. 684–694. ISSN: 0730725X. DOI: [10.1016/j.mri.2006.10.017](https://doi.org/10.1016/j.mri.2006.10.017). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0730725X06003080> (cit. 09. 07. 2023).
- [11] T. M. Cover a Joy A. Thomas. *Elements of information theory*. 2nd ed. OCLC: ocm59879802. Hoboken, N.J: Wiley-Interscience, 2006. 748 s. ISBN: 978-0-471-24195-9.
- [12] J.S. Damoiseaux et al. „Consistent resting-state networks across healthy subjects“. In: *Proceedings of the National Academy of Sciences of the United States of America* 103.37 (2006), s. 13848–13853. ISSN: 0027-8424. DOI: [10.1073/pnas.0601417103](https://doi.org/10.1073/pnas.0601417103).
- [13] Yuhui Du et al. „Artifact removal in the context of group ICA: A comparison of single-subject and group approaches“. In: *Human Brain Mapping* 37.3 (21. pros. 2015), s. 1005–1025. ISSN: 1065-9471. DOI: [10.1002/hbm.23086](https://doi.org/10.1002/hbm.23086). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5784424/> (cit. 29. 06. 2023).
- [14] Yuhui Du et al. „NeuroMark: An automated and adaptive ICA based pipeline to identify reproducible fMRI markers of brain disorders“. In: *NeuroImage: Clinical* 28 (2020), s. 102375. ISSN: 22131582. DOI: [10.1016/j.nicl.2020.102375](https://doi.org/10.1016/j.nicl.2020.102375). URL: <https://linkinghub.elsevier.com/retrieve/pii/S2213158220302126> (cit. 05. 06. 2023).

- [15] Nguyen Thanh Duc et al. „3D-Deep Learning Based Automatic Diagnosis of Alzheimer’s Disease with Joint MMSE Prediction Using Resting-State fMRI“. In: *Neuroinformatics* 18.1 (1. led. 2020), s. 71–86. ISSN: 1559-0089. DOI: [10.1007/s12021-019-09419-w](https://doi.org/10.1007/s12021-019-09419-w). URL: <https://doi.org/10.1007/s12021-019-09419-w> (cit. 10. 07. 2023).
- [16] M.D. Greicius et al. „Functional connectivity in the resting brain: A network analysis of the default mode hypothesis“. In: *Proceedings of the National Academy of Sciences of the United States of America* 100.1 (2003), s. 253–258. ISSN: 0027-8424. DOI: [10.1073/pnas.0135058100](https://doi.org/10.1073/pnas.0135058100).
- [17] Martijn P. van den Heuvel a Hilleke E. Hulshoff Pol. „Exploring the brain network: A review on resting-state fMRI functional connectivity“. In: *European Neuropsychopharmacology* 20.8 (1. srp. 2010), s. 519–534. ISSN: 0924-977X. DOI: [10.1016/j.euroneuro.2010.03.008](https://doi.org/10.1016/j.euroneuro.2010.03.008). URL: <https://www.sciencedirect.com/science/article/pii/S0924977X10000684> (cit. 09. 07. 2023).
- [18] J. Himberg a A. Hyvarinen. „Icasso: software for investigating the reliability of ICA estimates by clustering and visualization“. In: *2003 IEEE XIII Workshop on Neural Networks for Signal Processing (IEEE Cat. No.03TH8718)*. 2003 IEEE XIII Workshop on Neural Networks for Signal Processing (IEEE Cat. No.03TH8718). ISSN: 1089-3555. Zář. 2003, s. 259–268. DOI: [10.1109/NNSP.2003.1318025](https://doi.org/10.1109/NNSP.2003.1318025). URL: <https://ieeexplore.ieee.org/document/1318025?denied=> (cit. 10. 12. 2023).
- [19] A. Hyvärinen a E. Oja. „Independent component analysis: algorithms and applications“. In: *Neural Networks* 13.4 (1. čvn. 2000), s. 411–430. ISSN: 0893-6080. DOI: [10.1016/S0893-6080\(00\)00026-5](https://doi.org/10.1016/S0893-6080(00)00026-5). URL: <https://www.sciencedirect.com/science/article/pii/S0893608000000265> (cit. 27. 03. 2023).
- [20] Aapo Hyvärinen, Juha Karhunen a Erkki Oja. *Independent Component Analysis*. Google-Books-ID: 96D0ypDwAkkC. John Wiley & Sons, 5. dub. 2004. 506 s. ISBN: 978-0-471-46419-8.
- [21] Mark Jenkinson et al. „FSL“. In: *NeuroImage*. 20 YEARS OF fMRI 62.2 (15. srp. 2012), s. 782–790. ISSN: 1053-8119. DOI: [10.1016/j.neuroimage.2011.09.015](https://doi.org/10.1016/j.neuroimage.2011.09.015). URL: <https://www.sciencedirect.com/science/article/pii/S1053811911010603> (cit. 10. 07. 2023).
- [22] Václav Kautský et al. „Cramér–Rao Bounds for Complex-Valued Independent Component Extraction: Determined and Piecewise Determined Mixing Models“. In: *IEEE Transactions on Signal Processing* 68 (2020). Conference Name: IEEE Transactions on Signal Processing, s. 5230–5243. ISSN: 1941-0476. DOI: [10.1109/TSP.2020.3022827](https://doi.org/10.1109/TSP.2020.3022827).
- [23] David B. Keator et al. „The Function Biomedical Informatics Research Network Data Repository“. In: *NeuroImage* 124 (Pt B 1. led. 2016), s. 1074–1079. ISSN: 1095-9572. DOI: [10.1016/j.neuroimage.2015.09.003](https://doi.org/10.1016/j.neuroimage.2015.09.003).

- [24] Z. Koldovsky, P. Tichavsky a E. Oja. „Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound“. In: *IEEE Transactions on Neural Networks* 17.5 (zář. 2006). Conference Name: IEEE Transactions on Neural Networks, s. 1265–1277. ISSN: 1941-0093. DOI: [10.1109/TNN.2006.875991](https://doi.org/10.1109/TNN.2006.875991).
- [25] Zbyněk Koldovský a Petr Tichavský. „Gradient Algorithms for Complex Non-Gaussian Independent Component/Vector Extraction, Question of Convergence“. In: *IEEE Transactions on Signal Processing* 67.4 (ún. 2019). Conference Name: IEEE Transactions on Signal Processing, s. 1050–1064. ISSN: 1941-0476. DOI: [10.1109/TSP.2018.2887185](https://doi.org/10.1109/TSP.2018.2887185).
- [26] Zbyněk Koldovský et al. „Dynamic Independent Component/Vector Analysis: Time-Variant Linear Mixtures Separable by Time-Invariant Beamformers“. In: *IEEE Transactions on Signal Processing* 69 (2021). Conference Name: IEEE Transactions on Signal Processing, s. 2158–2173. ISSN: 1941-0476. DOI: [10.1109/TSP.2021.3068626](https://doi.org/10.1109/TSP.2021.3068626).
- [27] Dana Lahat, Tulay Adali a Christian Jutten. „Multimodal Data Fusion: An Overview of Methods, Challenges, and Prospects“. In: *Proceedings of the IEEE* 103.9 (zář. 2015), s. 1449–1477. ISSN: 0018-9219, 1558-2256. DOI: [10.1109/JPROC.2015.2460697](https://doi.org/10.1109/JPROC.2015.2460697). URL: <http://ieeexplore.ieee.org/document/7214350/> (cit. 03. 04. 2023).
- [28] Martin J McKeown, Lars Kai Hansen a Terrence J Sejnowsk. „Independent component analysis of functional MRI: what is signal and what is noise?“. In: *Current Opinion in Neurobiology* 13.5 (1. řij. 2003), s. 620–629. ISSN: 0959-4388. DOI: [10.1016/j.conb.2003.09.012](https://doi.org/10.1016/j.conb.2003.09.012). URL: <https://www.sciencedirect.com/science/article/pii/S0959438803001338> (cit. 27. 03. 2023).
- [29] Martin J. Mckeown et al. „Analysis of fMRI data by blind separation into independent spatial components“. In: *Human Brain Mapping* 6.3 (1998), s. 160–188. ISSN: 1097-0193. DOI: [10.1002/\(SICI\)1097-0193\(1998\)6:3<160::AID-HBM5>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1097-0193(1998)6:3<160::AID-HBM5>3.0.CO;2-1). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-0193%281998%296%3A3%3C160%3A%3AAID-HBM5%3E3.0.CO%3B2-1> (cit. 27. 03. 2023).
- [30] S.M. Plis et al. „Deep learning for neuroimaging: A validation study“. In: *Frontiers in Neuroscience* 8 (2014). ISSN: 1662-4548. DOI: [10.3389/fnins.2014.00229](https://doi.org/10.3389/fnins.2014.00229).
- [31] Muhammad Naveed Iqbal Qureshi, Jooyoung Oh a Boreom Lee. „3D-CNN based discrimination of schizophrenia using resting-state fMRI“. In: *Artificial Intelligence in Medicine* 98 (1. čvc. 2019), s. 10–17. ISSN: 0933-3657. DOI: [10.1016/j.artmed.2019.06.003](https://doi.org/10.1016/j.artmed.2019.06.003). URL: <https://www.sciencedirect.com/science/article/pii/S0933365719301393> (cit. 10. 07. 2023).
- [32] M.M. Rahman et al. „Interpreting models interpreting brain dynamics“. In: *Scientific Reports* 12.1 (2022). ISSN: 2045-2322. DOI: [10.1038/s41598-022-15539-2](https://doi.org/10.1038/s41598-022-15539-2).

- [33] M. E. Raichle et al. „A default mode of brain function“. In: *Proceedings of the National Academy of Sciences of the United States of America* 98.2 (16. led. 2001), s. 676–682. ISSN: 0027-8424. DOI: [10.1073/pnas.98.2.676](https://doi.org/10.1073/pnas.98.2.676).
- [34] Rangaraj Mandayam Rangayyan. *Biomedical signal analysis: a case-study approach*. IEEE press series on biomedical engineering. Piscataway, N.J. New York: IEEE press Wiley-interscience, 2002. ISBN: 978-0-471-20811-2.
- [35] Stephen M. Smith et al. „Correspondence of the brain’s functional architecture during activation and rest“. In: *Proceedings of the National Academy of Sciences of the United States of America* 106.31 (4. srp. 2009), s. 13040–13045. ISSN: 0027-8424. DOI: [10.1073/pnas.0905267106](https://doi.org/10.1073/pnas.0905267106). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2722273/> (cit. 04. 09. 2023).
- [36] Radek Tupý. „Functional magnetic resonance in relation to languages centers“. In: *Listy klinické logopedie* 2.2 (30. zář. 2018), s. 22–26. ISSN: 25706179, 25706179. DOI: [10.36833/lkl.2018.021](https://doi.org/10.36833/lkl.2018.021). URL: <http://casopis.aklcr.cz/doi/10.36833/lkl.2018.021.html> (cit. 13. 12. 2023).
- [37] J. Xu, M.N. Potenza a V.D. Calhoun. „Spatial ICA reveals functional activity hidden from traditional fMRI GLM-based analyses“. In: *Frontiers in Neuroscience* 7 (2013). ISSN: 1662-4548. DOI: [10.3389/fnins.2013.00154](https://doi.org/10.3389/fnins.2013.00154).