

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2024

Bc. Veronika Plocková



# VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY

A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

## POKROČILÁ ANALÝZA CHARAKTERISTICKÝCH PŘÍZNAKŮ TROMBŮ V MULTIFÁZICKÝCH CT DATECH

(PLOCKOVÁ VERONIKA) ADVANCED ANALYSIS OF SPECIFIC THROMBI FEATURES IN MULTIPHASE CT DATA

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Veronika Plocková

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Roman Jakubíček, Ph.D.

BRNO 2024

# Diplomová práce

magisterský navazující studijní program **Bioinženýrství**

Ústav biomedicínského inženýrství

**Studentka:** Bc. Veronika Plocková

**ID:** 211575

**Ročník:** 2

**Akademický rok:** 2023/24

**NÁZEV TÉMATU:**

## **Pokročilá analýza charakteristických příznaků trombů v multifázických CT datech**

**POKYNY PRO VYPRACOVÁNÍ:**

1) Nastudujte problematiku ischemické mozkové příhody, její diagnostiky pomocí CT a průběhu léčby. 2) Vyhledejte v odborné literatuře relevantní informace a vypracujte rešerši z oblasti analýzy trombu v CT datech a jejich vztahy k efektivitě léčby. 3) Seznamte se a zpracujte multifázická CT data poskytnutá vedoucím, a ve vhodném programovém prostředí extrahujte radiomické příznaky. 4) Navrhněte a realizujte alespoň jeden praktický experiment pro analýzu radiomických obrazových příznaků trombu včetně metodiky hodnocení experimentu. 5) Navrhněte a realizujte další alespoň dva experimenty vedoucí k redukci a selekci příznakového prostoru vzhledem k potřebám klinické praxe. 6) Vyhodnoťte a diskutujte výsledky dosažené v experimentech a vyvoďte závěry z provedených experimentů.

**DOPORUČENÁ LITERATURA:**

[1] THEODORIDIS, Sergios a KOUTROUMBAS, Konstantinos, c2009. Pattern recognition. 4th ed. Burlington, Mass.: Academic Press. ISBN 9781597492720.

[2] DUMITRIU LAGRANGE, Daniela; HOFMEISTER, Jeremy; ROSI, Andrea; VARGAS, Maria Isabel; WANKE, Isabel et al., 2023. Predictive value of clot imaging in acute ischemic stroke: A systematic review of artificial intelligence and conventional studies. Online. Neuroscience Informatics. Roč. 3, č. 1. ISSN 27725286. Dostupné z: <https://doi.org/10.1016/j.neuri.2022.100114>. [cit. 2023-08-14].

**Termín zadání:** 5.2.2024

**Termín odevzdání:** 22.5.2024

**Vedoucí práce:** Ing. Roman Jakubíček, Ph.D.

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

**UPOZORNĚNÍ:**

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

## **ABSTRAKT**

Tato diplomová práce se zabývá pokročilou analýzou trombu v multifázických CT datech. V první části je zpracována rešerše na téma cévní mozkové příhody, její léčby a vlivu struktury trombu na výslednou úspěšnost léčby. Dále se v rešerši objevuje stručná anatomie mozku a metody extrakce radiomických příznaků a následné možné metody analýzy trombu z obrazových dat. V poslední části teoretické práce jsou shrnuty extrahované příznaky použité v praktické části práce a metody jejich výpočtu. Následuje stručné představení statistických metod použitých také v praktické části práce. Praktická část se zaměřuje na extrakci statistických a texturních příznaků z CT dat. Experimenty se zabývají analýzou extrahovaných příznaků, diskriminačními vlastnostmi trombu a pozadí a navrhováním přístupu k segmentaci trombu z CT dat. Na závěr jsou výsledky experimentů diskutovány.

## **KLÍČOVÁ SLOVA**

Cévní mozková příhoda, trombus, CT angiografie, extrakce texturních příznaků, analýza příznaků, radiomika, segmentace.

## **ABSTRACT**

This diploma thesis focuses on advanced analysis of thrombus in multiphase CT data. The first part presents a literature review on ischemic stroke, its treatment, and the impact of thrombus structure on treatment success. Additionally, the literature review includes brief anatomy of the brain and methods for extracting radiomic features and possible methods for thrombus analysis from imaging data. In the last part of the theoretical work, the extracted features used in the practical part of the work and methods of their calculation are summarized. A brief introduction to statistical methods used in the practical part of the work follows. The practical part focuses on extracting statistical and textural features from CT data. Experiments deal with the analysis of extracted features, discriminatory properties of thrombus and background, and proposing an approach to thrombus segmentation from CT data. The results of the experiments are discussed.

## **KEYWORDS**

Stroke, thrombus, CT angiography, texture extraction, feature analysis, radiomics, segmentation.

PLOCKOVÁ, Veronika. *Pokročilá analýza charakteristických příznaků trombů v multi-fazických CT datech*. Diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství, 2024. Vedoucí práce: Ing. Roman Jakubíček, Ph.D

## Prohlášení autora o původnosti díla

**Jméno a příjmení autora:** Bc. Veronika Plocková  
**VUT ID autora:** 211575  
**Typ práce:** Diplomová práce  
**Akademický rok:** 2023/24  
**Téma závěrečné práce:** Pokročilá analýza charakteristických příznaků trombů v multifazických CT datech

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

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

Brno .....

.....

podpis autorky\*

---

\*Autor podepisuje pouze v tištěné verzi.

## PODĚKOVÁNÍ

Ráda bych poděkovala vedoucímu diplomové práce panu Ing. Romanu Jakubíčkoví, Ph.D. za trpělivost, ochotu, cenné rady a podnětné připomínky k práci.

# Obsah

Úvod	12
<b>1 Anatomie mozku a cévní mozková příhoda</b>	<b>13</b>
1.1 Cévní zásobení mozku	14
1.2 Cévní mozková příhoda	16
1.3 Diagnostika cévní mozkové příhody	16
1.4 Zobrazovací techniky v diagnostice CMP	17
1.5 Léčba cévní mozkové příhody	18
1.6 Charakteristika trombů a jejich patofyziologie	19
<b>2 Analýza trombu v CT datech</b>	<b>21</b>
2.1 Předzpracování	21
2.2 Analýza pomocí konvenčních statistických metod	21
2.3 Analýza pomocí strojového učení	22
<b>3 Metody zpracování dat</b>	<b>24</b>
3.1 Binární morfologické operace	24
3.2 Extrakce příznaků	25
3.2.1 Statistické metody prvního řádu	25
3.2.2 Co-occurrence matice (GLCM)	29
3.2.3 Gray Level Size Zone Matrix (GLSZM)	29
3.2.4 Gray Level Run Length Matrix (GLRLM)	29
3.2.5 Neighbouring Gray Tone Difference Matrix (NGTDM)	30
3.2.6 Gray Level Dependence Matrix (GLDM)	30
3.3 Statistická analýza	30
3.3.1 Pearsonův korelační koeficient	30
3.3.2 Mann-Whitney test	31
<b>4 Experimentální část</b>	<b>32</b>
4.1 Použitá data	32
4.1.1 Předzpracování dat a extrakce příznaků	33
4.2 Experiment 1: Korelační a PCA analýza radiomických příznaků voxelů trombu a pozadí	34
4.2.1 Návrh experimentu	34
4.2.2 Výsledky experimentu	36
4.2.3 Diskuze a závěr	38
4.3 Experiment 2: Diskriminace voxelů trombu a pozadí v nativních a multifázických snímcích	39

4.3.1	Použitá data a jejich předzpracování . . . . .	39
4.3.2	Návrh experimentu . . . . .	39
4.3.3	Výsledky experimentu . . . . .	40
4.3.4	Diskuze a závěr . . . . .	46
4.4	Experiment 3: Segmentace trombů pomocí metod strojového učení . .	46
4.4.1	Data a návrh experimentu . . . . .	46
4.4.2	Výsledky a diskuze . . . . .	49
4.4.3	Diskuze a závěr . . . . .	53
<b>Závěr</b>		<b>55</b>
<b>Literatura</b>		<b>57</b>
<b>Seznam symbolů a zkratk</b>		<b>63</b>
<b>Seznam příloh</b>		<b>64</b>
<b>A Přílohy</b>		<b>65</b>
A.1	Seznam příznaků . . . . .	65
A.2	t-SNE 10 vybraných příznaků . . . . .	68
A.3	t-SNE 15 vybraných příznaků . . . . .	71
<b>B Obsah elektronické přílohy</b>		<b>73</b>

# Seznam obrázků

1.1	Anatomie mozku . . . . .	14
1.2	Cévní systém mozku . . . . .	15
1.3	Zobrazení trombu . . . . .	17
3.1	Příklady morfologických operací. . . . .	25
4.1	Ukázka dostupných obrazových dat . . . . .	33
4.2	Předzpracovaná data . . . . .	34
4.3	Blokové schéma 1. experimentu . . . . .	35
4.4	Korelační matice - detail . . . . .	36
4.5	Korelační matice . . . . .	37
4.6	Scree plot . . . . .	38
4.7	PCA analýza hlavních komponent . . . . .	38
4.8	Blokové schéma 2. experimentu . . . . .	39
4.9	Histogramy příznaků . . . . .	41
4.10	t-SNE vizualizace vybraných příznaků . . . . .	43
4.11	t-SNE vizualizace vybraných příznaků pro vyrovnané počty dat . . .	43
4.12	t-SNE vizualizace vybraných příznaků pro vyrovnané počty dat . . .	44
4.13	Histogram příznaku - fúzovaný snímek . . . . .	44
4.14	t-SNE vizualizace vybraných příznaků pro vyrovnané počty dat . . .	45
4.15	Blokové schéma 3. experimentu . . . . .	47
4.16	Elbow metoda . . . . .	49
4.17	Segmentace K-means . . . . .	50
4.18	Segmentace Random forrest . . . . .	51
4.19	Segmentace Random forrest . . . . .	52
4.20	Segmentace Gradient boosting . . . . .	53
A.1	t-SNE 10 příznaků (pacient 112) . . . . .	68
A.2	t-SNE 10 příznaků (pacient 238) . . . . .	68
A.3	t-SNE 10 příznaků (pacient 243) . . . . .	68
A.4	t-SNE 10 příznaků (pacient 297) . . . . .	69
A.5	t-SNE 10 příznaků (pacient 319) . . . . .	69
A.6	t-SNE 10 příznaků (pacient 365) . . . . .	69
A.7	t-SNE 10 příznaků (pacient 394) . . . . .	70
A.8	t-SNE 10 příznaků (pacient 400) . . . . .	70
A.9	t-SNE 15 příznaků (pacient 112 a 238) . . . . .	71
A.10	t-SNE 15 příznaků (pacient 243 a 297) . . . . .	71
A.11	t-SNE 15 příznaků (pacient 319 a 365) . . . . .	71
A.12	t-SNE 15 příznaků (pacient 394 a 400) . . . . .	72

# Seznam tabulek

4.1	Tabulka s hodnotami DICE skóre pro jednotlivé pacienty. . . . .	42
4.2	Hodnocení úspěšnosti metody "K-means" a průměrné DICE skóre. . .	50
4.3	Hodnocení úspěšnosti metody "Random forrest" a průměrné DICE skóre.	50
4.4	Hodnocení úspěšnosti metody "Random forrest" a průměrné DICE skóre. Zvýšená trénovací množina 8 trénovací a 5 test . . . . .	51
4.5	Hodnocení úspěšnosti metody "Gradient Boosting" a průměrné DICE skóre. Menší trénovací množina . . . . .	53
A.1	Extrahované příznaky (část 1). . . . .	65
A.2	Extrahované příznaky (část 2). . . . .	66
A.3	Extrahované příznaky (část 3). . . . .	67

# Úvod

V současné době představuje cévní mozková příhoda (CMP) jeden z nejčastějších a velmi závažných zdravotních problémů s výraznými dopady na kvalitu života postižených jednotlivců. CMP je charakterizována náhlým narušením krevního zásobení mozku, což může mít vážné důsledky, včetně neurologických deficitů a v nejhorším případě trvalých postižení. Tato problematika se stává stále naléhavější, přičemž rychlá a přesná diagnostika hraje klíčovou roli při minimalizaci škodlivých efektů CMP.

Vědecký pokrok v oblasti obrazových diagnostických technologií, jako jsou výpočetní tomografie (CT) a magnetická rezonance (MRI), umožňuje detailnější studium cévní mozkové příhody a poskytuje cenné informace pro léčebné strategie. Tato diplomová práce se zaměřuje na současný výzkum a technologické inovace v oblasti diagnostiky CMP, s důrazem na analýzu radiomických příznaků trombu v multi-fázických CT datech. V teoretické části byla zpracovaná rešerše popisující CMP z medicínského hlediska a možnostmi její léčby. V další části teoretické rešerše byly představeny postupy analýzy obrazů. Praktická část práce je zaměřena na extrakci radiomických příznaků pomocí statistických metod prvního řádu, ale i pomocí složitějších metod texturní analýzy. Další část práce byla rozdělena na tři experimenty. Prvním experimentem byla analýza již extrahovaných příznaků. V dalším experimentu byla zkoumána diskriminační vlastnost trombu a pozadí pomocí statistických testů, jednoduché segmentace a analýzy t-SNE, kde byly porovnány výsledky pro nativní snímky a zfúzované snímky. Posledním experimentem bylo navrhnout přístupy segmentace trombu z CT dat. Na závěr byly všechny přístupy a jejich výsledky diskutovány.

# 1 Anatomie mozku a cévní mozková příhoda

Mozek, jakožto jeden z primárních pilířů lidského organismu, představuje vysoce komplexní orgán. Zajišťuje řízení myšlenek, uchovávání paměti, prožívání emocí, vnímání dotyků, pohybové schopnosti, vidění, dýchání, regulaci teploty, hladiny energie a veškerého řízení tělesných funkcí. Společně s míchou, která z něj vychází, tvoří centrální nervový systém, jinak zvaný CNS [1].

Anatomicky se mozek dá rozdělit na několik základních částí:

## Koncový mozek (*telencephalon*)

Koncový mozek je tvořen dvěma hemisférami, pravou a levou, které jsou spojeny tkáněmi nazývanými *corpus callosum*. Každá hemisféra má pět telencephalických laloků – čelní, týlní, spánkový, temenní a insulární, ty hrají klíčovou roli v různých aspektech kognitivních funkcí [1].

## Mozeček (*cerebellum*)

*Cerebellum* představuje největší segment zadní části mozku a je spojen s různými oblastmi mozku třemi páry mozečkových stonků. Získává informace z motorických i senzorických drah, což umožňuje synchronizaci aktuálního stavu pohybového aparátu s plánovanými pohyby. Zajímavý je fakt, že počet vstupních drah je až čtyřicetkrát větší, než počet výstupních drah, což znamená, že do mozečku přichází spousta informací z různých částí těla a díky tomu mozek provádí přesné pohyby. Dalo by se to přirovnat k centrální řídicí jednotce, která dostává spousta informací a na základě nich řídí, jak má tělo reagovat a jaké pohyby má vykonat [2].

*Cerebellum* je také důležitý pro udržování rovnováhy, správný svalový tonus, vertikální postoj a harmonickou koordinaci pohybů. Jeho činnost je aktivní i při zpracování informací o plánovaných pohybech, a má schopnost modifikovat jak pyramidové, tak mimopyramidové motorické impulsy, které následně směřují eferentními dráhami [2].

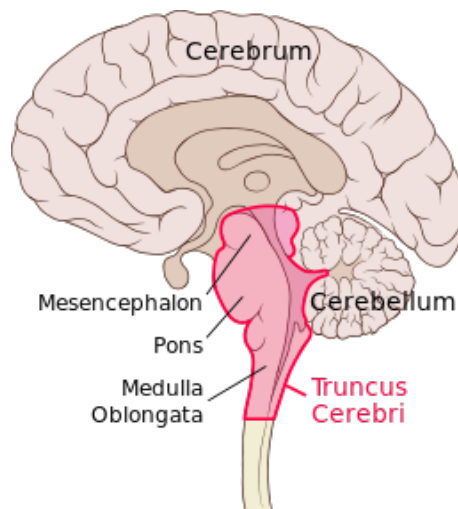
## Mezimozek

Mezimozek neboli *diencephalon* se skládá z pěti funkčně i morfologicky odlišných částí: *epithalamu*, *thalamu*, *metathalamu*, *subthalamu* a *hypothalamu*. Funkcí *thalamu* je integrace a převod signálů z nižších oddílů nervového systému a bazálních ganglií do mozkové kůry. *Hypotalamus*, další součást mezimozku, má významný vliv na regulaci emocí a chování. Tato část mozku je spojena s reakcemi na strach, radost, hlad, žízeň a další základní pocity. *Hypotalamus* také ovlivňuje funkci endokrinního systému, což má vliv na hormonální rovnováhu a fyziologické reakce. Mezi další

funkce mezimozku patří koordinace pohybu a motorické schopnosti a také produkce melatoninu v epifyze, který má vliv na regulaci spánku [3].

## Mozkový kmen

Mozkový kmen zahrnuje část mozku, která navazuje na míchu. Tato oblast se skládá ze tří částí: *medulla oblongata*, *pons Varoli* a *mesencephalon*. Většina zadní plochy mozkového kmene je zakryta mozečkem, které je spojen s jednotlivými částmi kmene skrze párové stonky zvané mozečkové stvoly (*pedunculi cerebellares*). Tato oblast kmene obsahuje bílou hmotu, která zahrnuje sestupné a vzestupné dráhy, a také šedou hmotu, která tvoří jádra, nazývaná *nuclei*, a retikulární formaci. V oblasti mozkového kmene nalezneme životně důležitá reflexní centra, dále zde také vystupují hlavové nervy a procházejí zde důležité sensorické i motorické dráhy [3].



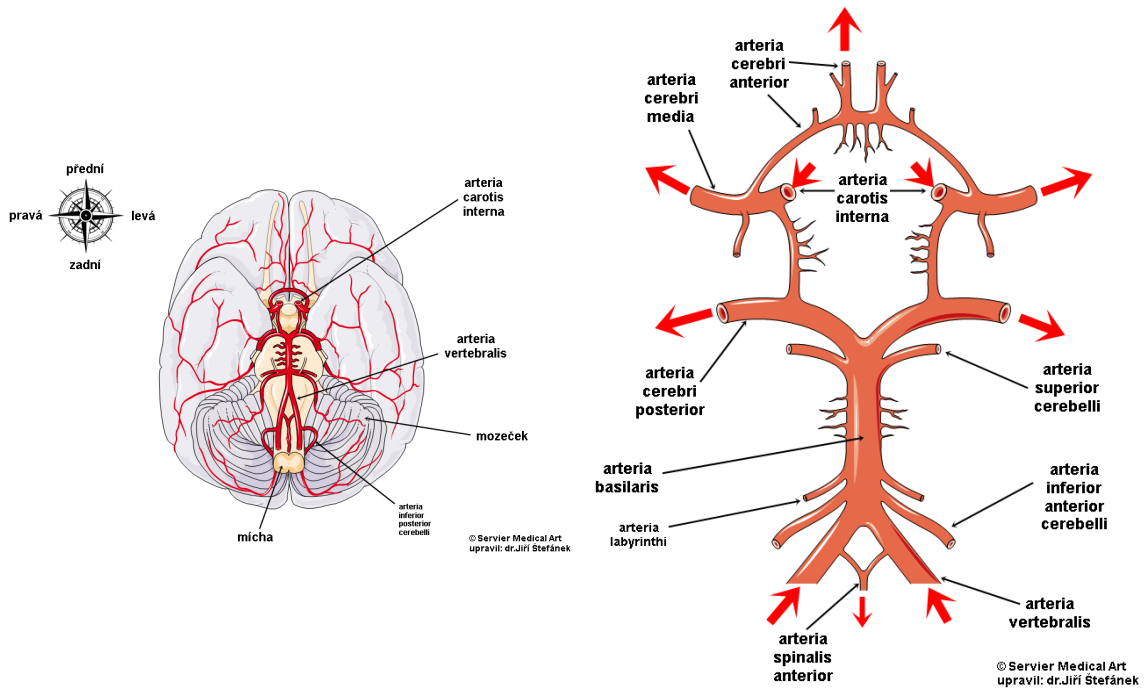
Obr. 1.1: Řez mozkem se znázorněným mozečkem (*cerebellum*), koncovým mozkem (*cerebrum*) a červeně vyznačeným mozkovým kmenem (*truncus encephali*). Převzato z: [4].

## 1.1 Cévní zásobení mozku

Cévní zásobení mozku je velmi důležité pro udržení jeho funkcí. Mozek potřebuje neustálý přísun kyslíku a živin, který je zajištěn složitou sítí krevních cév. Hlavními cévami zásobujícími mozek jsou vnitřní krkavice a vertebrální tepny.

Vnitřní krkavice (*carotis interna*) zásobuje velkou část mozku kyslíkem a živinami. Dělí se na pravou a levou vnitřní krkavici. Vertebrální tepny vstupují do lebky skrze krční páteř a spojují se do bazilární tepny, která zásobuje mozkový kmen a

zadní části velkého mozku. Propojením anteriorního a posteriorního cévního oběhu vzniká tzv. Willisův kruh [13].



Obr. 1.2: Vlevo: Mozek zespuď se základními tepnami. Převzato z: [14]. Vpravo: Schéma Willisova okruhu a přilehlých tepen, červené šipky ukazují směr toku krve. Převzato z: [14].

Willisův kruh slouží jako záložní systém pro krevní zásobení. Jeho funkce je klíčová při kompenzaci začátku ischemie [13].

## 1.2 Cévní mozková příhoda

Cévní mozková příhoda (CMP) je definována jako akutní onemocnění mozku zahrnující rychle se rozvíjející příznaky postižení určité části mozku, nebo celkovou poruchu mozkové funkce, trvající déle než 24 hodin, která může vést k úmrtí. Následky jsou různorodé a závisí na místě postižení. Mohou se projevovat poruchami hybnosti až ochrnutím, problémy s řečí, zvracením, bolestí hlavy, nebo narušeným viděním. Typickým znakem může být povislý koutek [5] [6].

Z hlediska etiologie, je rozeznávána CMP na ischemickou a hemoragickou formu. Nejčastější - ischemická forma CMP tvoří 85 % případů a je způsobena ucpáním cévy, nejčastěji trombem, což má za následek nedostatečné zásobení krví v daném úseku mozku. Hemoragická forma CMP je způsobena prasknutím cévy, což vyvolá krvácení přímo do mozku, či do mozkových obalů, tím opět dojde k nedostatečnému přívodu krve do dané mozkové oblasti [7].

Na vzniku CMP se podílí mnoho faktorů, ať už jsou to faktory ovlivnitelné, nebo neovlivnitelné. Mezi významná rizika patří například dlouhodobě zvýšený krevní tlak, který může vést ke strukturálním změnám v drobných cévách. Kardiovaskulární onemocnění může přispět k uvolnění krevních sraženin přímo do mozkových tepen, a porucha hemodynamiky může zvýšit riziko vzniku CMP. Ateroskleróza, tj. postupné ukládání tukových látek v cévách, má také značný podíl na zúžení hlavních mozkových tepen. Dalšími rizikovými faktory jsou kouření, *diabetes mellitus*, věk a dokonce i genetická predispozice [8].

## 1.3 Diagnostika cévní mozkové příhody

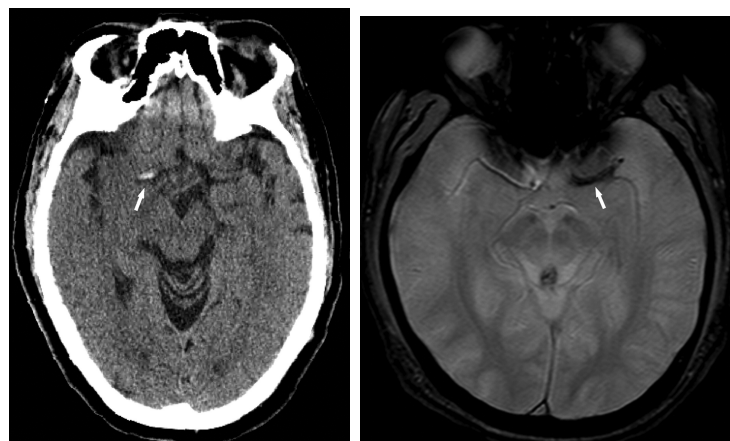
Molekulární události, které jsou spouštěny akutní fokální ischemií, lze seskupit do časově závislé kaskády, která se vyznačuje nízkou produkcí energie, nadměrnou stimulací neuronálních glutamátových receptorů (tzv. excitotoxicita), abnormálním hromaděním sodíkových, chloridových a vápenatých iontů uvnitř neuronů, poškozením mitochondrií, a nakonec buněčnou smrtí. Proto je nejdůležitější co nejrychleji obnovit normální průtok krve do mozku a ochránit neurony tím, že se zasáhne do průběhu ischemické kaskády, nebo ji zpomalí. Výzkumy s využitím magnetické rezonance (MRI) a pozitronové emisní tomografie (PET) naznačují, že během ischemie rychle vzniká jádro infarktové mozkové tkáně, které je obklopeno hypoxickou, ale potenciálně zachranitelnou tkání [9].

Jen přibližně třetina osob s CMP je schopna podle příznaků poznat, že se jedná o CMP. Pokud jsou zaznamenány symptomy, je důležité okamžitě kontaktovat záchrannou službu. Prvním krokem v hodnocení pacienta by mělo být zhodnocení dýchání, průchodnosti dýchacích cest a oběhu pacienta, známé jako "ABC" resuscitace.

U pacientů s podezřením na CMP se provádí měření kyslíkové saturace pomocí pulzní oxymetrie, měření hladiny glukózy, elektrolytů a renálních funkcí v séru, úplný krevní obraz, testy srážlivosti krve, elektrokardiografie a rentgen hrudníku. Neurologické vyšetření by mělo zahrnovat testy na dysfázii, homonymní hemianopii, hemiparézu, nebo hemisenzorickou ztrátu a další známky ložiskového poškození. Přítomnost ložiskového neurologického deficitu a anamnéza náhlého nástupu příznaků bez úrazu naznačují výskyt cévní mozkové příhody. Obecná léčba zahrnuje podávání doplňkového kyslíku a izotonických intravenózních tekutin a antipyretickou léčbu. Urychleně po příjezdu pacienta by měla být provedena diagnostická počítačová tomografie (CT) mozku, aby se odlišilo spontánní intracerebrální krvácení a subarachnoidální krvácení od ischemického infarktu [9] [10].

## 1.4 Zobrazovací techniky v diagnostice CMP

Pro vyšetření a diagnostiku CMP se dnes lékaři neobejdou bez zobrazovacích metod, jako je MRI a CT. Vzhledem k časové náročnosti MRI je stále upřednostňovanější metodou CT, a to i přes to, že u MRI nedochází k ozařování pacienta. CT vyšetření poskytuje detailní a přesné snímky cév případných trombů. Na obrázku 1.3 je zobrazen rozdíl mezi zobrazením CMP pomocí CT a MRI. Princip, na kterém CT systém funguje, je založen na využití rentgenového záření a počítačového zpracování dat [23].



Obr. 1.3: Zobrazení trombu pomocí CT (vlevo) a pomocí MRI (vpravo) s patrným "blooming artefaktem". Převzato z: [45].

CT systém se skládá z dvou hlavních komponent: rentgenového zářiče a detektoru. Tyto komponenty jsou umístěny naproti sobě na opačných stranách pacienta. Rentgenový zářič generuje tenký paprsek rentgenového záření, který je následně

směřován směrem k pacientovi. Toto rentgenové záření proniká tkáněmi a je buď absorbováno, nebo prochází skrz ně, v závislosti na hustotě tkání.

Detektor na opačné straně pacienta zachycuje rentgenové záření, které prošlo tkáněmi. Množství a intenzita rentgenového záření, které je zachyceno detektorem, se liší v závislosti na průchodu záření různými tkáněmi. Tato data jsou následně digitalizována a počítačově zpracována [22].

Počítačový systém CT provádí rekonstrukci obrazu pomocí Radonové transformace na základě dat z detektoru. Pomocí těchto výpočtů je možné rekonstruovat obrazy, které znázorňují různé vrstvy a struktury těla. Díky tomu lze získat detailní a trojrozměrný obraz tkání, orgánů a kostí v těle pacienta.

Jako zlatý standart je považováno nativní CT. To se využívá k zobrazení hemoragie, kdy ložisko je zobrazováno denzitě 30 až 60 HU. Takovéto hodnoty se objevují většinou v prvních minutách krvácení, kdy není ložisko jednoznačně rozlišitelné od mozkové tkáně. Teprve po několika minutách má hemoragické ložisko denzitu 60 až 100 HU [21]. Rovněž se občas v CT snímcích může objevit jev známý jako "swirl sign". Tento jev se projevuje jako proužek, nebo malé ložisko s vysokou hustotou, které má v okolí hypodenzní, nebo izodenzní oblast (což znamená, že je podobné hustotě mozkového parenchymu). Swirl sign indikuje stále probíhající krvácení z prasklé tepny a ukázalo se, že je nezávislým prediktorem pacientovy mortality a špatného klinického stavu během tří měsíců po cévní mozkové příhodě [21].

Při nálezu ischemie je vhodné vyšetření doplnit o angiografické CT vyšetření intrakraniálních tepen za účelem vyloučení aneuryzmatu či cévní malformace. Před angiografií se do pacienta vpraví jodová kontrastní látka. Tyto kontrastní látky mají vysokou hustotu a dobře pohlcují rentgenové záření, což umožňuje získat detailní obrazy cév a tkání. Pro provedení angiografie je často nutné zavést tenký katétr do tepny, obvykle na místo přístupu v blízkosti stehenní tepny, nebo pažní tepny.

Jakmile je katetr zaveden, lékař postupně naviguje katétr cévami směrem k mozku. Během tohoto procesu jsou pořizovány rentgenové snímky nebo CT snímky s kontrastem. To umožňuje zobrazení průtoku krve a detekci eventuálních abnormalit, jako jsou krevní sraženiny, stenózy, nebo aneuryzma [24].

## 1.5 Léčba cévní mozkové příhody

Jak již bylo zmíněno, při léčbě CMP je jedním z nejdůležitějších faktorů ovlivňující úspěšnost léčby a délku rekonvalescence čas a poloha zúžených cest současně s úrovní zapojení kolaterálních cest [11].

Tradiční terapie pro mozkovou ischemickou cévní příhodu dříve zahrnovala rozpuštění krevních sraženin trombolytickými látkami a případně chirurgický zákrok hemikraniektomie. Nicméně výsledky tohoto přístupu byly omezené. Pouze polovina

pacientů na léčbu reagovala pozitivně a ve značné části rozpouštění trombu nezařadilo vůbec. Tato situace vedla k aktivnímu hledání nových metod léčby. Od roku 2015 došlo k revoluci v léčbě mozkové ischemické cévní příhody především díky výsledkům endovaskulární léčby. Tento nový přístup, spojený s tradičními léčebnými postupy a někdy je i nahrazuje a byl rychle přijat v klinické praxi. Doporučeným postupem Cerebrovaskulární sekce České neurologické společnosti České lékařské společnosti Jana Evangelisty Purkyně je do 4,5 hodin od začátku příznaků podání aktivátoru tkáňového plazminogenu, známého také jako t-PA [16]. Ten hraje klíčovou roli v přeměně plazminogenu na plazmin. Plazmin poté rozkládá fibrin, což vede k procesu fibrinolýzy. Rekombinantní rt-PA se aplikuje přímo na místo trombu pomocí mikrokatestrů, které se zavádějí přes femorální tepnu, aortu nebo krční tepnu. Pokud tato léčba je neúčinná, musí se přistoupit k endovaskulární léčbě, tedy vyjmutí trombu pomocí extrakčních systémů [16] [18].

Nejčastěji využívané jsou "stent retrievery". Nejprve je katétr zaveden do cévy v noze pacienta a navigován až do místa, kde trombus blokuje mozkovou cévu. Poté je stent retriever roztažen tak, aby zachytil a obklopil trombus. Jakmile je trombus pevně zachycen, lékař jej opatrně vytáhne z těla pacienta. Nedávné studie ukázaly, že trombektomie pomocí stahovatelného stentu, nebo flexibilního aspiračního katétru byla spojena s vysokou mírou rekanalizace a nízkou mírou komplikací [17] [18].

Rozvoj novějších zařízení a probíhající intenzivní výzkum nám v endovaskulárních procedurách přináší nové výzvy. Jednou z hlavních překážek je rozmanitost sraženin, co se týče konzistence a odstranitelnosti. Někdy můžeme narazit na sraženinu, kterou lze snadno odstranit pomocí aspirace, nebo trombektomie, avšak jindy se můžeme potýkat se sraženinou, která je odolná vůči dostupným nástrojům. Případy pacientů s CMP nám ukazují, že vlastnosti sraženin jsou různorodé a někdy nepředvídatelné [15].

## 1.6 Charakteristika trombů a jejich patofyziologie

Trombus vzniká obvykle v důsledku poškození cévní stěny a následně aktivace krevních destiček. Poškození cévní stěny může být způsobeno různými faktory, včetně aterosklerózy, cévních zánětů, nebo jiných cévních onemocnění. Toto poškození vytváří místo, kde se krevní složky shlukují, kde začínají vytvářet základní strukturu trombu. Současně s aktivací destiček dochází k aktivaci koagulačního systému krve, který zahrnuje různé krevní proteiny, které spolupracují na vytvoření fibrinové sítěviny. Tato síťovina zpevňuje trombus a zadržuje další krevní buňky a plazmatické složky. Trombus postupně roste, jak se více krevních destiček a fibrinové sítěviny přidává k jeho struktuře, což může vést k blokádě cév.

Složení trombu je obecně nehomogenní a můžeme tedy tromby podle hlavního podílu některé ze složek rozdělit do tří základních skupin, a to jsou tromby složené převážně z červených krvinek, fibrinu anebo bílých krvinek. Pokroky v endovaskulární trombektomii a souvisejících zobrazovacích metodách vytvořily jedinečnou možnost analyzovat tromby odstraněné z mozkových tepen. Znalost biologické povahy trombu nám umožňuje zkoumat vztahy mezi složením trombu, mechanismu vzniku trombu a úspěšností léčby, i přesto, že prozatím není žádný jednotný závěr vlivu složení trombu na úspěšnost léčby [12].

V jedné z publikovaných studií bylo uvedeno, že tromby z kardioembolizmu měly významně vyšší podíl červených krvinek ve srovnání s tromby z aterosklerózy velkých tepen. Podíl červených krvinek byl vyšší u sraženin se známkou náchylnosti cév než u sraženin bez ní, zatímco podíl fibrinu a trombocytů byl vyšší u sraženin bez známky náchylnosti cév, než u těch s ní [18].

V posledních letech se velké množství studií nejen z medicínské oblasti týká strojového učení. To se začíná využívat k rychlé a přesné diagnostice a prognóze nemocí z lékařských snímků CT a MRI. Možností využití strojového je nepřeborné množství, a tak se zcela přirozeně začal výzkum zabývat možnostmi analýzy trombů pomocí těchto metod. Důležitou snahou klinického neurozobrazování u cévní mozkové příhody je získat informace týkající se složení trombu ze zobrazovacích snímků pacienta a předpovědět složení trombu a jeho pravděpodobnou odpověď na léčbu.

## 2 Analýza trombu v CT datech

### 2.1 Předzpracování

Předzpracování surových snímků získaných z CT systémů je důležitým krokem ve všech přístupech strojového učení a procesu segmentace. Ačkoli není pro předzpracování CT skenů definován standartní postup, často je třeba snížit dynamický rozsah obrazu. K tomu je využívána metoda tzv. "windowing". Tato technika umožňuje manipulovat s hodnotami pixelů u šedé škály CT obrazu za účelem zvýraznění specifických struktur. Úpravou úrovně okna je možné měnit jas obrazu. Kontrast lze upravit úpravou šířky okna. Větší šířka okna zobrazí širší rozsah hodnot CT. Dalším běžným krokem je odstranění lebky a krku, protože by mohly být zdrojem chyb při segmentaci. Jedním z možných postupů je shlukování objemových dat do tří skupin (pozadí, lebka a tkáň včetně hemoragie). Následně je vhodné provést erozi, nebo dilataci, podle našich požadavků. Obrazy jsou běžně znehodnoceny šumem. Šum se může projevovat jako širokopásmový, nebo impulzní. Šedý šum může vznikat v důsledku tepelných fluktuací systému, nedokonalostí zobrazovacího zařízení, nebo operací jako je kvantování. Tento typ šumu je náhodný a lze ho potlačit pomocí Gaussovského filtru. Naopak impulzní šum je eliminován pomocí mediánového filtru [26] [27].

Registrace obrazů neboli slícování je dalším krokem zpracování obrazu. Při registraci dochází k zarovnání a porovnání snímků z různých úhlů, nebo časových bodů. Snímek mozku může být pořízený z různých úhlů, nebo v různých časech náchylný k malým posunům, nebo deformacím v důsledku pohybu pacienta, změněné pozice hlavy, nebo otřesům. Nejčastěji je používána geometrická transformace pro hrubé zarovnání snímku a pro preciznější zarovnání se používá transformace afinní [28] [25].

### 2.2 Analýza pomocí konvenčních statistických metod

Při analýze trombu v CT datech můžeme rozlišit dva základní přístupy. Jedním z nich je konvenční statistika, která využívá tradiční statistické metody, jako jsou například statistické testy, nebo regresní analýza. Vědci v této studii se zaměřili na zpracování dat a analýzu statistických vlastností trombu CT a MRI snímků [20].

Mezi zkoumané parametry patří například délka, objem, hustota, propustnost, ale i složitější radiomické příznaky, které nejsou z obrazu na první pohled patrné. Dle studie publikované v časopisu *Neuroradiology* [19] u některých parametrů byla zjištěna souvislost v efektivitě léčby. Předchozí studie ukazovaly, že jednoduché parametry mohou mít vliv na predikci úspěchu léčby, jejich možnosti, však byly omezené.

Jedním ze závěrů bylo například, že menší sraženiny se snadněji odstraňují pomocí mechanické trombektomie, zatímco krátké, ale propustnější sraženiny se lépe rekanalizují pomocí intravenózní alteplázy [19].

V tomto článku autoři nejdříve získali CT snímky pacientů. Následně byla provedena segmentace trombů a extrakce 293 radiomických příznaků. Tyto radiomické příznaky byly rozděleny do tří skupin, do skupiny zkoumající tvar a velikost, statistické příznaky prvního řádu a statistické příznaky vyšších řádů. Byla provedena analýza a vývoj klasifikačního modelu pro predikci úspěchu mechanické trombektomie. Pro selekci důležitých radiomických příznaků byla použita Spearmanova korelace pro filtraci validních příznaků, testy normálního rozložení a multivariační analýza. Po provedení těchto testů bylo vybráno 6 neredundantních radiomických příznaků, které byly použity pro trénink a validaci modelu. Mezi vybrané příznaky patří: "Large Dependence Low Gray-Level Emphasis", koeficient špičatosti, matice plošných šedotónových závislostí, co-occurrence matice a korelační koeficient. Poté byly tromby zkoumány histologicky a výsledky porovnány. Výsledky studie ukazují, že sraženiny s většími homogenními oblastmi erytrocytů a menšími oblastmi agregátů fibrinu a destiček byly spojeny s pozitivními výsledky mechanické trombektomie. Výsledkem tedy je, že sraženiny pocházející z kardiogenního zdroje mohou mít tendenci mít strukturu, která lépe reaguje na mechanickou trombektomii a také to, že výše uvedených 6 radiomických příznaků má lepší prediktivní hodnotu, než tradiční obrazové příznaky jako jsou délka, nebo objem [19] [20].

## 2.3 Analýza pomocí strojového učení

Druhým přístupem je použití metod strojového učení. Jednou z publikovaných studií využívajících metod strojového učení je třeba i studie využívající software MenthinksLVO [35]. Studie zahrnovala 1453 pacientů s podezřením na akutní mozkovou příhodu, kteří podstoupili nekontrastní počítačovou tomografii (NCCT) a počítačovou tomografickou angiografii (CTA). Dalším krokem byla segmentace trombu. Obrazy CTA byly automaticky zarovnány s obrazy NCCT pomocí pevné registrace s 3D Slicer [29]. Kvalita registrace byla posouzena lidským expertem. Poté bylo provedeno manuální ohraničení trombu vrstva po vrstvě z axiálních pohledů na NCCT obrazy pomocí ITK-SNAP [40]. Tato ruční metoda segmentace byla využita jako primární analýza. Pro zhodnocení variability způsobené ruční segmentací provedl další lidský expert manuální segmentaci stejných dat. Výsledky získané touto metodou byly srovnány s výsledky primární analýzy. Dále byla provedena analýza Pearsonovy korelace ohledně délky trombu a objemu. Pro výběr příznaků byla použita lineární diskriminační analýza. Byly vybrány tři skupiny příznaků a byly provedeny všechny možné kombinace, kdy nakonec byly seřazeny podle důležitosti. Analýza ROC byla

využita k porovnání rysů v každé skupině. Výsledné příznaky byly poté seřazeny od nejvyšší po nejnižší důležitost v každé skupině na základě hodnoty plochy pod křivkou (AUC). Optimální počet příznaků pro každou skupinu byl určen pomocí kombinace příznaků, která dosáhla nejvyšší hodnoty AUC a zbytek příznaků, které byly považovány za redundantní, byly eliminovány. Poté byla provedena pětinasobná křížová validace, při které byli subjekty náhodně rozděleni do pěti skupin stejné velikosti. V rámci těchto experimentů byl model strojového učení optimalizován na čtyřech z pěti skupin a následně byl vyhodnocen na subjektech ve zbývající páté skupině. Všechny křížové validační experimenty byly opakovány desetkrát za účelem zhodnocení stability vybraných příznaků.

Výsledkem je, že příznaky z NCCT jsou nejpřesnější s  $AUC\ 0,78 \pm 0,02$ , následované šesti příznaky definující změnu radiomiky s  $AUC\ 0,77 \pm 0,04$ , zatímco patnáct příznaků CTA mají nejnižší přesnost s  $AUC\ 0,74 \pm 0,03$ . Kombinace těchto příznaků do různých skupin výrazně zvyšuje přesnost predikce. Zejména dvanáct příznaků vybraných ze skupiny kombinující NCCT, CTA a radiomické změny poskytuje nejlepší přesnost predikce ( $AUC\ 0,85 \pm 0,03$ ) a byly významně lepší, než klinicky měřitelné příznaky, jako je délka trombu, objem trombu, permeabilita trombu, měřená jako průměrný nárůst absorpce od NCCT do CTA, maximální absorpce v CTA, nebo maximální nárůst absorpce od NCCT do CTA. [36] [37]

## 3 Metody zpracování dat

V této kapitole jsou představeny metody zpracování obrazů a dat, které jsou využity v praktické části práce. V první podkapitole jsou představeny binární morfologické operace. Binární morfologické operace jsou klíčovými nástroji v oblasti zpracování obrazu, umožňujícími manipulaci a analýzu obrazových dat na základě jejich tvarů a struktur. Tyto operace se používají k úpravě obrazů, odstranění šumu, segmentaci objektů, a dalším úpravám, které zlepšují jejich interpretaci[31].

### 3.1 Binární morfologické operace

#### Eroze

Eroze je morfologická operace v oblasti zpracování obrazu, která se používá k zmenšení objektů na obraze tím, že odstraňuje jejich okraje. Během eroze se operátor posouvá po obraze. Pokud alespoň jeden pixel v okolí tohoto vzoru patří k pozadí, tedy jeho hodnota je nula, je daný pixel na výstupním obraze také nastaven na nulu. Tato operace se využívá při odstraňování malých detailů a zpřesňování hranic objektů [31].

#### Dilatace

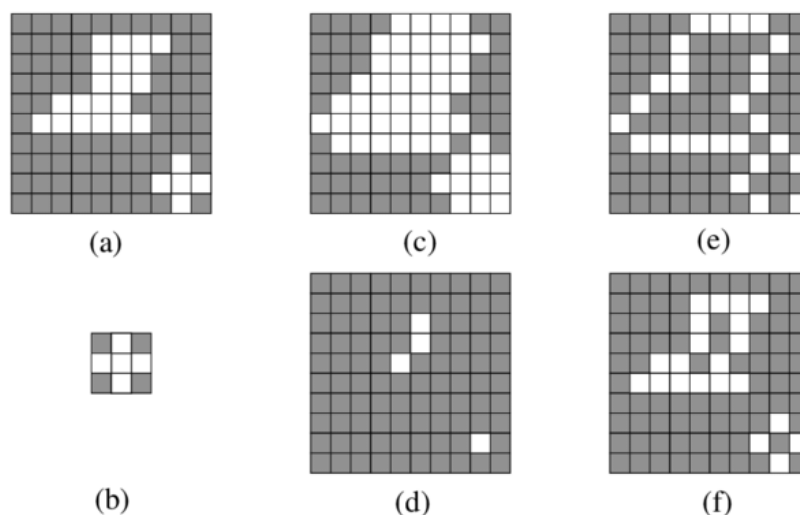
Dilatace je opakem eroze a zvětšuje objekty na obraze tím, že rozšiřuje jejich okraje. Pokud alespoň jeden pixel v okolí operátoru patří k objektu a má tedy hodnotu pixelu 1, je pixel na výstupním obraze nastaven na jedničku. Dilatace je užitečná pro spojování blízkých objektů a vyplňování malých mezer mezi nimi [31].

#### Uzavření

Uzavření je další morfologická operace, která kombinuje dilataci a erozi. Nejprve se provádí dilatace, která uzavírá malé mezery a praskliny v objektech, a poté na výsledek dilatace následuje eroze. Tato operace pomáhá spojit praskliny v objektech a uzavřít malé mezery, což vede k úplnému uzavěru objektů na obraze [31].

#### Otevření

Otevření je morfologická operace, která také kombinuje erozi a dilataci. Nejprve se provádí eroze, která odstraní malé objekty a šum z obrazu, a poté na výsledek eroze následuje dilatace. Otevření je efektivní při oddělování blízkých objektů a odstranění jemného šumu [31].



Obr. 3.1: Příklady morfologických operací. Vstupní pole (a) Originální obraz (b) morfologický operátor (c) dilatace (d) dilatace (e) otevření (f) uzavření. Převzato z: [30].

## 3.2 Extrakce příznaků

Extrakce radiomických a texturálních příznaků z medicínských obrazů umožňuje získání kvantitativních informací o struktuře a vlastnostech tkání. Příznaky mohou být extrahovány jak voxelově v případě 3D obrazů, nebo pouze z pixelů oblasti zájmu ze 2D obrazů. V následujícím textu představíme metody používané při extrakci příznaků [32].

### 3.2.1 Statistické metody prvního řádu

Jednou z nejzákladnějších metod analýzy textur v obraze je výpočet statistik prvního řádu, kde se zaměřujeme na distribuci intenzit pixelů v dané oblasti zájmu v obraze. Tyto příznaky poskytují základní informace o celkovém chování intenzit v určité oblasti. K extrakci radiomických příznaků může být použit například balíček Pyradiomics [34]. Níže jsou uvedeny základní příznaky prvního řádu dostupné v knihovně Pyradiomics, které byly použité v praktické části práce [33].

#### Energie

Měří celkovou intenzitu pixelů v oblasti zájmu. Vyšší hodnota indikuje vyšší součet druhých mocnin intenzit. Lze definovat takto:

$$\text{Energie} = \sum_{i=1}^n (X(i)+c)^2 \quad (3.1)$$

kde  $n$  je počet pixelů v oblasti zájmu a  $X(i)$  je intenzita  $i$ -tého pixelu.

### **Celková energie**

Podobná energii, ale normalizovaná objemem voxelu v krychlových milimetrech. Tato hodnota se používá k eliminaci efektu velikosti oblasti zájmu. Jednou z možných interpretací je tato rovnice:

$$\text{Celková energie} = V_{\text{voxel}} \sum_{i=1}^n X(i) + c^2 \quad (3.2)$$

kde  $n$  je počet pixelů v oblasti zájmu,  $X_i$  je intenzita  $i$ -tého pixelu.

### **Entropie**

Měří míru náhodnosti ve vzorech intenzit. Vyšší entropie naznačuje větší rozmanitost hodnot.

$$\text{Entropie} = - \sum_{i=1}^n p_i \log_2(p_i) \quad (3.3)$$

kde  $n$  je počet různých intenzit v oblasti zájmu a  $p_i$  je pravděpodobnost výskytu intenzity  $i$  v oblasti zájmu.

### **Minimum**

Nejnižší hodnota intenzity v oblasti zájmu.

$$\text{Minimum} = \min(X) \quad (3.4)$$

kde  $X$  jsou všechny intenzity v oblasti zájmu.

### **10. percentil**

Hodnota, pod kterou padá 10 % hodnot intenzit v oblasti zájmu.

### **90. percentil**

Hodnota, pod kterou padá 90 % hodnot intenzit v oblasti zájmu.

### **Maximum**

Nejvyšší hodnota intenzity v oblasti zájmu.

$$\text{Maximum} = \max(X) \quad (3.5)$$

kde  $n$  je počet pixelů v oblasti zájmu a  $X$  je intenzita pixelu.

## Průměr

Průměrná hodnota intenzity v oblasti zájmu.

$$\text{Průměr} = \frac{1}{Np} \sum_{i=1}^{Np} X(i) \quad (3.6)$$

kde  $Np$  je počet pixelů v oblasti zájmu a  $X(i)$  je intenzita  $i$ -tého pixelu.

## Medián

Střední hodnota rozdělení intenzit v oblasti zájmu.

## Interkvartilový rozsah

Rozdíl mezi 75. percentilem a 25. percentilem. Měří rozsah, ve kterém se nachází střední 50 % hodnot.

$$\text{Interkvartilový rozsah} = Q3 - Q1 \quad (3.7)$$

kde  $Q3$  je 75. percentil a  $Q1$  je 25. percentil.

## Rozsah

Rozdíl mezi nejvyšší a nejnižší hodnotou intenzity v oblasti zájmu

$$\text{Rozsah} = \text{Max} - \text{Min}. \quad (3.8)$$

## Průměrná absolutní odchylka

Průměrná vzdálenost všech hodnot intenzit od průměrné hodnoty

$$\text{Průměrná absolutní odchylka} = \frac{1}{n} \sum_{i=1}^n |X(i) - \bar{X}|. \quad (3.9)$$

## Robustní průměrná absolutní odchylka

Průměrná vzdálenost hodnot od průměrné hodnoty, ale počítána pouze pro 10 % - 90 % rozsah intenzit

$$\text{Robustní průměrná absolutní odchylka} = \frac{1}{m} \sum_{i=1}^m |X(i) - \bar{X}| \quad (3.10)$$

kde  $m$  je počet hodnot v 10 % - 90 % rozsahu intenzit.

### Kvadratický průměr

Jedná se o druhou odmocninu aritmetického průměru druhých mocnin intenzit v oblasti zájmu

$$\text{Kvadratický průměr} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X(i) + c)^2}. \quad (3.11)$$

### Směrodatná odchylka

Měří množství rozptylu hodnot od průměru

$$\text{Směrodatná odchylka} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X(i) - \bar{X})^2}. \quad (3.12)$$

### Koeficient šikmosti

Měří asymetrii rozdělení hodnot kolem průměru

$$\text{Koeficient šikmosti} = \frac{\frac{1}{n} \sum_{i=1}^n (X(i) - \bar{X})^3}{\left(\frac{1}{n} \sum_{i=1}^n (X(i) - \bar{X})^2\right)^{3/2}}. \quad (3.13)$$

### Koeficient špičatosti

Měří špičatost rozdělení hodnot v oblasti zájmu

$$\text{Koeficient špičatosti} = \frac{\frac{1}{n} \sum_{i=1}^n (X(i) - \bar{X})^4}{\left(\frac{1}{n} \sum_{i=1}^n (X(i) - \bar{X})^2\right)^2}. \quad (3.14)$$

### Variance

Měří průměr druhých mocnin vzdáleností od průměru. Je to ukazatel rozptylu hodnot.

$$\text{Variance} = \frac{1}{n} \sum_{i=1}^n (X(i) - \bar{X})^2 \quad (3.15)$$

kde  $n$  je počet pixelů v oblasti zájmu a  $X(i)$  je intenzita  $i$ -tého pixelu.

### Uniformita

Měří homogenitu hodnot intenzit. Vyšší hodnota značí, že hodnoty jsou vyrovnanější [34]

$$\text{Uniformita} = \sum_{i=1}^n (p(i))^2. \quad (3.16)$$

### 3.2.2 Co-occurrence matice (GLCM)

Co-occurrence matice, někdy také v literatuře uváděny jako "matice současného výskytu", jsou po statistikách prvního řádu pravděpodobně nejznámější metodou extrakce příznaků. Jedná se o statistickou metodu druhého řádu, zachycující prostorové vztahy intenzit pixelů v oblasti obrazu omezené maskou. Tuto matici lze chápat jako 2D sdružený histogram dvou obrazů A a B stejných rozměrů s  $q$  a  $r$  stupni šedi, definovaný vztahem: [34]

$$AB(l, m) = \sum_{i=1}^q \sum_{k=1}^r \delta(i-l, k-m) \times A(i, k) \times B(i + \Delta x, k + \Delta y) \quad (3.17)$$

kde  $\delta$  je delta funkce s hodnotami 1, 0, nebo -1 pro  $i-l = 0$ ,  $i-l = 1$  a  $i-l = -1$  a  $k-m = 0$ ,  $k-m = 1$  a  $k-m = -1$ , a  $\Delta x$ ,  $\Delta y$  jsou parametry určující posunutí pixelů. [38]

Při analýze obrazu s  $q$  úrovněmi šedi jsou zkoumána okolí pixelů  $(x, y)$  a  $(x + \Delta x, y + \Delta y)$ , ze kterých se vypočítává 2D sdružený histogram. Výsledná matice  $q \times q$  představuje co-occurrence matici, která závisí na hodnotách parametrů definujících posunutí  $\Delta x$  a  $\Delta y$ . V praktických aplikacích se obvykle počítá s více maticemi pro různé hodnoty posunutí  $\Delta x$  a  $\Delta y$  a ve více směrech ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  a  $135^\circ$ ).

Z co-occurrence matic lze odvodit různé parametry, jako je energie, entropie, kontrast, modus a korelační koeficienty, které mohou sloužit jako další příznaky [38].

### 3.2.3 Gray Level Size Zone Matrix (GLSZM)

GLSZM je další metodou výpočtu příznaků z obrazu. GLRLM je metoda, která definuje délku/počet po sobě jdoucích pixelů, jejichž hodnota úrovně šedi je stejná. Na rozdíl od jiných metod jako GLCM a GLRLM, GLSZM je nezávislá na rotaci, generuje pouze jednu matici pro všechny směry v oblasti zájmu [38].

GLSZM matice poté slouží k výpočtu dalších příznaků jako je např. variance, entropie, a mnoho dalších [34] [39].

### 3.2.4 Gray Level Run Length Matrix (GLRLM)

GLRLM je sofistikovanou metodou v oblasti extrakce texturálních vlastností z obrazu, kde se zaměřuje na délky po sobě jdoucích pixelů s identickými úrovněmi šedi. GLRLM umožňuje kvantifikovat distribuci délek běhů pixelů s konkrétní šedou úrovní a tím zachytit charakteristické texturální vlastnosti oblasti zájmu.

Princip GLRLM spočívá v tvorbě matice, kde každý prvek reprezentuje délku běhu pro konkrétní úroveň šedi. Tato matice následně umožňuje odvodit statistiky, které reflektují texturální charakteristiky obrazu.

Výpočet rysů zahrnuje hodnocení GLRLM pro každý úhel samostatně, následované výpočtem průměru těchto hodnot [34] [39].

### 3.2.5 Neighbouring Gray Tone Difference Matrix (NGTDM)

Tato matice se zaměřuje na rozdíly mezi okolními šedými tóny v obraze a umožňuje zachytit jemné texturální charakteristiky, které nemusí být snadno patrné na první pohled. NGTDM vytváří matici, kde každý prvek odpovídá počtu pixelů s určitým rozdílem šedé intenzity mezi okolními pixely v obraze. Tato informace je následně využita k extrakci texturálních příznaků, jako jsou průměrné šedé rozdíly, variabilita šedých rozdílů a další.

NGTDM kvantifikuje rozdíl mezi hodnotou šedé barvy a průměrnou šedou hodnotou okolí ve vzdálenosti  $\delta$ . Součet absolutních rozdílů pro úroveň šedé barvy  $i$  je uložen v matici. Označme  $X_{gl}$  jako sadu segmentovaných voxelů a  $x_{gl}(j_x, j_y, j_z) \in X_{gl}$  jako šedou úroveň voxelu na pozici  $(j_x, j_y, j_z)$ . Průměrná šedá úroveň okolí je dána vzorcem:

$$\bar{A}_i = \bar{A}(j_x, j_y, j_z) = \frac{1}{W} \sum_{k_x=-\delta}^{\delta} \sum_{k_y=-\delta}^{\delta} \sum_{k_z=-\delta}^{\delta} x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z), \quad (3.18)$$

kde  $(k_x, k_y, k_z) \neq (0, 0, 0)$  a  $x_{gl}(j_x + k_x, j_y + k_y, j_z + k_z) \in X_{gl}$ . Zde  $W$  představuje počet voxelů v okolí, které jsou také v  $X_{gl}$ . [34] [39]

### 3.2.6 Gray Level Dependence Matrix (GLDM)

GLDM kvantifikuje závislosti úrovní šedé barvy na snímku. Závislost úrovně šedé barvy je definována jako počet spojených voxelů ve vzdálenosti  $\delta$ , které jsou závislé na středovém voxelu. Sousední voxel s úrovní šedé barvy  $j$  je považován za závislý na středovém voxelu s úrovní šedé barvy  $i$ , pokud  $|i - j| \leq \alpha$ . V matici závislosti úrovně šedé barvy  $P(i, j)$  popisuje  $(i, j)$ -tý prvek početkrát, kdy se voxel s úrovní šedé barvy  $i$  s  $j$  závislými voxely ve svém okolí objevuje na snímku [34] [39].

## 3.3 Statistická analýza

### 3.3.1 Pearsonův korelační koeficient

Pearsonův koeficient korelace představuje statistický nástroj pro hodnocení síly lineární asociace mezi dvěma proměnnými, označovaným jako "r". Tento koeficient nabývá hodnot od +1 do -1, přičemž hodnota 0 naznačuje žádnou korelaci, kladná

hodnota ukazuje pozitivní korelaci, a záporná hodnota ukazuje negativní korelaci [41].

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (3.19)$$

kde  $r$  je Pearsonův koeficient korelace, (3.20)

$n$  je počet pozorování (3.21)

### 3.3.2 Mann-Whitney test

Mann-Whitneyho test, známý také jako Wilcoxonův test pro dva nezávislé vzorky, je neparametrický statistický test používaný k porovnání mediánů dvou nezávislých skupin dat. Na rozdíl od parametrických testů, jako je t-test, Mann-Whitneyho test nevyžaduje předpoklady o normálním rozdělení dat a je robustní vůči odlehlým hodnotám.

Princip Mann-Whitneyho testu spočívá v převodu hodnoty pozorovaných dat na pořadí a následném porovnání součtu pořadí mezi dvěma skupinami. Hypotézy Mann-Whitneyho testu jsou formulovány takto:

- Nulová hypotéza ( $H_0$ ): Mediány obou skupin jsou shodné.
- Alternativní hypotéza ( $H_1$ ): Mediány obou skupin jsou různé.

Výsledek testu poskytuje informaci o tom, zda můžeme zamítnout nulovou hypotézu ve prospěch alternativní hypotézy. P-hodnota, kterou obdržíme po provedení Mann-Whitneyho testu, měří pravděpodobnost pozorovaných dat za předpokladu platnosti nulové hypotézy. Pokud je p-hodnota nižší, než námi zvolená hladina významnosti (např. 0,05), můžeme nulovou hypotézu zamítnout [41] [42].

## 4 Experimentální část

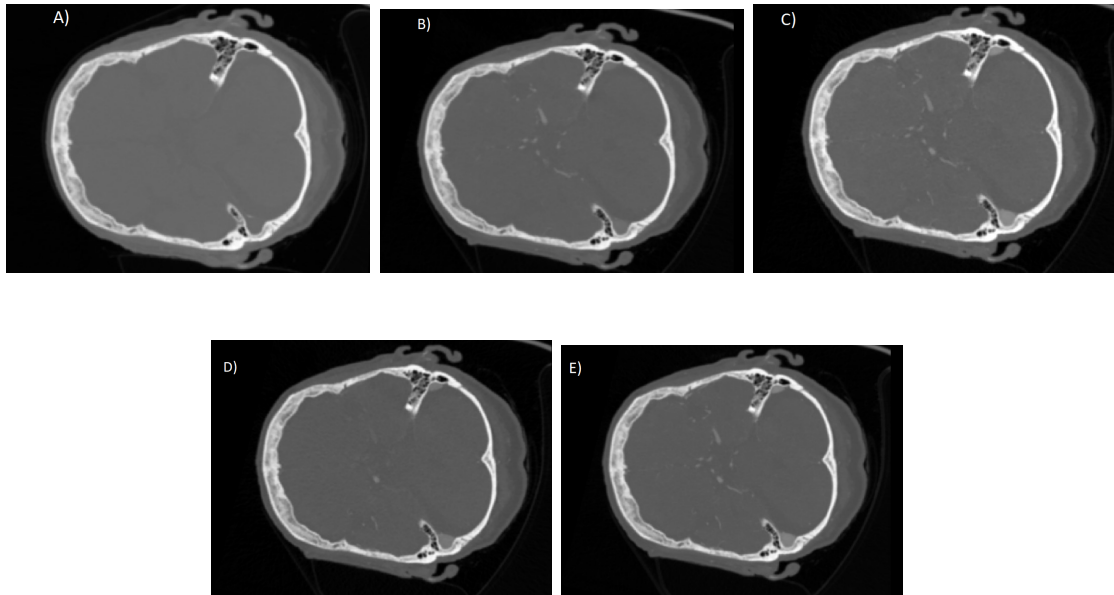
V této kapitole jsou popsány všechny kroky a metody, použité v praktické části práce. Cílem praktické části práce je pokročilá analýza multifázických CT dat mozku, kdy jsou analyzována data pacientů s CMP. Praktická část práce byla navržena v programovacím jazyce Python a k extrakci příznaků byl použit balíček Pyradiomics [34].

### 4.1 Použitá data

Použitý dataset jsou anonymizované snímky osmi pacientů s diagnostikovanou cévní mozkovou příhodou. Jedná se o nativní CT skeny mozku, kontrastní multifázické CTA snímky ve třech časových okamžicích a fúzovaný snímek všech výše zmíněných. Dataset byl nasnímán a poskytnut nemocnicí FNUSA v rámci smluvní výzkumné spolupráce FNUSA a UBMI. Data byla nasnímána pomocí přístroje Brilliance iCT 256 (Phillips Medical Systems, Eindhoven, Nizozemsko).

Postup nasnímání dat zahrnoval nekontrastní CT vyšetření mozku (nCT), více-fázovou CT angiografii (mCTA) hlavy a krku a perfuzní CT mozku. Nekontrastní CT vyšetření pokrývalo oblast mozku pomocí 0,9 mm axiálních překrývajících se řezů. Při mCTA byly řezy snímány s časovým odstupem 8 sekund a s překrývajících se řezy o velikosti 0,9 mm. Perfuzní snímky zobrazovaly 8 cm výřez mozku s 10 souvislými řezy. Během celého vyšetření bylo provedeno 37 cyklů s dobou cyklu 1,8 sekundy. [46]

Poskytnutá data jsou již předzpracovaná a sjednocená ve formátu NIFTI. Vizualizace dostupných dat je uvedena na přiloženém obrázku: 4.1.

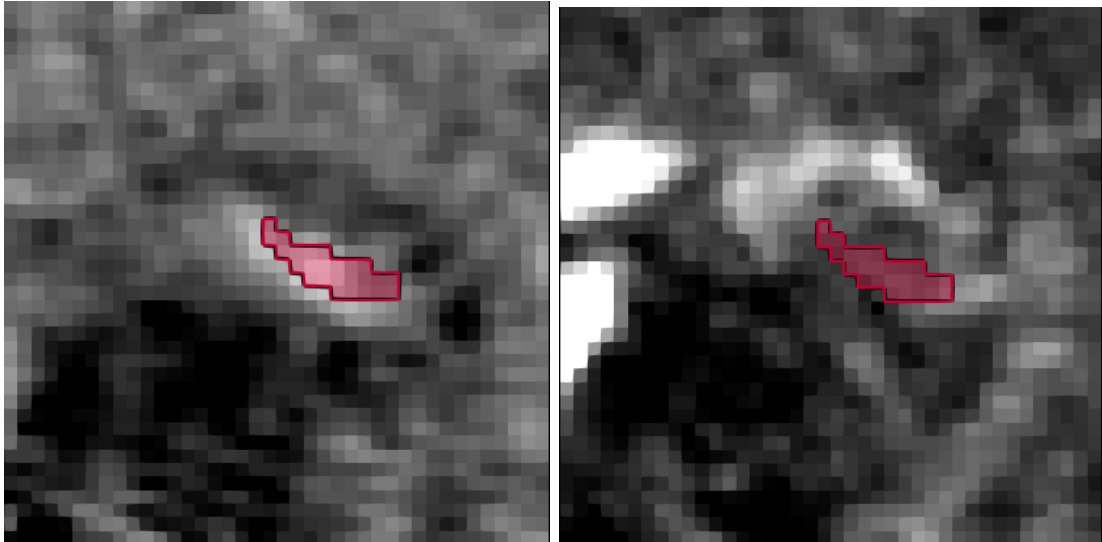


Obr. 4.1: a) Nativní CT, b) CTA fáze 1, c) CTA fáze 2, d) CTA fáze 3, e) Fúzovaný snímek.

#### 4.1.1 Předzpracování dat a extrakce příznaků

Cílem této analýzy je detailně zkoumat charakteristiky trombu, okolních měkkých tkání a jejich vzájemné vztahy v kontextu získaných příznaků a následná segmentace trombu. Před samotnou analýzou bylo nezbytné provést předzpracování dat. Dataset obsahoval již hotové binární masky anotované lékaři pro každého pacienta. V binární masce jsou tromby označeny hodnotou jedna, zatímco okolní tkáně jsou označeny hodnotou nula. Nejdříve byly v binární masce detekovány hodnoty 1 a následně byly CT skeny oříznuty na velikost  $40 \times 40 \times 20$  voxelů a střed byl zarovnán s centrem binární masky. Vybraná velikost byla zvolena tak, aby byla snížena výpočetní náročnost s cílem zajistit, aby při oříznutí nedocházelo ke ztrátě voxelů trombu. V průběhu analýzy byly využívány nástroje jako napari [43], SimpleITK [44] a MITK [47] k vizualizaci a manipulaci s daty. Na obrázku 4.1.1 jsou zobrazena data, ze kterých již byly extrahovány radiomické příznaky. V nativních snímcích dosahuje trombus vyšších jasových hodnot a trombus má tendence v obrazech "svítit" zatímco ve fúzovaných snímcích, je trombus pozorovatelný hůře.

Pro extrakci radiomických příznaků byl použit předdefinovaný soubor parametrů, který obsahuje konfiguraci extraktoru příznaků. Proces extrakce příznaků byl prováděn s lineárním interpolátorem a s hodnotou vzdálenosti mezi vzorky po přeškálování na  $[1, 1, 1]$ . Extrahované příznaky jsou popsány v kapitole 3.2, celkově jich bylo vypočítáno 91. Detailní přehled příznaků je uveden v příloze A.1, ve zbytku práce budou označovány z důvodu větší přehlednosti čísly použitými v tabulce. Výsledky



Obr. 4.2: Ukázka předzpracovaných dat ve stejném řezu pro nativní snímek s patrným trombem, dosahujíc vyšších jasových hodnot (vlevo) a fúzovaný (vpravo) snímek, kde naopak vyšší jasové hodnoty odpovídají cévám.

byly uloženy ve formátu NRRD. Dalším krokem, bylo odstranění odlehlých hodnot, pomocí z-score a jejich standardizace. Standardizace byla provedena na každém pacientovi zvlášť.

## 4.2 Experiment 1: Korelační a PCA analýza radiomických příznaků voxelů trombu a pozadí

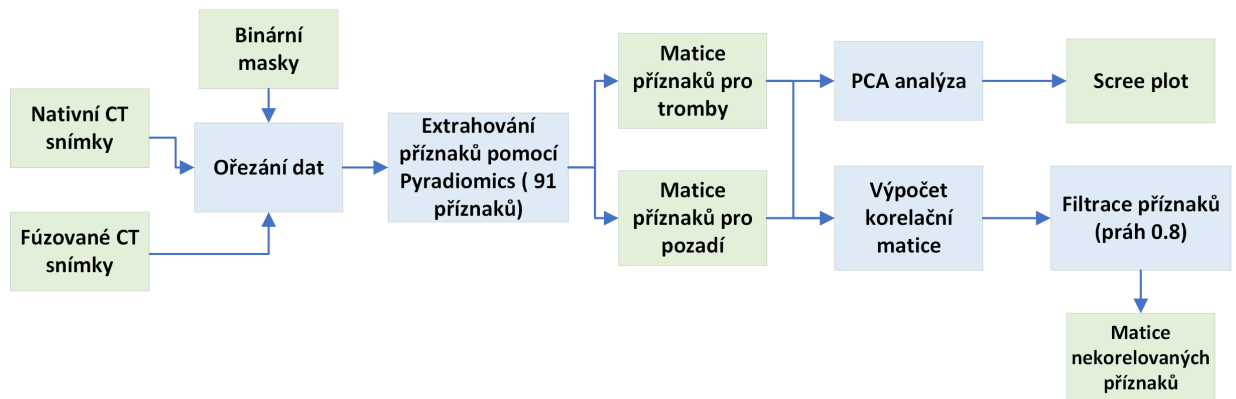
Tento experiment se zabývá analýzou jednotlivých příznaků, jejich korelací a metodami výběru relevantních příznaků. Analýza příznaků je důležitá především pro to, že odstraníme-li korelované příznaky zvýšíme efektivitu modelů, snížíme přeučení modelu, zvýšíme interpretovatelnost a v neposlední řadě snížíme výpočetní náročnost.

### 4.2.1 Návrh experimentu

Analýza příznaků má za úkol zjednodušit výpočetní proces a zlepšit klasifikační model tím, že identifikuje nevýznamné příznaky a odstraní je, aniž by došlo ke ztrátě klasifikační síly. K tomuto účelu se využívá Principal Component Analysis (PCA), která odstraňuje příznaky s minimální informační hodnotou a korelační analýza, která identifikuje korelované příznaky. Je důležité si uvědomit, že PCA pracuje s předpokladem, že nejdůležitější jsou příznaky s největší variabilitou, což nemusí vždy platit.

Odstranění korelovaných příznaků umožňuje vyřadit prediktory, které popisují stejné vlastnosti. Dalším důvodem pro použití korelační analýzy je optimalizace procesu učení metod strojového učení, kterým se budeme věnovat v posledním experimentu. Pro výpočet korelace je použit Pearsonův korelační koeficient. Byla použita data nativních snímků, již standardizovaná a předzpracovaná viz výše. Korelace byla provedena na vyextrahovaných datech trombu a pozadí zvlášť. Po vytvoření korelační matice jsou identifikovány páry korelovaných příznaků, jejichž absolutní hodnoty jsou vyšší než 0,8 a následně příznaky s absolutní hodnotou vyšší než 0,8 jsou odstraněny. Pro odstranění příznaků, které přesahují tento práh, se vytvoří množina indexů korelovaných příznaků. Tato množina indexů obsahuje indexy příznaků, které mají korelaci nad zvoleným prahem s alespoň jedním jiným příznakem.

Po odstranění korelovaných příznaků je vytvořena nová korelační matice pouze pro zbývající příznaky. Dále byla provedena jednoduchá analýza hlavních komponent pomocí metody PCA.



Obr. 4.3: Blokové schéma 1. experimentu.

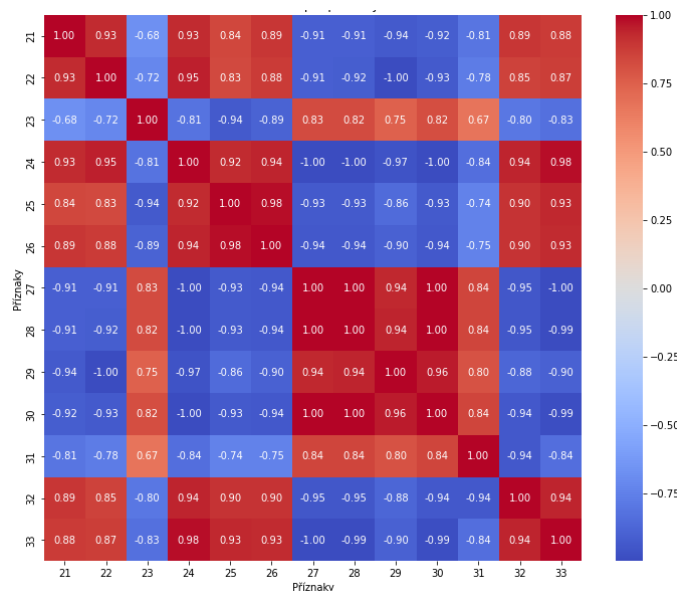
## 4.2.2 Výsledky experimentu

Mezi nejkorelovanější skupinou příznaků mezi sebou byly příznaky 27, 28, 29, 30, 31, 32, 33. (27 - GLCM - Id, 28 - GLCM - Idm, 29 - GLCM - Idmn, 30 - GLCM - Idn, 31 - GLCM - Imc1, 32- GLCM - Imc2, 33 - GLCM inverse variance) Jedná se o příznaky počítané pomocí GLCM matic. Je to logické, protože příznaky 27, 28, 29, 30 počítají tzv. homogenitu obrazu s malými obměnami jako je přidání sousedních voxelů, nebo přidáním normalizace, zatímco 31, 32, 33 počítají korelaci a rozptyl hodnot v obraze.

Další výraznou skupinou korelovaných příznaků jsou příznaky 21, 22, 24, 25, 26. (21 - GLCM - cluster tendency, 22 - GLCM - contrast, 23 - GLCM correlation, 24 - GLCM - difference average, 25 - GLCM - difference entropy, 26 - GLCM - difference variance)

Příznak 21 měří směr shlukování hodnot v obraze, zatímco příznak 22 měří rozdíl mezi intenzitami sousedních voxelů. Tyto dva příznaky by mohly korelovat, pokud má obraz tendenci mít výrazné shluky, které zároveň vykazují výrazný kontrast mezi sousedícími oblastmi. Zbylé příznaky měří různé charakteristiky rozdílů mezi hodnotami sousedních voxelů. 24 měří průměrný rozdíl mezi sousedními voxely, 25 měří entropii rozdílů mezi sousedními voxely a 26 měří rozptyl těchto rozdílů.

Zároveň výše zmíněné dvě skupiny, které byly diskutovány, měly mezi sebou vysokou hodnotu negativní korelace, na výřez z korelační matice s těmito příznaky se můžete podívat zde: 4.4.

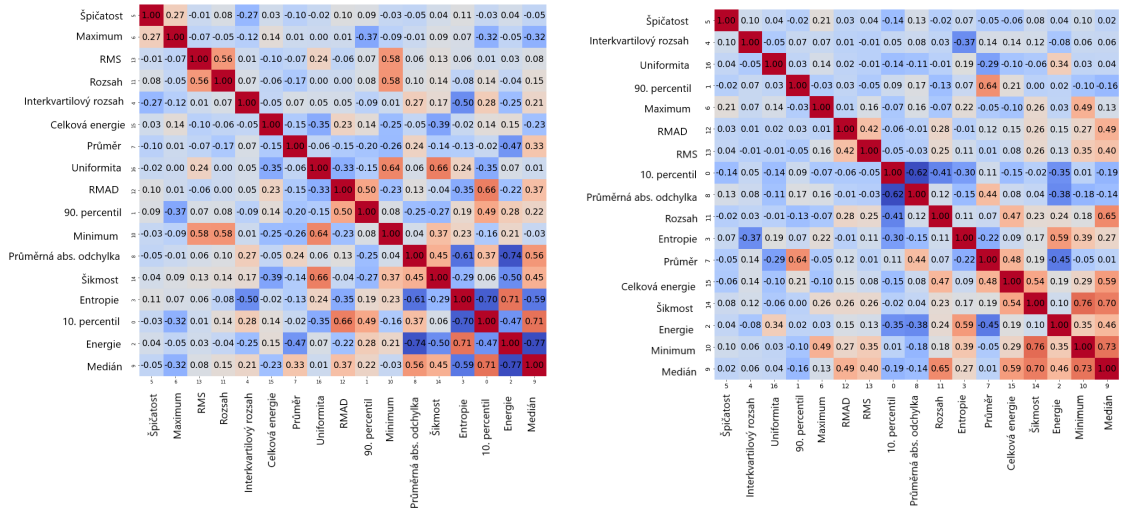


Obr. 4.4: Detail korelační matice pro příznaky č. 21-33 pro pacienta č. 112.

Na obrázku 4.5 můžeme vidět již filtrované příznaky. Počet nekorelovaných příznaků byl v průměru 20 pro trombus a 15 pro pozadí. Zajímavé je, že matice se skládá pouze z příznaků prvních statistik a GLCM matic. Dalšími příznaky, které vykazovaly významnou korelaci, byly všechny typy entropie, a to bez ohledu na to, zda byly vypočítány z původních obrazových dat nebo pomocí matic. Předpoklad, že medián a průměr budou vzájemně korelované se překvapivě nepotvrdil.

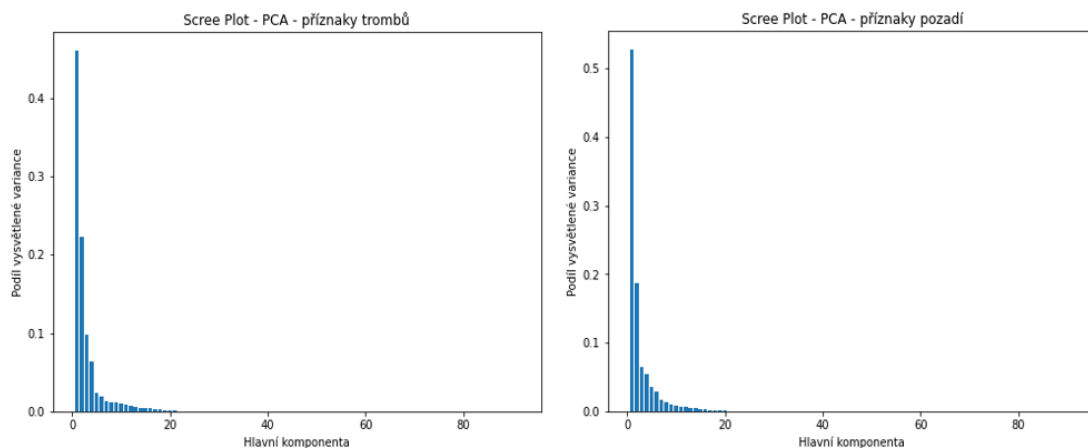
Během filtrace bylo zjištěno, že příznaky s indexem nižším než 30 měly obecně nižší míru korelace. Navíc bylo zjištěno, že mezi nekorelovanými příznaky nebyly pozorovány významné rozdíly mezi daty trombů, pozadí a pacienty.

Na obrázku 4.5 si můžeme všimnout opačné korelace 10. percentilu (0) a entropie (3) Pokud entropie a 10. percentil mají opačnou korelaci, znamená to, že v datech s vyšší entropií je pravděpodobnost, že je hodnota na 10. percentilu je nižší. Toto se stává v případech kde vysoká míra náhodnosti v datech vede k rozptýlení hodnot a tím i k nižším hodnotám na určitých percentilových úrovních.



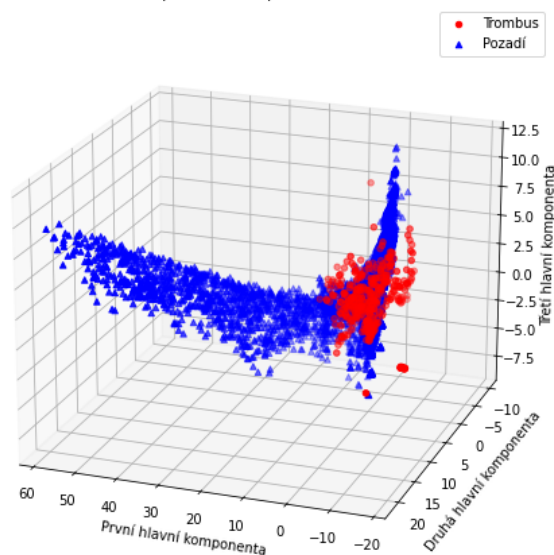
Obr. 4.5: Ukázka korelačních matic pro nejméně korelované příznaky (trombus - nahoře, pozadí - dole) pro pacienta č. 112.

Výsledkem PCA bylo zjištění, že stačí maximálně 15 hlavních komponent pro zachování více jak 95 % variability, a to jak pro data trombu, tak pro data pozadí, kde se výsledky nelišily. S menším počtem příznaků se snižuje riziko přeplnění modelu a zlepšuje se interpretace dat. I když může být obtížné interpretovat každou komponentu zvlášť, toto snížení dimenzionality může být velmi užitečné pro rychlejší a efektivnější analýzu dat.



Obr. 4.6: Ukázka Scree plotu PCA analýzy všech příznaků (trombus - vlevo, pozadí - vpravo) pro pacienta č. 112.

Vizualizace prvních tří hlavních komponent v 3D prostoru s odlišením mezi trombem a pozadím



Obr. 4.7: Vizualizace prvních tří hlavních komponent v 3D prostoru s odlišením mezi trombem a pozadím.

### 4.2.3 Diskuze a závěr

V rámci tohoto experimentu byla provedena analýza korelace mezi jednotlivými příznaky voxelů trombů a pozadí. Výsledky ukazují existenci silně korelovaných skupin příznaků, které mohou být redundantní např. příznaky odvozené z GLCM mají významnou korelaci mezi sebou.

Aplikace PCA prokázala, že umožňuje efektivní snížení dimenzionality dat a zároveň zachování významné části variability.

V tomto experimentu byl potvrzen význam provádění korelační analýzy a PCA při výběru relevantních příznaků pro případné trénování modelů strojového učení.

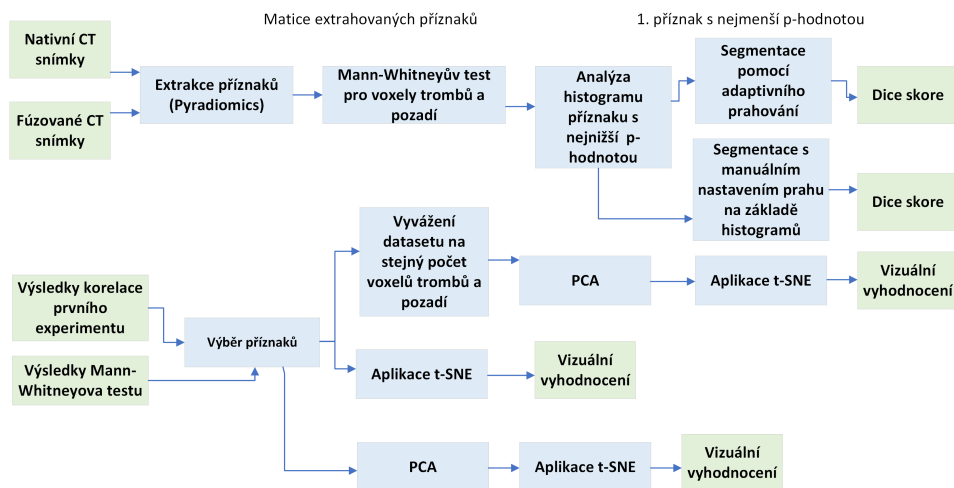
## 4.3 Experiment 2: Diskriminace voxelů trombu a pozadí v nativních a multifázických snímcích

### 4.3.1 Použitá data a jejich předzpracování

Byla použita data z celkem 8 pacientů. Stejně jako v předchozím experimentu byly použity extrahované radiomické příznaky na ořezaných datech o velikosti  $40 \times 40 \times 20$  voxelů. V tomto experimentu byly použity dva druhy dat, a to nativní a fúzované snímky.

### 4.3.2 Návrh experimentu

Vycházíme z předpokladu, že tromby se od okolní tkáně budou odlišovat. Kvantitativní hodnocení daných příznaků bylo realizováno pomocí Mann-Whitneyu testu, který porovnává voxely trombu a voxely okolních tkání. Tento statistický test určuje, zda existuje statisticky významný rozdíl mezi danými pozorováními či nikoliv. Obecným postupem je stanovení nulové hypotézy  $H_0$ , která byla stanovena jako, že trombus a okolní tkáně nevykazují statisticky významné rozdíly. Následně bylo na základě přechozího experimentu vybrána kombinace více příznaků a byla provedena analýza pomocí metody t-SNE.



Obr. 4.8: Blokové schéma 2. experimentu.

## Stanovení hypotéz H0 a H1 pro Mann-Whitneyův test

Stanovená hypotéza H0: Neexistuje žádný rozdíl mezi hodnotou příznaků u skupiny trombů a okolních tkání. Alternativní hypotéza H1: Existuje rozdíl mezi hodnotou příznaků u skupiny trombů a okolních tkání.

Pro testování hypotéz je klíčové stanovení p-hodnoty. V našem případě byla zvolena hladina 0,05, což znamená, že interval spolehlivosti tohoto testu je 95 %. Jinými slovy, 95 % všech rozdílů mezi mediány vzorků trombu a okolních tkání budou obsaženy v tomto intervalu.

Interval spolehlivosti je rozmezí hodnot, ve kterém můžeme s danou pravděpodobností považovat naši hypotézu za platnou. Nastavení p-hodnoty na hodnotu např.  $p = 0,02$  by vedlo k relativně přesnějším výsledkům Mann-Whitney testu, ale zároveň by se zvětšil rozsah konfidenčního intervalu. Proto je při volbě hladiny významnosti nutné hledat kompromis mezi přesností výsledků a šířkou intervalu.

Příznaky byly seřazeny na základě p-hodnoty, od nejmenší k největším a byla provedena analýza příznaků u všech pacientů.

## Výběr příznaků a jejich analýza

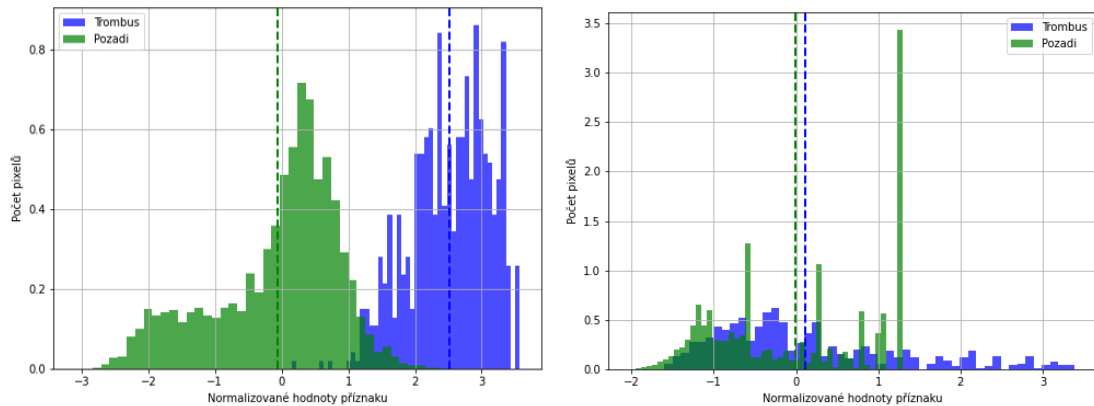
Na základě předchozího experimentu byly analyzovány histogramy nejvýznamnějších příznaků a byla provedena jednoduchá segmentace trombu pro objektivní zhodnocení výsledků. Celý postup experimentu je znázorněn v blokovém schématu 4.15. Dále bylo vybráno deset příznaků (0, 1, 5, 6, 8, 10, 12, 16, 17, 33) a následně byla pro vizualizaci aplikována metoda t-SNE s nastavenými parametry perplexity 40 a počtu iterací 1000, které v průběhu celého experimentu zůstaly neměnné. Pro ještě lepší vizualizaci dat, byla aplikována na příznakovou matici metoda PCA. Dalším krokem bylo rozšíření příznakové sady o dalších 6 příznaků (0, 1, 5, 6, 8, 10, 12, 16, 17, 33, 49, 47, 59, 60, 61, 68). Celý postup byl proveden jak na nativních, tak na fúzovaných snímcích a výsledky byly porovnány.

### 4.3.3 Výsledky experimentu

#### Diskriminační schopnost v nativních snímcích

Z výsledků statistického testu je zřejmé, že pro valnou většinu příznaků u většiny pacientů je možné zamítnout nulovou hypotézu a přijmout hypotézu alternativní H1: Existuje rozdíl mezi hodnotou příznaků u skupiny trombů a okolních tkání. Z těchto výsledků by se mohlo na první pohled zdát, že pro spolehlivou klasifikaci, by bylo možné použít kterýkoliv ze získaných příznaků. Z hodnot příznaků byly vytvořeny normalizované histogramy. Normalizace byla nutná, neboť počet voxelů okolní tkáně byl mnohonásobně vyšší než počet voxelů trombu. Jak je vidět z obr. 4.9

histogramy voxelů trombů a okolních tkání se v některých případech překrývají, z čehož je patrné, že všechny příznaky pro klasifikaci a segmentaci trombu nelze použít. Dalším zajímavým faktem je, že voxely trombu mají tendenci mít vyšší hodnoty než okolní tkáň, což odpovídá předpokladu hyperdenszity v nativních snímcích. Po vypočtení p-hodnoty pro všechny příznaky všech pacientů, byly příznaky seřazeny podle vzestupné tendence od příznaků s nejmenší p-hodnotou po tu největší.



Obr. 4.9: Ukázka histogramů příznak s nejnižší (č. 1 - vlevo) a nejvyšší (č. 47 - vpravo) p-hodnotou pro pacienta č. 394 s vyznačenou průměrnou hodnotou.

Ač příznak č. 47 A.1 vyšel ve statistice jako významný, histogramy hodnot obou skupin vykazují výrazné překrývání což můžeme vidět na obr 4.9, to znamená, že rozdíly mezi skupinami mohou být menší, než by se mohlo zdát na základě statistického testu. Tyto odchylky mohly být způsobeny případným nedostatečným odstraněním odlehlých hodnot, i přesto, že pro Mann-Whitneyův test je typická jistá robustnost vůči odlehlým hodnotám.

Pro již zmíněný příznak č.1 se tedy nabízí otestování našeho předpokladu pomocí objektivního hodnocení, a tak byla provedena segmentace pomocí globálního a adaptivního prahování pomocí tohoto příznaku pro všechny pacienty. Tento příznak byl demonstrativně vybrán především kvůli tomu, že měl tendenci být v 10. příznacích s nejmenší p-hodnotou pro sedm z osmi pacientů, jiné významné příznaky se u každého pacienta více či méně odlišovaly. Nejdříve bylo provedeno globální prahování, kdy na základě dat z histogramů, byl nastaven pevný práh. Tento práh byl na základě zobrazení pomocí histogramu nastaven pro každého pacienta individuálně. Dále bylo provedeno adaptivní prahování. Navrhovaná metoda využívá statistické vlastnosti příznaků (konkrétně jednoho příznaku) a kombinuje průměr a standardní odchylku hodnot příznaků v dané oblasti obrazu. Jak je patrné z tabulky, nelze zcela jednoznačně posoudit, která metoda prahování je výrazně lepší, protože u dvou pacientů obě metody selhávaly. Mírně lepších výsledků však dosahuje metoda globálního prahování. V případě pacienta č. 400 je to pravděpodobně způsobeno tím, že tento

Tab. 4.1: Tabulka s hodnotami DICE skóre pro jednotlivé pacienty.

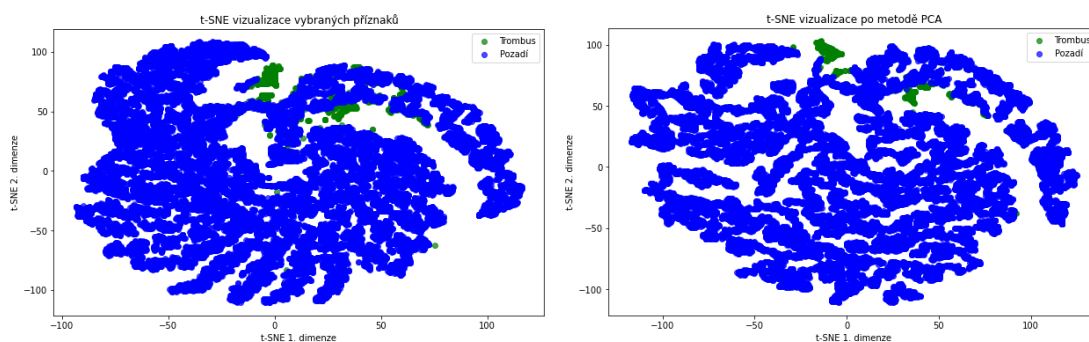
Pacient	Globální prahování Dice [%]	Adaptivní prahování Dice [%]
112	0,50	0,75
238	0,19	0,00
243	0,70	0,64
297	0,53	0,39
319	0,76	0,71
365	0,63	0,02
394	0,83	0,69
400	0,14	0,00
Průměrná hodnota	0,54	0,40

příznak je pro tohoto pacienta méně statisticky významný, a jako u jediného pacienta se příznak č. 1 neobjevil mezi 10 příznaky s nejmenší p-hodnotou. To však neplatí pro pacienta č. 238, kdy se tento příznak objevuje jako příznak s nejmenší p-hodnotou. Naopak u pacientů č. 394, 319 a 112 dokázaly kombinace buď globálního, nebo adaptivního prahování, rozlišit tkáň a trombus celkem spolehlivě. Toto je však pouze prvotní analýza, kdy bylo ověřeno, že extrahované příznaky mají schopnost v jisté míře tkáň od trombu rozlišit. Výsledky pomocí globálního a adaptivního prahování jednoho příznaku nebylo vždy úplně spolehlivé, proto bylo vybráno nejdříve 10 a poté 16 nejlepších příznaků, vybraných tak, aby reprezentovaly spolehlivé příznaky s nejnižší p-hodnotou a zároveň nebyly redundantní pro většinu pacientů.

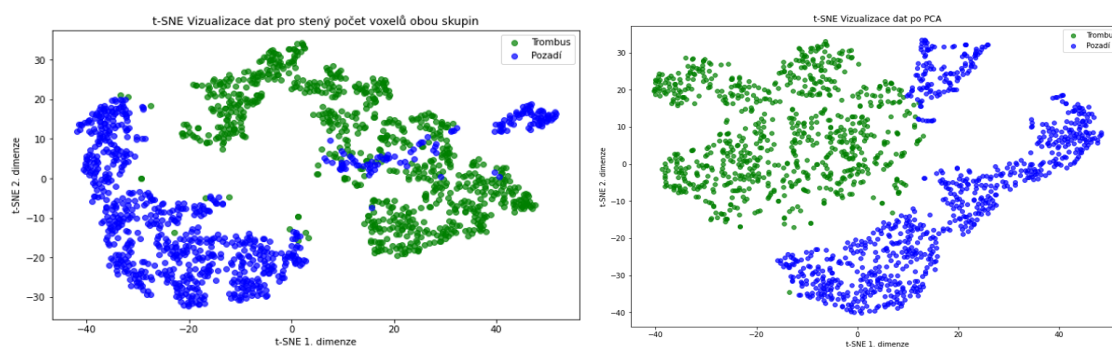
Na obrázku 4.10 můžeme vidět ukázkou vizualizace dat pomocí metody t-SNE. V případě vybraných příznaků, je zajímavé, že statistický test vyhodnotil jako nejvýznamnější příznaky počítané pomocí jednoduchých statistik. Jediný příznak, který nepochází ze skupiny jednoduchých statistik je 35, jedná se o tzv. "GLCM joint average".

Nejprve byly vizualizovány pomocí t-SNE všechny voxely z celého snímku. Vzhledem k tomu, že skupina voxelů představující pozadí byla výrazně početnější, t-SNE nedokázala efektivně rozlišit mezi těmito dvěma skupinami. I přesto, že je patrné, že trombus má tendenci tvořit shluky, t-SNE nedokázalo tuto vlastnost účinně zachytit. Dokonce ani redukce dimenzionality pomocí PCA nepřinesla požadovaný efekt v odlišení těchto skupin. Je tak otázka, zda není nezbytné mít vyvážené datasey, aby tyto metody dosáhly optimálních výsledků.

Následně byla provedena regularizace a byly vybrány vždy všechny voxely trombů spolu se stejným počtem náhodně vybraných voxelů pozadí. Tato strategie přinesla výrazně lepší výsledky, jak v původní dimenzionalitě, tak po redukcii dimenziona-

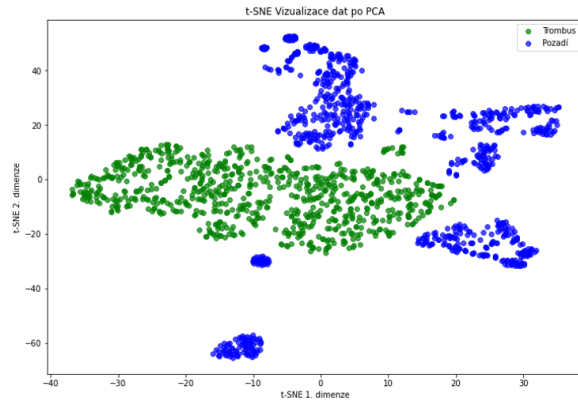


Obr. 4.10: Ukázka t-SNE vizualizace vybraných příznaků pacienta č. 400 bez redukce dimenzionality (vlevo) a s metodou PCA (vpravo).



Obr. 4.11: Ukázka t-SNE vizualizace vybraných příznaků pacienta č. 400 bez redukce dimenzionality (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.

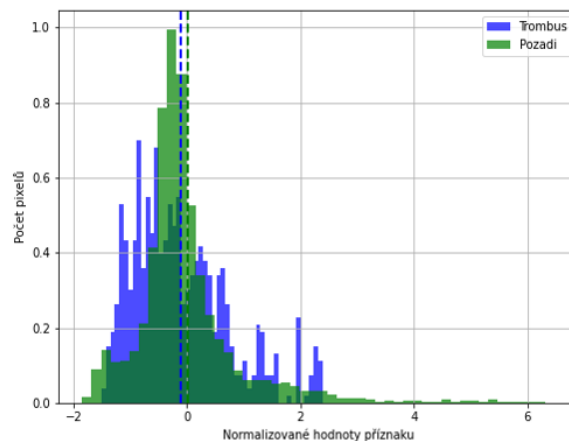
lity pomocí PCA. Jak lze pozorovat na následujícím obrázku 4.11, skupiny voxelů trombů již vykazují tendenci vytvářet shluky. Pro ověření stability tohoto postupu byla tato metoda opakována pětkrát, přičemž bylo zkoumáno, zda se výsledky mění při volbě různých náhodných voxelů pozadí. Závěrem bylo zjištěno, že i když jsou výsledky mírně variabilní, jejich odlišnosti nejsou příliš výrazné a v každém případě je patrná tendence k tvorbě shluků. Jak je vidět v příloze A.2, u čtyř pacientů se stávající sada příznaků ukázala jako nedostatečná, což vedlo k rozhodnutí rozšířit příznakový prostor o dalších 6 příznaků. V tomto kroku byly zahrnuty i příznaky z oblasti co-okurenčních matic. Na obrázku 4.12 jsou prezentovány výsledky po přidání těchto šesti nových příznaků. Tato modifikace výrazně přispěla k lepší identifikaci, jak lze vidět v příloze A.3, kde tato skupina příznaků již umožňuje úspěšně rozlišit obě skupiny.



Obr. 4.12: Ukázka t-SNE vizualizace 15 vybraných příznaků pacienta č. 400 s metodou PCA.

### Diskriminační schopnost ve fúzovaných snímcích snímcích

Dále byly podrobeny analýze i fúzované obrazy. Postup byl naprosto stejný jako u nativních snímků, tedy prvním krokem byl výpočet statistických testů. Poté byla provedena analýza histogramů nejvýznamnějších příznaků a poté vizualizace pomocí t-SNE. Výsledkem statistických testů, kdy byla zkoumána možná diskriminace trombů od okolních tkání, bylo u většiny příznaků zamítnutí  $H_0$ , tedy že statistické rozdíly jsou významné. Pokud se však podíváme na histogramy statisticky nejsilnějších příznaků viz 4.13, můžeme pozorovat jejich výrazné překrytí. Tento trend byl patrný u všech pacientů.



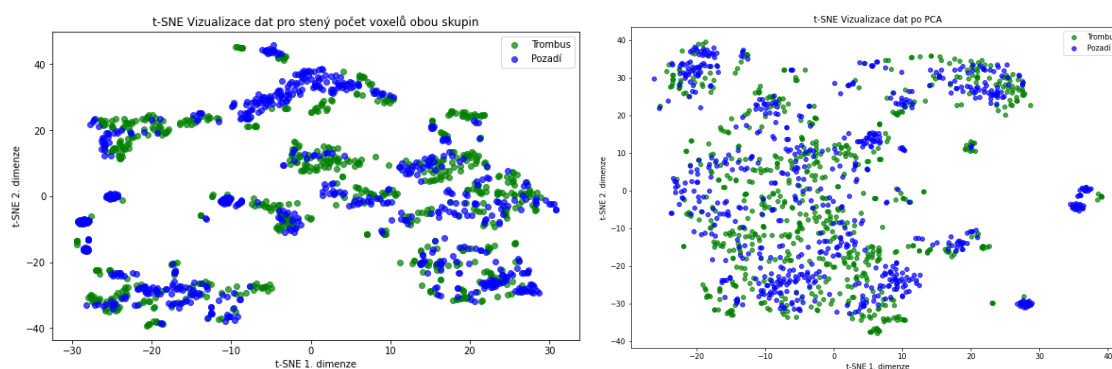
Obr. 4.13: Ukázka histogramu příznaku s nejnižší p-hodnotou pro pacienta č. 394 s vyznačenou průměrnou hodnotou.

Dále bylo provedeno ověření našeho předpokladu, tedy, že není možné diskriminovat tromby od pozadí u fúzovaných obrazů na jednoduché segmentaci u nejsilnějšího příznaku. Jak je patrné z tabulky, ani segmentace na základě histogramů, nebo

adaptivního prahování není schopna tromby od pozadí spolehlivě rozlišit.

Pacient	Globální prahování Dice [%]	Adaptivní prahování Dice [%]
112	0,07	0,03
238	0,02	0,00
243	0,06	0,29
297	0,03	0,06
319	0,15	0,07
365	0,02	0,04
394	0,03	0,03
400	0,00	0,00
<b>Průměrná hodnota</b>	<b>0.05</b>	<b>0.03</b>

Dále byla provedena vizualizace pomocí t-SNE, kdy byly použity stejné příznaky jako u nativních snímků. Jak je patrné na obrázku 4.14 tak fúzovaná data nemají tendenci tvořit výrazné shluky.



Obr. 4.14: Ukázka t-SNE vizualizace vybraných příznaků pacienta č. 394 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.

### 4.3.4 Diskuze a závěr

Z výsledků tohoto experimentu vyplývá, že při analýze diskriminační schopnosti příznaků v nativních snímcích je možné u většiny pacientů zamítnout nulovou hypotézu a přijmout hypotézu alternativní, že existuje rozdíl mezi hodnotami příznaků u skupiny trombů a okolních tkání. I přesto, že výsledky naznačují možnost spolehlivé klasifikace pomocí jednotlivých příznaků, analýza histogramů ukazuje, že ne všechny příznaky jsou vhodné pro segmentaci trombu. Zejména je zřejmé, že histogramy hodnot voxelů trombu a okolních tkání se v některých případech výrazně překrývají.

Přestože některé příznaky vyšly jako statisticky významné, výsledky vizualizace t-SNE naznačují, že rozdíly mezi skupinami mohou být menší, než by se mohlo zdát na základě statistických testů. Prvním krokem k ověření tohoto předpokladu bylo použití globálního a adaptivního prahování na příznaky. Výsledky těchto metod však nebyly vždy úplně spolehlivé, což vedlo k rozšíření příznakového prostoru o další relevantní příznaky.

Při analýze fúzovaných snímků bylo zjištěno, že i když statistické testy ukazují na možné rozdíly mezi tromby a okolními tkáněmi, analýza histogramů a vizualizace t-SNE dokazují, že fúzované obrazy nejsou vhodné pro spolehlivou segmentaci trombů od pozadí.

## 4.4 Experiment 3: Segmentace trombů pomocí metod strojového učení

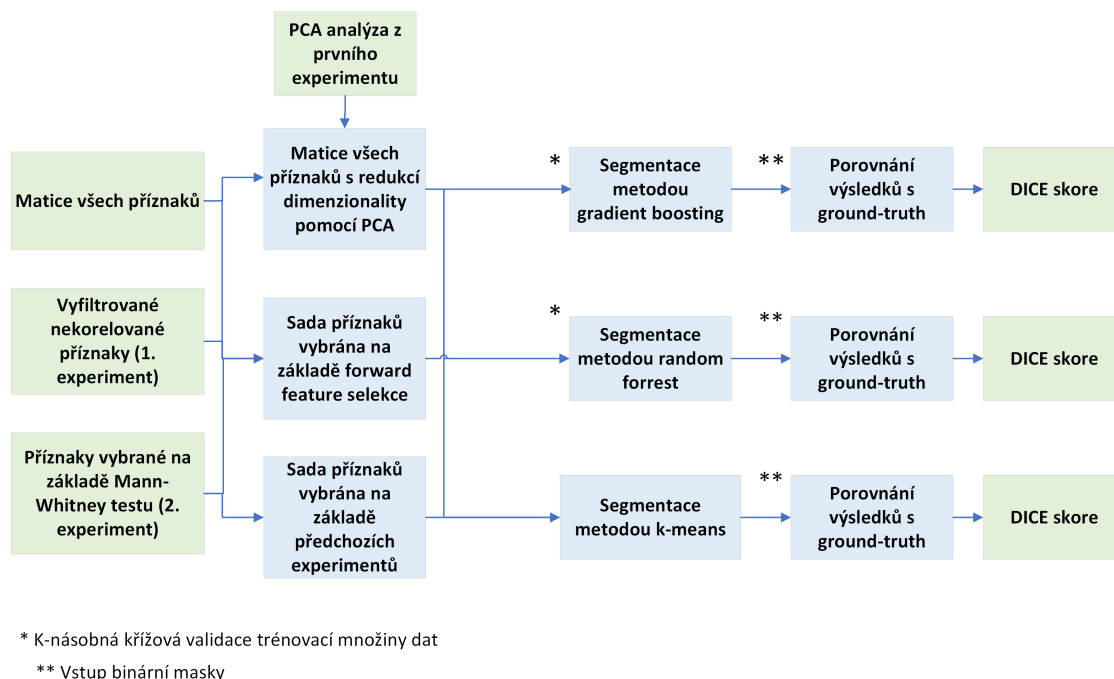
### 4.4.1 Data a návrh experimentu

Byla použita data z celkem osmi pacientů, která byla později rozšířena o dalších pět. Pro trénování modelů byly využity již předzpracované a analyzované příznaky z předchozích experimentů. V tomto experimentu byla použita data pouze z nativních snímků. Byl proveden experiment s různými sadami příznaků a velikostmi trénovacích dat.

Celkem byly implementovány tři metody segmentace: "K-means", Random Forest a Gradient Boosting. Experiment byl prováděn na třech sadách příznaků, které byly vybrány na základě předchozích analýz, jako je korelační analýza a diferenciací analýza trombů a pozadí pomocí Mann-Whitneyova testu ve spojení s metodou forward feature selection.

První dataset obsahoval příznakový vektor vybraný pouze metodou feature selection s 4 příznaky (0, 1, 2, 6). Začalo se s jedním příznakem, a další byly postupně

přidávány. Úspěšnost segmentace byla vyhodnocena pomocí DICE skóre. Druhý dataset byl rozšířen o dalších 6 příznaků (0, 1, 5, 6, 8, 10, 12, 16, 17, 33). Experimentální přidání dalších příznaků vedlo ke zhoršení výsledků modelu. Ve třetím datasetu bylo použito všech 91 příznaků, a to včetně metody PCA s 15 hlavními komponentami.



Obr. 4.15: Blokové schéma 3. experimentu.

### Aplikace metody „k-means“

Algoritmus "K-means" byl implementován pomocí knihovny scikit-learn a použit pro segmentaci trombů. Kód byl navržen tak, aby načítal nativní příznakový vektor pacienta a jejich binární masky.

Počet shluků byl nejdříve nastaven na základě "Elbow metody", která slouží k výběru optimálního počtu shluků viz. ukázka na obrázku: 4.16. Tento graf zobrazuje hodnoty setrvačnosti, v závislosti na počtu shluků. Typicky má tvar podobný lokti, a to z důvodu, že s rostoucím počtem shluků setrvačnost klesá, avšak po určitém bodě již klesá méně výrazně, což vytváří "loket" na grafu. Číslo, kde se tento "loket" vytvoří, je potom často považován za optimální počet shluků, protože přidání dalších shluků již nepřináší značné zlepšení v setrvačnosti. Počet shluků je nastaven na hodnotu 7, kdy předpokladem je, že 1 shluk bude obsahovat voxely trombu a zbylých 6 voxely pozadí. Výběr shluku odpovídající trombům byl proveden na základě výpočtu největší intenzity, protože po výsledném shlukování není předem známo,

který shluk patří, které třídě. Po provedení segmentace jsou vyhodnoceny výsledky pomocí konfuzní matice a DICE koeficientu.

Nakonec jsou aplikovány morfologické operace eroze a dilatace na finální segmentovaný obraz za účelem jemného vylepšení výsledků.

### **Aplikace metody „random forrest“**

K segmentaci trombů byla použita metoda Random Forest s následným postprocesingem. Algoritmus byl trénován na dvou různých trénovacích množinách, přičemž v prvním případě bylo použito 5 pacientů pro trénink a 8 pacientů pro testování, zatímco ve druhém případě bylo použito 8 pacientů pro trénink a 5 pacientů pro testování, kdy byl sledován vliv velikosti trénovací množiny na výslednou segmentaci. Byly použity dostupné knihovny a již implementovaná funkce RandomForestClassifier. Nejdříve byla provedeno rozdělení dat na trénovací, testovací množiny. Trénovací množina byla dále rozdělena pomocí třínásobné křížové validace na trénovací a validační dataset. Úspěšnost modelu byla hodnocena pomocí Dice skóre. Vybrané příznaky byly stejné jako u analýzy "K-means". V modelu byl počet stromů nastaven na 100. Nastavení randomstate=42 zajišťuje reprodukovatelnost výsledků, což znamená, že při opakovaném spuštění kódu budou výsledky stejné. K minimalizaci impurity v uzlech je použita "giny impurity". Všechny parametry modelu byly nastavené pomocí metody grid search. Posledním pokusem bylo provedení PCA na všech 91 příznaků s počtem hlavních komponent nastavených na 15 viz. první experiment. Nakonec jsou aplikovány morfologické operace eroze a dilatace na finální segmentovaný obraz za účelem jemného vylepšení výsledků. Nakonec je znovu vypočten DICE koeficient.

### **Aplikace metody „gradient boosting“**

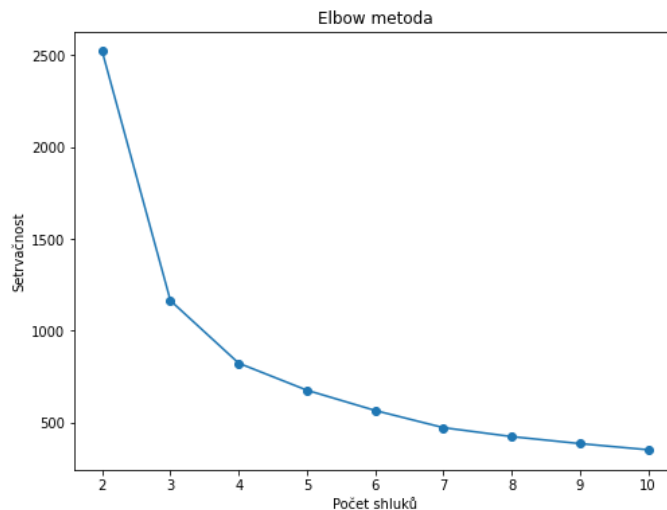
Poslední použitou metodou bylo "Gradient boosting". Gradient Boosting je algoritmus strojového učení, který umožňuje vytvořit silný model z jednoduchých prediktorů. Jeho schopnost zachytit komplexní nelineární vztahy v datech a odolnost vůči přeučení ho činí populární volbou pro řešení problémů klasifikace a regrese. Samotný Gradient Boosting je algoritmus ensemble learningu, který kombinuje více slabých modelů (v tomto případě rozhodovací stromy, jejichž počet byl nastaven na 100) za účelem vytvoření silného modelu. Při každé iteraci se nový model přidává do sestaveného ensemble tak, aby se zaměřil na nedostatky předchozích modelů. Trénovací dataset byl už byl rozdělen pouze na osm pacientů pro trénování a pět pro testování. Trénovací množina byla dále rozdělena třínásobnou křížovou validací na trénovací a validační množinu. Poté je inicializován a natrénován model Gradient Boosting,

Jako ztrátová funkce byla použita "LogitBoost". Následuje trénování Gradient Boosting modelu, na stejných příznakových sadách jako v předchozích metodách. V závěrečné fázi je provedena vizualizace výsledků a vyhodnocení pomocí DICE skóre.

## 4.4.2 Výsledky a diskuze

### Výsledky metody k-means

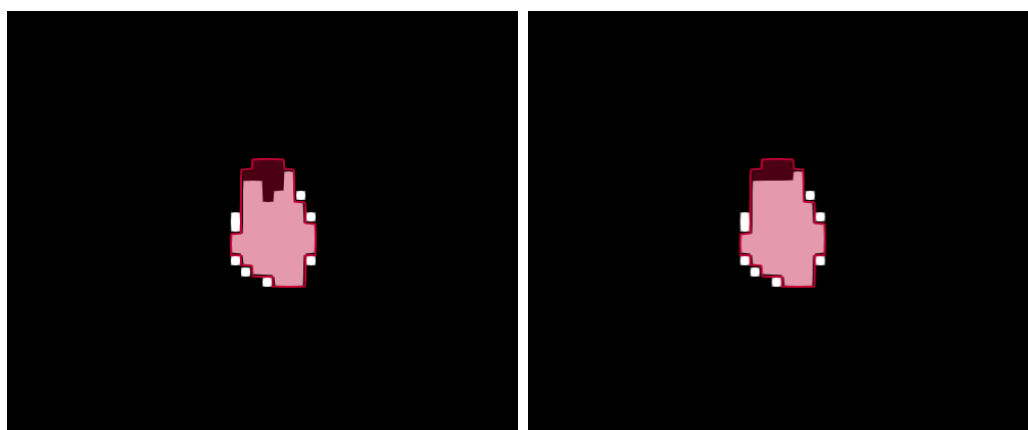
Limitací metody "k-means" může být výběr shluku na základě nejvyšší intenzity, protože vzhledem k nehomogenosti hmoty trombu nemusí být vybraný shluk optimální. K tomu pravděpodobně došlo u pěti pacientů kdy se hodnota DICE pohybovala okolo 0. Zbylé pacienty metoda segmentovala velmi dobře, protože se jejich DICE skóre bez postprocessingu pohybovalo okolo 0,70. Bohužel selhání u výše zmíněných pacientů výrazně snižuje celkovou přesnost segmentace. Výsledky jsou zaznamenány v tabulce 4.2. Výsledek posprocessingu v průměru vždy vylepšil výslednou segmentaci, zato navýšení počtu příznaků segmentaci ztlačilo. Úplně nejhorších výsledků bylo dosaženo po aplikaci PCA, kdy byly použity všechny příznaky a počet hlavních komponent byl nastaven na 15. Další aspekt, který ovlivňoval segmentaci, bylo nastavení počtu shluků. Metoda byla testována s různými nastaveními a nejlepších výsledků bylo dosaženo právě se sedmi shluky. Zajímavým faktem také bylo, že pokud trombus nebylo možné segmentovat a DICE bylo okolo 0, ani různá nastavení nebyla schopna segmentaci vylepšit. Výsledek segmentace a význam posprocessingu je znázorněn na obrázku: 4.17.



Obr. 4.16: Ukázka výběrů počtu shluků pomocí "Elbow metody".

Tab. 4.2: Hodnocení úspěšnosti metody "K-means" a průměrné DICE skóre.

Příznaky	Bez "post-processingu" Dice [%]	S "postprocessingem" Dice [%]
První sada	0,37	0,40
Druhá sada	0,29	0,31
PCA	0,11	0,13



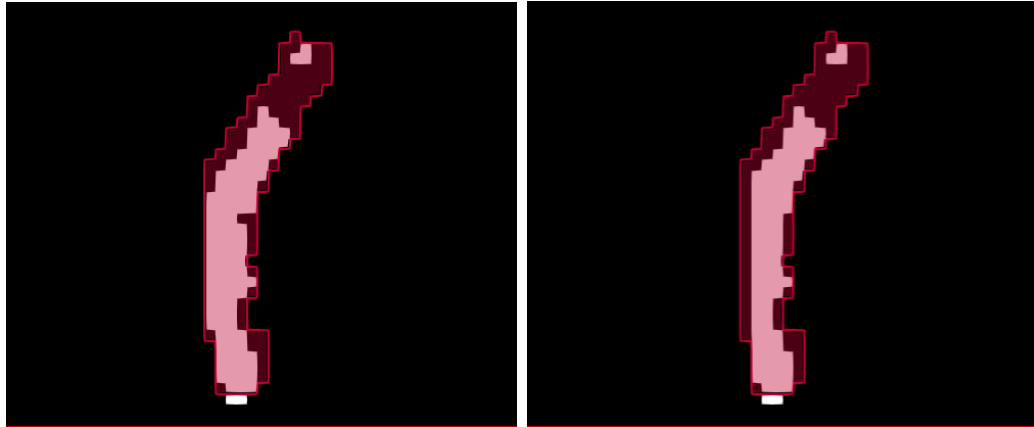
Obr. 4.17: Segmentace pacienta č. 243 metodou Kmeans. Obrázek vlevo - bez post procesingu (bíle označen segmentovaný trombus, červeně maska), vpravo po aplikaci "posprocesingu".

### Výsledky metody random forrest

Dataset byl rozdělen celkem na tři příznakové sady, kdy příznaková matice se shodovala s příznaky použitými u metody "k-means". Dále byl proveden pokus s vlivem velikosti trénovací a testovací sady na úspěšnost segmentace. Aby nedošlo k promíchání dat, bylo 8 snímků pacientů použito pouze pro testování. Zbýlých pět snímků bylo dále rozděleno na trénovací a validační dataset kdy byla provedena třinásobná křížová validace.

Tab. 4.3: Hodnocení úspěšnosti metody "Random forrest" a průměrné DICE skóre.

Příznaky	Bez "post-processingu" Dice [%]	S "postprocessingem" Dice [%]	Validace Dice [%]
První sada	0,49	0,51	0,55
Druhá sada	0,50	0,50	0,56
PCA	0,43	0,44	0,58



Obr. 4.18: Segmentace pacienta č. 365 metodou Random forrest. Obrázek vlevo - bez post procesingu (bíle označen segmentovaný trombus, červeně maska), vpravo po aplikaci "posprocessingu". První sada příznaků.

Pro první sadu příznaků bylo dosaženo Dice skóre bez post-procesingu 0,49 a 0,51 po aplikaci post-procesingu. Pro druhou sadu příznaků bylo dosaženo skóre 0,50. PCA přístup k redukci dimenzionality měl úspěšnost nejnižší a ani posprocessing výsledky výrazně nezlepšil.

Tab. 4.4: Hodnocení úspěšnosti metody "Random forrest" a průměrné DICE skóre. Zvýšená trénovací množina 8 trenovací a 5 test

Příznaky	Bez "post-processingu" Dice [%]	S "postprocessingem" Dice [%]	Validace Dice [%]
První sada	0,57	0,47	0,53
Druhá sada	0,59	0,59	0,55
PCA	0,55	0,57	0,53

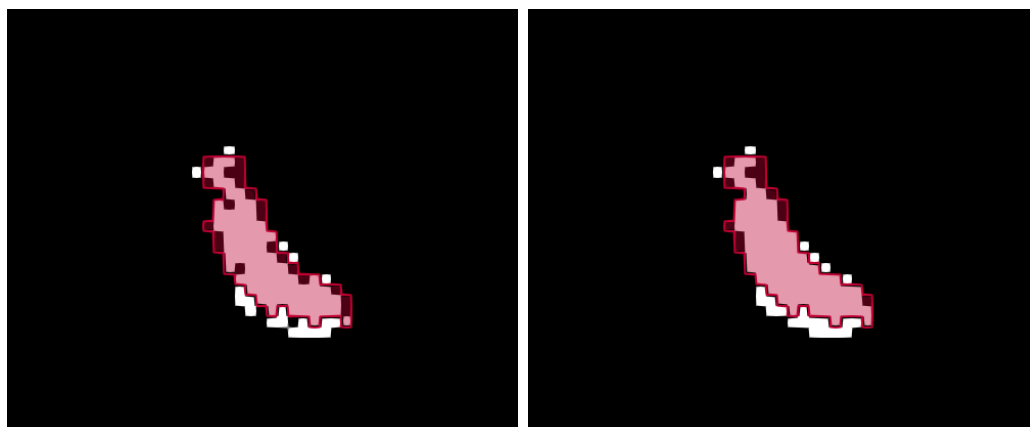
V druhém experimentu byla zvýšena trénovací množina na 8 a testovací na 5 pacientů. Trénovací množina byla znovu rozdělena třinásobnou křížovou validací na trénovací a validační množinu dat.

Pro první sadu příznaků bylo dosaženo Dice skóre bez post-procesingu ve výši 0,57 a 0,47 po aplikaci post-procesingu. V tomto případě posprocessing průměrně výrazně zhoršil výsledky. Pro druhou sadu příznaků bylo dosaženo skóre 0,59. PCA přístup k redukci dimenzionality měl Dice skóre 0,55 bez post-procesingu a 0,57 s ním.

Zvětšení trénovací množiny vedlo k vylepšení výsledků pro obě sady příznaků, s výraznějším zlepšením pro druhou sadu, která dosáhla nejvyššího skóre v obou experimentech. Post-processing způsobil mírné snížení úspěšnosti segmentace pro

první sadu, avšak zlepšil výsledky pro sadu s PCA. PCA redukce dimenzionality sice nepřinesla výrazné zlepšení výsledků, ale stále poskytla konkurenceschopná skóre.

Příčinou malé průměrné úspěšnosti post-processingu může být skutečnost, že použité morfologické operace upravovaly vždy pouze jeden voxel. Pokud by byl post-processing upraven individuálně pro každého pacienta, bylo by možné dosáhnout lepších výsledků. Avšak každý trombus byl segmentován trochu jinak, a proto se zdála být nejvhodnější metodou eroze o jeden voxel a následná dilatace o jeden voxel.



Obr. 4.19: Segmentace pacienta č. 31 metodou Random forrest. Obrázek vlevo - bez post procesingu (bíle označen segmentovaný trombus, červeně maska), vpravo po aplikaci "posprocesingu". Druhá datová sada.

### Výsledky metody "Gradient Boosting"

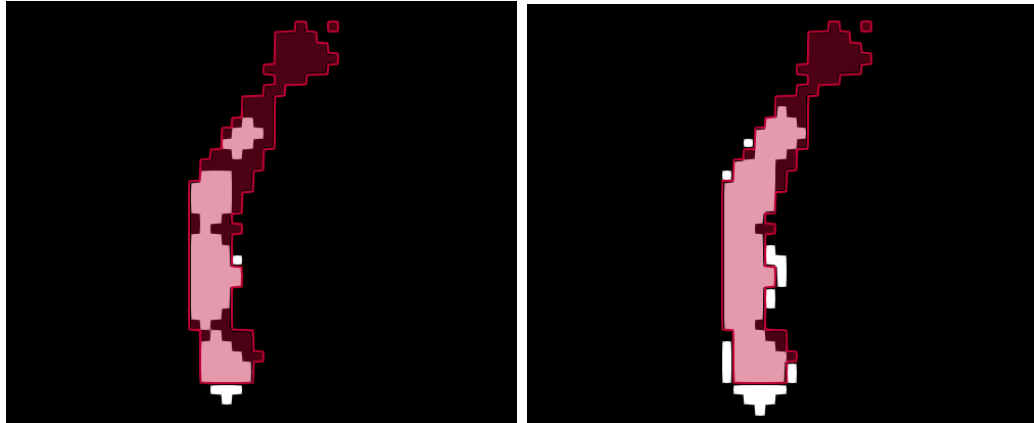
U metody gradient boosting už byla rozdělena trénovací a testovací sada na 8 pacientů pro trénování a 5 pro testování, kdy trénovací dataset byl rozdělen tří násobnou křížovou validací.

V prvním experimentu s menší příznakovou množinou, obě sady příznaků dosáhly totožných výsledků. První sada příznaků dosáhla Dice skóre 0,47 bez post-procesingu a 0,48 s ním. Použití PCA přístupu vedlo k výraznému zlepšení úspěšnosti s Dice skóre 0,61 a postprocessing vylepšil skóre o jednu setinu.

Gradient Boosting přinesl lepší výsledky porovnání s Random Forest, přičemž PCA redukce dimenzionality měla výrazný vliv na zlepšení úspěšnosti segmentace, i když za cenu nižší validace.

Tab. 4.5: Hodnocení úspěšnosti metody "Gradient Boosting" a průměrné DICE skóre. Menší trnovaci množina

Příznaky	Bez "post-processingu" Dice [%]	S "postprocessingem" Dice [%]	Validace Dice [%]
První sada	0,47	0,48	0,53
Druhá sada	0,47	0,48	0,53
PCA	0,61	0,62	0,37



Obr. 4.20: Segmentace pacienta č. 319 metodou Gradient boosting. Obrázek vlevo - bez post procesingu (bíle označen segmentovaný trombus, červeně maska), vpravo po aplikaci "posprocesingu". První datová sada.

### 4.4.3 Diskuze a závěr

Příznaky pro jednotlivé metody byly vybrány na základě analýzy příznaků a dále pomocí forward feature selection, kdy byly obě metody zkombinovány pro nalezení vhodných příznaků. Výše uvedené pokusy byly provedeny i na fúzovaných obrazech, tam však docházelo spíše k segmentaci cév, než trombů a při segmentaci trombů metody selhávaly.

Metoda k-means dosahovala nižší úspěšnosti ve srovnání s Random Forest a Gradient Boosting. Hlavní problém může spočívat v tom, že výběr shluku na základě nejvyšší intenzity nemusel být optimální pro nehomogenní tromby. Navíc, výsledky této metody byly citlivé na počet shluků a počet použitých příznaků.

Random Forest a Gradient Boosting dosáhly konkurenceschopných výsledků. PCA redukce dimenzionality výrazně zlepšila výsledky Gradient Boostingu, ale vedla ke snížení validace. Z výsledků lze vyvodit, že metoda Gradient Boosting ve spojení s PCA redukcí dimenzionality a vhodným post-processingem poskytuje nejlepší segmentaci trombu v obrazech z CT. Post-processing je klíčový pro zlepšení kvality segmentace a zmenšení chyb.

Důvodem proč žádná z metod nedosáhla vyšších skóre, není dán nepřesnou lokalizací trombu jako takového, ale tím, že ani jeden z modelů nebyl schopen detekovat okrajové voxely trombu.

Segmentace trombu v obrazech z CT jsou náročným úkolem kvůli nehomogenitě tkáně a různým faktorům ovlivňujícím kontrast a intenzitu obrazu. I přes to jsou výsledky těchto metod srovnatelné s manuální detekcí lidských expertů, jejichž průměrné DICE koeficienty se pohybovaly v rozmezí 0,39 - 0,61. Rozdíly v segmentaci měli i stejní experti mezi svými opakovanými pokusy [48]. Je tedy zřejmé, že manuální segmentace trombů je náročným úkolem a výzvou, kvůli přirozeným charakteristikám trombu jako je jeho hustota a průchodnost sraženiny, ovlivňují jeho viditelnost jak v nativních, tak i fúzovaných skenech [48].

Další výzkum by se mohl zaměřit na další optimalizaci sady příznaků a metod segmentace, stejně jako na rozšíření studie na větší soubory dat a různé typy trombóz. Dalším možným navázáním této diplomové práce by mohlo být použití konvolučních sítí.

# Závěr

Tato diplomová práce se zabývá problematikou CMP. V teoretické části je stručně popsána anatomie mozku a cév. Další část se zaměřuje na rešerši mechanismu vzniku CMP, metody její léčby a následně shrnuje závěry z vědeckých článků, které se týkají histologické analýzy trombů.

V další kapitole práce jsou rozebrány postupy z několika vědeckých článků, které prezentují aktuální metody pro analýzu trombů z obrazových dat. Závěry ukazují, že segmentace a kvalifikace trombů pomocí strojového učení je teoreticky možná. Nicméně, omezením jsou malé dostupné datasety a nedostatek anotací.

V rámci teoretického přehledu jsou stručně shrnuty metody zobrazování, konkrétně pomocí MRI a CT.

Dále jsou shrnuty metody extrakce příznaků a dalších statistických postupů, které byly využity v praktické části práce. To zahrnuje nejen jednoduché statistické metody, ale i pokročilejší přístupy, jako je extrakce příznaků pomocí Pyradiomics.

Praktická část práce se skládá ze tří experimentů. V první části byla provedena extrakce příznaků z CT dat pacientů s CMP. První experiment se zabývá korelační analýzou extrahovaných příznaků, rozdílů mezi pacienty a následnou PCA analýzou pro stanovení vhodného počtu hlavních komponent. Na základě vypočtené korelační matice jsou filtrovány příznaky s hodnotou vyšší než 0,8. Druhý experiment se zabývá diferenciačními vlastnostmi trombu a pozadí pomocí Mann-Whitneyova testu, který potvrdil odlišnost mezi příznaky trombů a okolních tkání. Dále bylo provedeno segmentování trombu s využitím adaptivního prahování a prahování s tvrdým prahem, přičemž byl použit jediný příznak, který se jevil jako nejsilnější podle korelační analýzy z prvního experimentu a Mann-Whitneyova testu.

Následně byla aplikována metoda t-SNE s více příznaky na vizualizaci dat ve snaze odlišit tromby od okolních tkání. První pokusy s touto metodou nepřinesly dostatečně jasné výsledky, avšak po vyvážení dat a rozšíření sady příznaků bylo dosaženo lepší identifikace a tvorby shluků.

Třetí experiment se zaměřil na segmentaci trombu. K segmentaci byly použity metody jako Random Forest, Gradient Boosting a K-means. U každé metody byl zkoumán vliv příznakové sady na úspěšnost segmentace. Vliv trénovací sady byl zkoumán prostřednictvím křížové validace a porovnáním výsledků mezi modely natrénovanými na různých podmnožinách dat. Z výsledků je patrné, že správná volba trénovací sady významně ovlivňuje schopnost správně klasifikovat tromby.

Random Forest a Gradient Boosting dosáhly podobných výsledků, přičemž oba modely dosáhly průměrného DICE koeficientu okolo 0,5. U segmentace K-means se ukázalo, že tato metoda není tak účinná jako metody strojového učení.

I přes to, že se výsledky DICE mohou zdát příliš nízké, nebyl problém v lokalizaci trombu, ale v detekci jeho okrajových voxelů, což je výzvou i pro lidské experty.

# Literatura

- [1] Fiala, Pavel, Jiří Valenta a Lada Eberlová. *Stručná anatomie člověka*. Praha: Univerzita Karlova v Praze, nakladatelství Karolinum, 2015. ISBN 978- 80-246-2693-2.
- [2] Číhák, R.; Drvga, R.; Grim, M.: *Anatomie 3*. Praha: Grada, druhé vydání, 2004, ISBN 978-80-247-1132-4.
- [3] Kikalová, Kateřina, Libor Machálek, Vladimír Holibka, Miroslav Kuttal a Milada Bezdeková. *NEUROANATOMIE I. - Struktury centrálního nervového systému*. [online]. 2008 [cit. 2023-09-12]. Dostupné z: <http://www.nan.upol.cz/neuro/index.html>
- [4] Lynch, Patrick. Brain sagittal section stem highlighted. In: Wikimedia Commons [online]. 2009-06-19 [cit. 2022-09-13].
- [5] Web-FNUSA: Cévní mozková příhoda detailně [online]. [2022-11-05], Fakultní nemocnice u sv. Anny v Brně, informace pro širokou veřejnost. Dostupné z: <https://www.fnusa.cz/o-nemocnici/cmp/cevni-mozkova-prihoda-detailne/>
- [6] Vaníček, J.: *Mechanická rekanalizace v léčbě ischemické cévní mozkové příhody* [online]. Habilitační práce, Klinika zobrazovacích metod LF MU Brno, Brno, 2019, [2022-11-14]. Dostupné z: <https://www.med.muni.cz>
- [7] Abdu, H.; Tadese, F.; Seyoum, G.; aj.: Comparison of Ischemic and Hemorrhagic Stroke in the Medical Ward of Dessie Referral Hospital, Northeast Ethiopia. *Neurology Research International*, ročník 2021, 2021-6-28: s. 1–9, ISSN 2090-1860, doi:10.1155/2021/9996958. Dostupné z: <https://doi.org/10.1155/2021/9996958>
- [8] Dufek, M.: Cévní mozkové příhody, obecný úvod a klasifikace. *Interní medicína – mezioborové přehledy*, ročník 2002, č. 6, 2002: s. 5–10, online. [2022-11-05]. Dostupné z: <https://www.internimedicina.cz/pdfs/int/2002/06/10.pdf>
- [9] Brott, Thomas a Julien Bogousslavsky. Treatment of Acute Ischemic Stroke. *New England Journal of Medicine*. 2000, (10), 710-722.
- [10] Potter, Christopher A., Achala S. Vagal, Mayank Goyal, Diego B. Nunez, Thabile Leslie-Mazwi a Michael H. Lev. CT for Treatment Selection in Acute Ischemic Stroke: A Code Stroke Primer. *RadioGraphics*. 2019, 6(39).

- [11] Školoudík, D.; Kulíha, M.; Jonszta, T.; aj.: Endovaskulární léčba ischemické cévní mozkové příhody. *Cesk Slov Neurol N* [online], ročník 2012, č. 75/108(6), 2012: str. 669–683, [2022-11-14]. Dostupné z: [https://www.researchgate.net/profile/Roman-Herzig/publication/283157773\\_Endovascular\\_Treatment\\_of\\_an\\_Ischemic\\_Cerebrovascular\\_Event/links/58398d2a08ae3a74b49e87f9/Endovascular-Treatment-of-an-Ischemic-Cerebrovascular-Event.pdf](https://www.researchgate.net/profile/Roman-Herzig/publication/283157773_Endovascular_Treatment_of_an_Ischemic_Cerebrovascular_Event/links/58398d2a08ae3a74b49e87f9/Endovascular-Treatment-of-an-Ischemic-Cerebrovascular-Event.pdf)
- [12] Meyer, S. F. D.; Andersson, T.; Baxter, B.; aj.: Analyses of thrombi in acute ischemic stroke. *International Journal of Stroke*, ročník 12, č. 6, 2017: s. 606–614, ISSN 1747-4930, doi:10.1177/1747493017709671.
- [13] Konan LM, Reddy V, Mesfin FB. Neuroanatomy, Cerebral Blood Supply. [Aktualizováno 2023 Jul 24]. In: *StatPearls* [online]. Treasure Island (FL): StatPearls Publishing; 2023. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK532297/>
- [14] Štefánek, Jiří. Mozek - cévní zásobení. In: *Medicína, nemoci, studium na 1. LF UK* [online]. 2011 [cit. 2023-09-14]. Dostupné z: <https://www.stefajir.cz/mozek-cevni-zasobeni>
- [15] Brouwer, Patrick A., Waleed Brinjikji a Simon F. De Meyer. Clot Pathophysiology: Why Is It Clinically Important? *Neuroimaging Clin N Am*. 2018, 28(4), 611-623. Dostupné také z: doi: 10.1016/j.nic.2018.06.005
- [16] Léčba pacientů s ischemickou CMP a TIA - odborná část. In: *Cerebrovaskulární sekce ČNS JEP* [online]. [cit. 2023-09-15]. Dostupné z: <https://www.cmp.cz/guidelines-a-postupy>
- [17] Gascou G, Lobotesis K, Machi P, Maldonado I, Vendrell JF, Riquelme C, Eker O, Mercier G, Mourand I, Arquizan C, Bonafé A, Costalat V. Stent Retrievers in Acute Ischemic Stroke: Complications and Failures during the Perioperative Period. *AJNR Am J Neuroradiol* [online]. 2014, 35(4), 734-740 [cit. 2023-09-15]. Dostupné z: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7965801/>
- [18] Kim SK, Yoon W, Kim TS, Kim HS, Eoo TW, Park MS. Histologic Analysis of Retrieved Clots in Acute Ischemic Stroke: Correlation with Stroke Etiology and Gradient-Echo MRI. *AJNR Am J Neuroradiol* [online]. 2015, 36(9) [cit. 2023-09-15]. Dostupné z: doi:10.3174/ajnr.A4402

- [19] Patel TR, Santo BA, Baig AA, Wagas M, Monterio A, Levy EI, Siddiqui AH, Tutino VM. Histologically interpretable clot radiomic features predict treatment outcomes of mechanical thrombectomy for ischemic stroke. *Neuroradiology*. 2023, roč. 65, č. 4, s. 737-749.
- [20] Dumitriu Lagrange, Daniela; Hofmeister, Jeremy; Rosi, Andrea; Vargas, Maria Isabel; Wanke, Isabel et al., 2023. Predictive value of clot imaging in acute ischemic stroke: A systematic review of artificial intelligence and conventional studies. Online. *Neuroscience Informatics*. Roč. 3, č. 1. ISSN 27725286. Dostupné z: <https://doi.org/10.1016/j.neuri.2022.100114>. [cit. 2023-08-14].
- [21] Unnithan, A.K.A., M. Das, J., Mehta, P. (2023). Hemorrhagic Stroke. [online] StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; [Aktualizované 8. března 2023]. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK559173/>
- [22] Goldman, L. W.: Principles of CT and CT Technology. *Journal of Nuclear Medicine Technology*, ročník 35, č. 3, 2007-09-01: s. 115–128, ISSN 0091-4916, doi:10.2967/jnmt.107.042978.
- [23] Kalvach, P.: *Mozkové ischemie a hemoragie*. Praha: Grada, třetí vydání, 2010, ISBN 978-80-247-2765-3.
- [24] Reif, Michal; Goldemund, David a Mikulík, Robert. Nejdůležitější metody v diagnostice akutní cévní mozkové příhody. *Kardiologická revue – Interní medicína*. 2013, roč. 15, č. 1, s. 11-25.
- [25] Fidgerald S, Wang S, Dai D, Murphee DH Jr, Pandit A, Douglas A, Rizvi A, Kadivel R, Gilvarry M, McCarthy R, Stritt M, Gounis MJ, Brinjikji W, Kallmes DF, Doyle KM. Orbit image analysis machine learning software can be used for the histological quantification of acute ischemic stroke blood clots. Online. 2019, roč. 14, č. 12. Dostupné z: <https://doi.org/10.1371/journal.pone.0225841>. [cit. 2023-10-01].
- [26] Bhanu Prakash, K. N.; Zhou, Shi; Morgan, Tim C.; Hanley, Daniel F. a Nowinski, Wieslaw. Segmentation and quantification of intra-ventricular/cerebral hemorrhage in CT scans by modified distance regularized level set evolution technique. Online. *Int J Comput Assist Radiol Surg*. 2012, roč. 7, č. 5, s. 785-798. Dostupné z: <https://doi.org/10.1007/s11548-012-0670-0>. [cit. 2023-10-14].
- [27] Khan MM, Chowdhury MEH, Arefin ASMS, Podder KK, Hossain MSA, Alqahtani A, Murugappan M, Khandakar A, Mushtak A, Nahiduzzaman M.

- A. A Deep Learning-Based Automatic Segmentation and 3D Visualization Technique for Intracranial Hemorrhage Detection Using Computed Tomography Images. Online. *Diagnostics (Basel)*. 2023, roč. 13, č. 15. Dostupné z: <https://doi.org/10.3390/diagnostics13152537>. [cit. 2023-10-14].
- [28] Jan, J.: *Medical image processing, reconstruction and analysis*. London: CRC Press, druhé vydání, 2020, ISBN 9781138310285.
- [29] Fedorov A., Beichel R., Kalpathy-Cramer J., Finet J., Fillion-Robin J-C., Pujol S., Bauer C., Jennings D., Fennessy F., Sonka M., Buatti J., Aylward S.R., Miller J.V., Pieper S., Kikinis R. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. Online. *Magnetic Resonance Imaging*. 2013, roč. 30, č. 9, s. 1323-41. Dostupné z: <https://www.slicer.org/>. [cit. 2023-10-14].
- [30] Deweerth, Stephen P. a Morris, T.G. *Analog VLSI Arrays for Morphological Image Processing*. Online. *International Conference on Application-Specific Systems, Architectures and Processors, Proceedings*. Dostupné z: <https://doi.org/10.1109/ASAP.1994.331809>. [cit. 2023-10-28].
- [31] Chhikara, Prateek. *Understanding Morphological Image Processing and Its Operations*. Online. *Medium*. 2022. Dostupné z: <https://towardsdatascience.com/understanding-morphological-image-processing-and-its-operations-7bcf1ed11756>. [cit. 2023-10-28].
- [32] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: Extracting more information from medical images using advanced feature analysis. Online. *Eur J Cancer*. 2012, roč. 48, č. 4, s. 441-446. Dostupné z: <https://doi.org/10.1016/j.ejca.2011.11.036>. [cit. 2023-11-18].
- [33] Mahmoudi, S., Martin, S.S., Ackermann, J. et al. Potential of high dimensional radiomic features to assess blood components in intraaortic vessels in non-contrast CT scans. *BMC Med Imaging* 21, 123, 2021. <https://doi.org/10.1186/s12880-021-00654-9>. [cit. 2023-11-18].
- [34] van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Aerts, H. J. W. L. (2017). Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*, 77(21), 104–107. DOI: 10.1158/0008-5472.CAN-17-0339

- [35] Marta Olive-Gadea, Carlos Crespo, Cristina Granes, Maria Hernandez-Perez, Natalia Pérez de la Ossa, Carlos Laredo, Xabier Urrea, Juan Carlos Soler, Alexander Soler, Paloma Puyalto, Patricia Cuadras, Cristian Marti, a Marc Ribo. "Deep Learning Based Software to Identify Large Vessel Occlusion on Non-contrast Computed Tomography." *Stroke*, 2020, roč. 51, s. 3133–3137. DOI: 10.1161/STROKEAHA.120.030326.
- [36] Fitzgerald, S., Wang, S., Dai, D., Murphree, D. H. Jr., Pandit, A., Douglas, A., Rizvi, A., Kadirvel, R., Gilvarry, M., McCarthy, R., Stritt, M., Gounis, M. J., Brinjikji, W., Kallmes, D. F., a Doyle, K. M. "Orbit image analysis machine learning software can be used for the histological quantification of acute ischemic stroke blood clots." *PLoS One*, 2019, roč. 14, č. 12, e0225841. DOI: 10.1371/journal.pone.0225841. PMID: 31805096; PMCID: PMC6894878.
- [37] Patel, T. R., Santo, B. A., Baig, A. A., Waqas, M., Monterio, A., Levy, E. I., Siddiqui, A. H., a Tutino, V. M. "Histologically interpretable clot radiomic features predict treatment outcomes of mechanical thrombectomy for ischemic stroke." *Neuroradiology*, 2023, roč. 65, č. 4, s. 737-749. DOI: 10.1007/s00234-022-03109-2. Epub 2023 Jan 4. PMID: 36600077.
- [38] Athanasiou, L. S., Fotiadis, D. I., a Michalis, L. K. (2017). "Plaque Characterization Methods Using Intravascular Ultrasound Imaging." In: Athanasiou, L. S., Fotiadis, D. I., a Michalis, L. K. (Eds.), *Atherosclerotic Plaque Characterization Methods Based on Coronary Imaging* (str. 71-94). Academic Press. ISBN 9780128047347. DOI: 10.1016/B978-0-12-804734-7.00004-X. Dostupné z: <https://www.sciencedirect.com/science/article/pii/B978012804734700004X>
- [39] Al-Areqi, F., & Konyar, M. Z. (2022). Effectiveness Evaluation of Different Feature Extraction Methods for Classification of Covid-19 from Computed Tomography Images: A High Accuracy Classification Study. *Biomedical Signal Processing and Control*, 103662. <https://doi.org/10.1016/j.bspc.2022.103662>.
- [40] Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3), 1116-1128. DOI: 10.1016/j.neuroimage.2006.01.015.
- [41] Holčík, Jiří, Komenda, Martin (eds.) a kol. *Matematická biologie: e-learningová učebnice* [online]. 1. vydání. Brno: Masarykova univerzita, 2015. ISBN 978-80-210-8095-9.

- [42] McClenaghan, E. (2022). Mann-Whitney U Test: Assumptions and Example. Technology Networks. [Online] Dostupné z: <https://www.technologynetworks.com/informatics/articles/mann-whitney-u-test-assumptions-and-example-363425>, [cit. 2023-11-18].
- [43] Ahlers, J., et al. (2023). napari: a multi-dimensional image viewer for Python. Zenodo. Dostupné z: <https://doi.org/10.5281/zenodo.3555620>
- [44] Beare, R., Lowekamp, B. C., Yaniv, Z. (2018). Image Segmentation, Registration and Characterization in R with SimpleITK. Journal of Statistical Software, 86(8). Dostupné z: <https://doi.org/10.18637/jss.v086.i08>
- [45] Liebeskind, D. S., et al. (2011). CT and MRI Early Vessel Signs Reflect Clot Composition in Acute Stroke. Stroke, 42(5), 1237–1243. <https://doi.org/10.1161/STROKEAHA.110.605576>
- [46] Nemcekova, P., Holecek, T., Chmelik, J., Ourednicek, P., Valis, K., Jakubicek, R. (Nepublikováno). Establishing the Optimal Standard for Preprocessing Head CT Data in Diagnostic Analysis. Department of Biomedical Engineering, FEEC, Brno University of Technology, Technicka 3082/12, 616 00, Brno, Czech Republic; Department of Imaging Methods, St. Anne’s University Hospital Brno, Pekarska 664/53, 656 91, Brno, Czech Republic.
- [47] Wolf, I.; Vetter, M.; Wegner, I.; aj.: The Medical Imaging Interaction Toolkit (MITK) [online]. [2022-11-22]. Univerzita Heidelberg. URL [https://www.mitk.org/wiki/The\\_Medical\\_Imaging\\_Interaction](https://www.mitk.org/wiki/The_Medical_Imaging_Interaction)
- [48] Nemcekova, P., Holecek, T., Chmelik, J., Ourednicek, P., Valis, K., Jakubicek, R. (2024). Establishing the Optimal Standard for Preprocessing Head CT Data in Diagnostic Analysis. In A. Badnjević & L. Gurbeta Pokvić (Eds.), *MEDICON’23 and CMBEBIH’23* (pp. 162-169). Cham: Springer Nature Switzerland. DOI: 10.1007/978-3-031-49068-2\_18
- [49] OpenAI. (2024). ChatGPT – jazykový model GPT-4 free byl v práci použit jako asistenční nástroj pro hledání chyb v kódu a jejich úpravu. Dostupné z: <https://chat.openai.com> [2024-05-21].

## Seznam symbolů a zkratek

<b>AUC</b>	Area Under the Curve
<b>CMP</b>	Cévní mozková příhoda
<b>CT</b>	Výpočetní tomografie
<b>CTA</b>	Počítačová tomografická angiografie
<b>FNUSA</b>	Fakultní nemocnice u sv. Anny v Brně
<b>nCT</b>	Nekontrastní CT vyšetření mozku
<b>mCTA</b>	Vícefázová CT angiografie
<b>GLCM</b>	Gray Level Co-occurrence Matrix
<b>GLDM</b>	Gray Level Dependence Matrix
<b>GLRLM</b>	Gray Level Run Length Matrix
<b>GLSZM</b>	Gray Level Size Zone Matrix
<b>MRI</b>	Magnetická rezonance
<b>NCCT</b>	Nekontrastní výpočetní tomografie
<b>NGTDM</b>	Neighboring Gray Tone Difference Matrix
<b>PET</b>	Pozitronová emisní tomografie
<b>ROC</b>	Receiver Operating Characteristic
<b>PCA</b>	Principal Component Analysis

# Seznam příloh

<b>A Přílohy</b>	<b>65</b>
A.1 Seznam příznaků . . . . .	65
A.2 t-SNE 10 vybraných příznaků . . . . .	68
A.3 t-SNE 15 vybraných příznaků . . . . .	71
<b>B Obsah elektronické přílohy</b>	<b>73</b>

# A Přílohy

## A.1 Seznam příznaků

Index	Název příznaku
0	original_firstorder_10Percentile.nrrd
1	original_firstorder_90Percentile.nrrd
2	original_firstorder_Energy.nrrd
3	original_firstorder_Entropy.nrrd
4	original_firstorder_InterquartileRange.nrrd
5	original_firstorder_Kurtosis.nrrd
6	original_firstorder_Maximum.nrrd
7	original_firstorder_Mean.nrrd
8	original_firstorder_MeanAbsoluteDeviation.nrrd
9	original_firstorder_Median.nrrd
10	original_firstorder_Minimum.nrrd
11	original_firstorder_Range.nrrd
12	original_firstorder_RobustMeanAbsoluteDeviation.nrrd
13	original_firstorder_RootMeanSquared.nrrd
14	original_firstorder_Skewness.nrrd
15	original_firstorder_TotalEnergy.nrrd
16	original_firstorder_Uniformity.nrrd
17	original_firstorder_Variance.nrrd
18	original_glcmm_Autocorrelation.nrrd
19	original_glcmm_ClusterProminence.nrrd
20	original_glcmm_ClusterShade.nrrd
21	original_glcmm_ClusterTendency.nrrd
22	original_glcmm_Contrast.nrrd
23	original_glcmm_Correlation.nrrd
24	original_glcmm_DifferenceAverage.nrrd
25	original_glcmm_DifferenceEntropy.nrrd
26	original_glcmm_DifferenceVariance.nrrd
27	original_glcmm_Id.nrrd
28	original_glcmm_Idm.nrrd
29	original_glcmm_Idmn.nrrd

Tab. A.1: Extrahované příznaky (část 1).

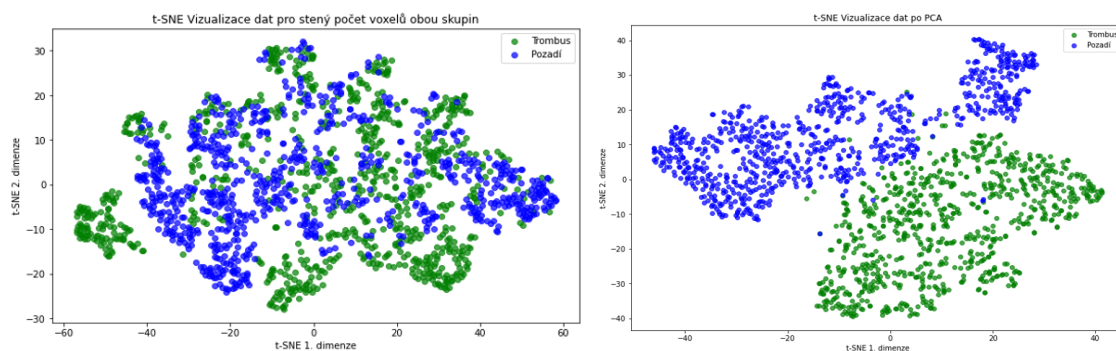
Index	Název příznaku
30	original_glcm_Idn.nrrd
31	original_glcm_Imc1.nrrd
32	original_glcm_Imc2.nrrd
33	original_glcm_InverseVariance.nrrd
34	original_glcm_JointAverage.nrrd
35	original_glcm_JointEnergy.nrrd
36	original_glcm_JointEntropy.nrrd
37	original_glcm_MaximumProbability.nrrd
38	original_glcm_SumEntropy.nrrd
39	original_glcm_SumSquares.nrrd
40	original_gldm_DependenceEntropy.nrrd
41	original_gldm_DependenceNonUniformity.nrrd
42	original_gldm_DependenceNonUniformityNormalized.nrrd
43	original_gldm_DependenceVariance.nrrd
44	original_gldm_GrayLevelNonUniformity.nrrd
45	original_gldm_GrayLevelVariance.nrrd
46	original_gldm_HighGrayLevelEmphasis.nrrd
47	original_gldm_LargeDependenceEmphasis.nrrd
48	original_gldm_LargeDependenceHighGrayLevelEmphasis.nrrd
49	original_gldm_LargeDependenceLowGrayLevelEmphasis.nrrd
50	original_gldm_LowGrayLevelEmphasis.nrrd
51	original_gldm_SmallDependenceEmphasis.nrrd
52	original_gldm_SmallDependenceHighGrayLevelEmphasis.nrrd
53	original_gldm_SmallDependenceLowGrayLevelEmphasis.nrrd
54	original_glrlm_GrayLevelNonUniformity.nrrd
55	original_glrlm_GrayLevelNonUniformityNormalized.nrrd
56	original_glrlm_GrayLevelVariance.nrrd
57	original_glrlm_HighGrayLevelRunEmphasis.nrrd
58	original_glrlm_LongRunEmphasis.nrrd
59	original_glrlm_LongRunHighGrayLevelEmphasis.nrrd
60	original_glrlm_LongRunLowGrayLevelEmphasis.nrrd
61	original_glrlm_LowGrayLevelRunEmphasis.nrrd
62	original_glrlm_RunEntropy.nrrd
63	original_glrlm_RunLengthNonUniformity.nrrd
64	original_glrlm_RunLengthNonUniformityNormalized.nrrd

Tab. A.2: Extrahované příznaky (část 2).

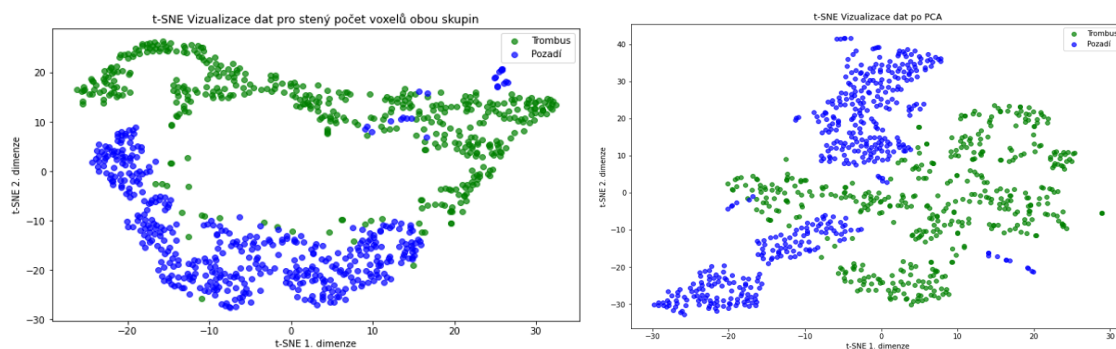
Index	Název příznaku
65	original_glrlm_RunPercentage.nrrd
66	original_glrlm_RunVariance.nrrd
67	original_glrlm_ShortRunEmphasis.nrrd
68	original_glrlm_ShortRunHighGrayLevelEmphasis.nrrd
69	original_glrlm_ShortRunLowGrayLevelEmphasis.nrrd
70	original_glszm_GrayLevelNonUniformity.nrrd
71	original_glszm_GrayLevelNonUniformityNormalized.nrrd
72	original_glszm_GrayLevelVariance.nrrd
73	original_glszm_HighGrayLevelZoneEmphasis.nrrd
74	original_glszm_LargeAreaEmphasis.nrrd
75	original_glszm_LargeAreaHighGrayLevelEmphasis.nrrd
76	original_glszm_LargeAreaLowGrayLevelEmphasis.nrrd
77	original_glszm_LowGrayLevelZoneEmphasis.nrrd
78	original_glszm_SizeZoneNonUniformity.nrrd
79	original_glszm_SizeZoneNonUniformityNormalized.nrrd
80	original_glszm_SmallAreaEmphasis.nrrd
81	original_glszm_SmallAreaHighGrayLevelEmphasis.nrrd
82	original_glszm_SmallAreaLowGrayLevelEmphasis.nrrd
83	original_glszm_ZoneEntropy.nrrd
84	original_glszm_ZonePercentage.nrrd
85	original_glszm_ZoneVariance.nrrd
86	original_ngtdm_Busyness.nrrd
87	original_ngtdm_Coarseness.nrrd
88	original_ngtdm_Complexity.nrrd
89	original_ngtdm_Contrast.nrrd
90	original_ngtdm_Strength.nrrd

Tab. A.3: Extrahované příznaky (část 3).

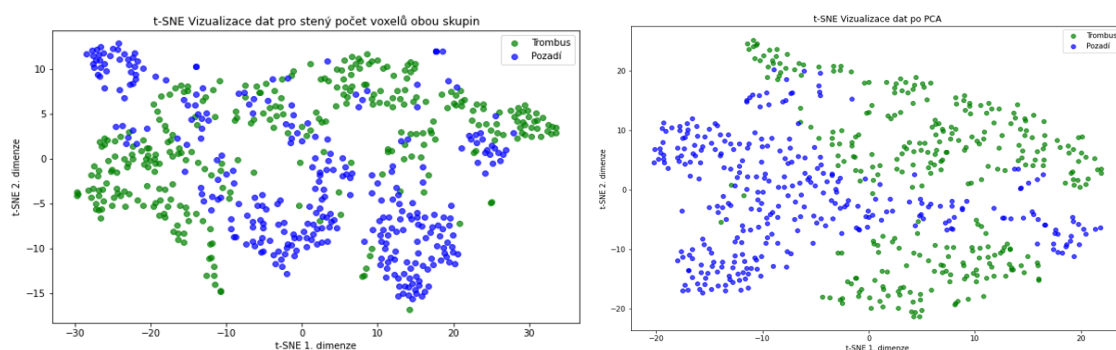
## A.2 t-SNE 10 vybraných příznaků



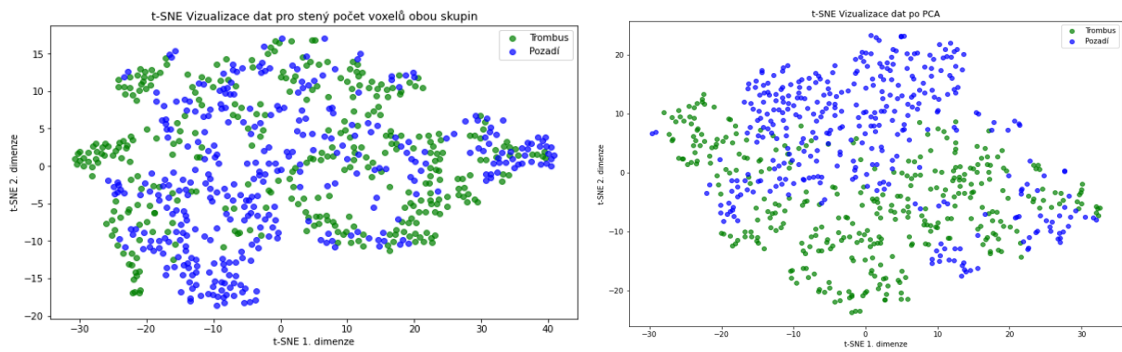
Obr. A.1: t-SNE vizualizace vybraných příznaků pacienta č. 112 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.



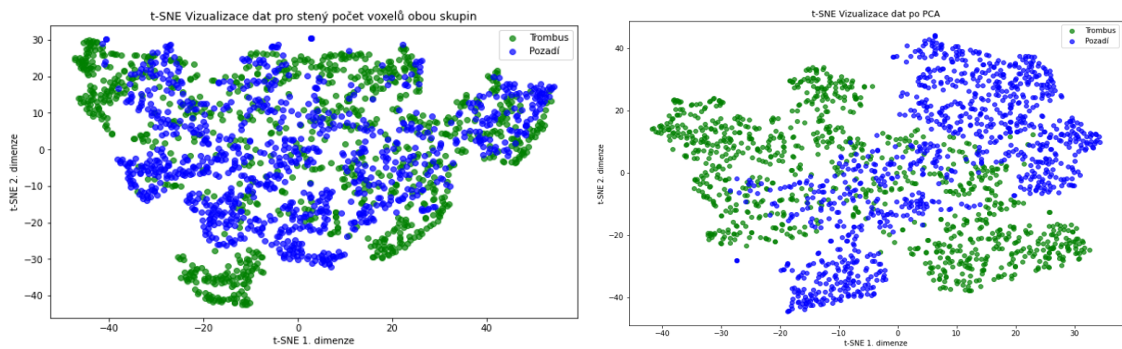
Obr. A.2: t-SNE vizualizace vybraných příznaků pacienta č. 238 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.



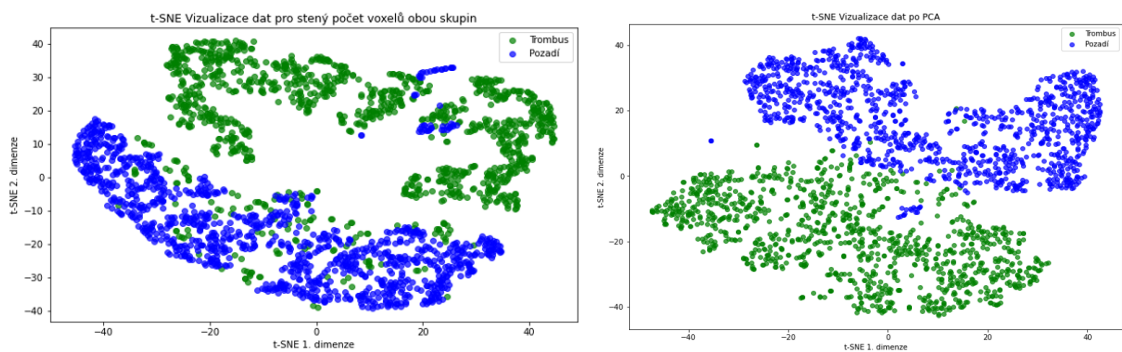
Obr. A.3: t-SNE vizualizace vybraných příznaků pacienta č. 243 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.



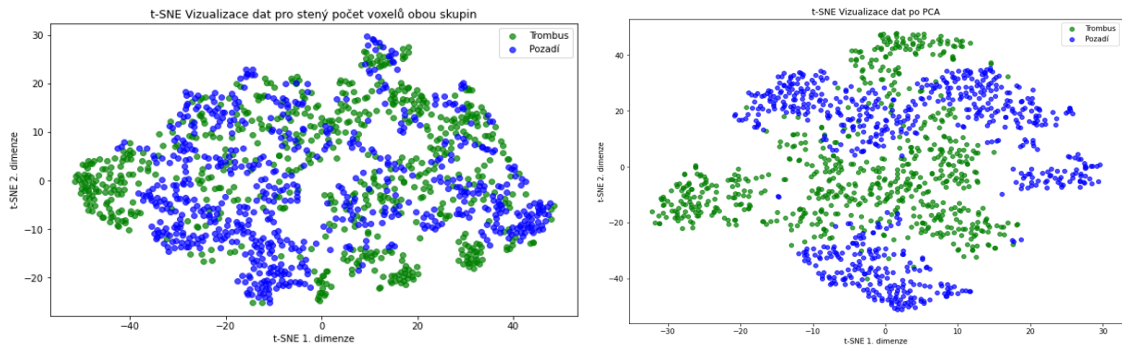
Obr. A.4: t-SNE vizualizace vybraných příznaků pacienta č. 297 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.



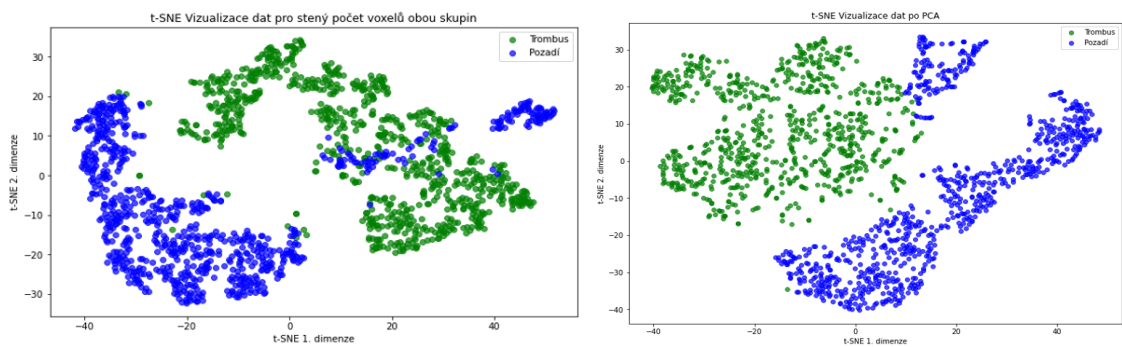
Obr. A.5: t-SNE vizualizace vybraných příznaků pacienta č. 319 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.



Obr. A.6: t-SNE vizualizace vybraných příznaků pacienta č. 365 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.

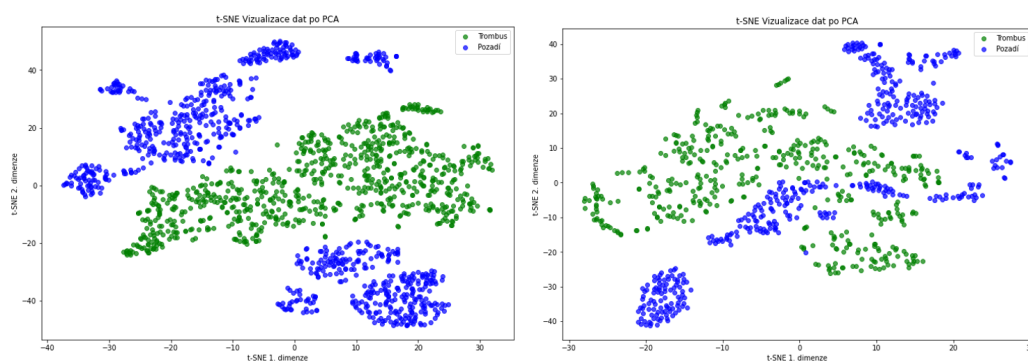


Obr. A.7: t-SNE vizualizace vybraných příznaků pacienta č. 394 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.

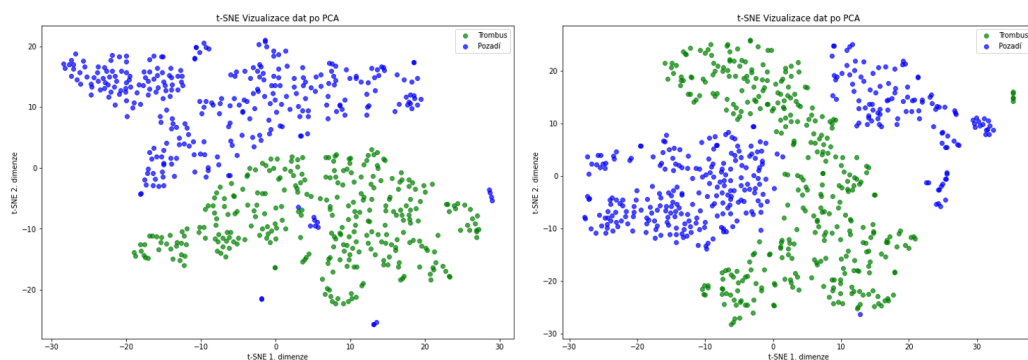


Obr. A.8: t-SNE vizualizace vybraných příznaků pacienta č. 400 bez redukce dimenzionality pomocí PCA (vlevo) a s metodou PCA (vpravo) pro vyrovnané počty dat.

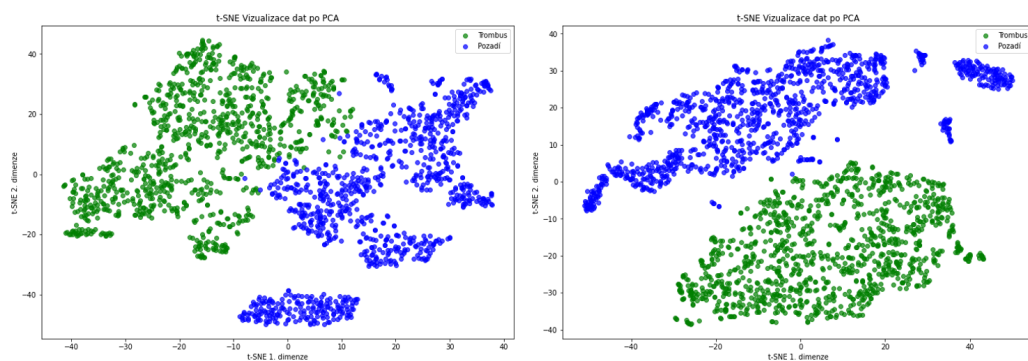
### A.3 t-SNE 15 vybraných příznaků



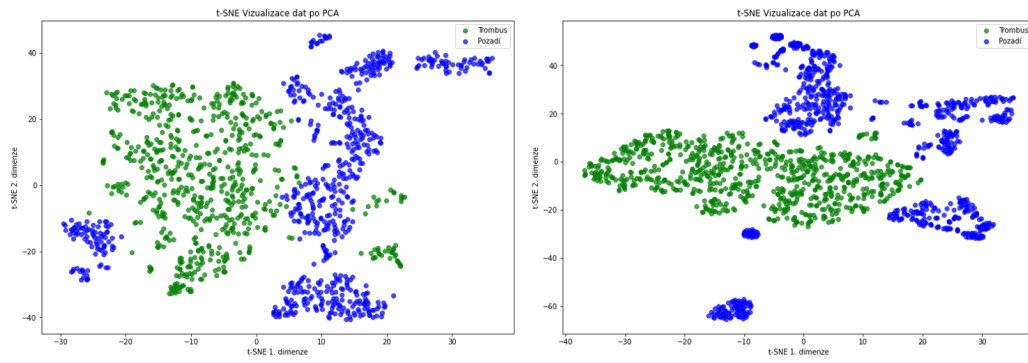
Obr. A.9: t-SNE vizualizace vybraných příznaků pacienta č. 112 (vlevo) a č. 238 (vpravo) s redukcí dimenzionality pomocí PCA.



Obr. A.10: t-SNE vizualizace vybraných příznaků pacienta č. 243 (vlevo) a č. 297 (vpravo) s redukcí dimenzionality pomocí PCA.



Obr. A.11: t-SNE vizualizace vybraných příznaků pacienta č. 319 (vlevo) a č. 365 (vpravo) s redukcí dimenzionality pomocí PCA..



Obr. A.12: t-SNE vizualizace vybraných příznaků pacienta č. 394 (vlevo) a č. 400 (vpravo) s redukcí dimenzionality pomocí PCA.

## B Obsah elektronické přílohy

priloha .....	Kořenový adresář
├─ experiment_1.py	
├─ experiment_2.py	
├─ experiment_2_tsne.py	
├─ extrakce_priznaku.py	
├─ gradient_boosting	
├─ k_means.py	
├─ random_forrest.py	