



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

METODY HLUBOKÉHO UČENÍ PRO SEGMENTACI ELEKTROKARDIOGRAFICKÝCH SIGNÁLŮ

SEGMENTATION OF ELECTROCARDIOGRAPHIC SIGNALS USING DEEP LEARNING METHODS

DIZERTAČNÍ PRÁCE

DOCTORAL THESIS

AUTOR PRÁCE

AUTHOR

Ing. Jakub Hejč

ŠKOLITEL

SUPERVISOR

doc. Ing. Jana Kolářová, Ph.D.

BRNO 2023

ABSTRAKT

Práce se zabývá metodami hlubokého učení pro segmentaci elektrokardiografických záznamů z povrchového a intrakardiálního snímání se zaměřením na síňovou aktivitu. Teoretická část seznamuje se současnými přístupy segmentace a s problematikou spojenou s využíváním standardních EKG databází pro vývoj hlubokých modelů. V práci byl navržen řetězec metod pro zpracování multimodálních dat z elektrofyziologického vyšetření. Tyto metody byly následně využity pro sestavení trénovacích datových sad. V práci byl navržen hluboký model pro segmentaci intrakardiálních záznamů založený na modifikované architektuře s reziduálním spojením. Byla provedena série experimentů sledující vliv nastavení modelu a předzpracování datové sady na kvalitu segmentace. Jako zásadní činitel se ukázala přítomnost a způsob anotace záznamů s fibrilací síní. Mezi další významné parametry patřilo nastavení penalizační funkce a způsob augmentace trénovacích dat. V práci byla dále navržena nová metoda segmentace P vln využívající neúplné reference. Přístup byl inspirován metodou hlubokého kontrastního učení, která byla modifikována pro odlišení lokálních úseků signálů na různých úrovních abstrakce extrahovaných příznakových map. Výsledky byly analyzovány pomocí standardních metrik kvality a post-hoc vizuální analýzou. V dílčích případech bylo provedeno statistické srovnání experimentů pro různá nastavení. Výsledky práce ukázaly, že je možné využít intrakardiální signály pro zabudování vektorové reprezentace lokální síňové aktivity do hlubokých modelů.

KLÍČOVÁ SLOVA

Metody hlubokého učení, konvoluční neuronové sítě, automatická segmentace signálů, intervenční elektrofyziologické vyšetření, intrakardiální elektrogramy, lokální aktivační čas, síňová aktivita, skryté P vlny, poruchy srdečního rytmu, databáze signálů.

ABSTRACT

The thesis deals with deep learning methods for the segmentation of surface and intracardiac electrocardiographic recording with focus on atrial activity. The theoretical part introduces current segmentation approaches of electrocardiographic signals. Issues related to the development of deep learning models in context of standard ECG databases were also discussed. We proposed a pipeling for processing multimodal electrophysiology data from interventional procedures in order to build reliable training datasets. A deep model for segmentation of intracardiac recordings based on a modified residual architecture was proposed. A series of experiments was conducted to evaluate the effect of both model and dataset properties on segmentation quality. The annotation methodology of recordings with atrial fibrillation proved to be a crucial factor. Properties of loss function and type of data augmentation were revealed as secondary important parameters. A novel P wave segmentation method for incomplete references was proposed in the thesis. The approach was inspired by the deep contrast learning. It was modified to distinguish local segments of signals at different levels of abstraction of the extracted feature maps. Results were analyzed using standard quality metrics and post-hoc visual analysis. In some cases, a statistical comparison of experiments for different settings was performed. The results of the work showed that it is possible to use intracardiac signals for embedding a vector representation of local atrial activation into deep models.

KEYWORDS

Deep learning, convolutional neural networks, segmentation, cardiac mapping, intracardiac electrograms, local activation time, atrial activity, hidden P waves, arrhythmias.

HEJČ, Jakub. *Metody hlubokého učení pro segmentaci elektrokardiografických signálů*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství, 2023, 131 s. Disertační práce. Vedoucí práce: doc. Ing. Kolářová Jana, Ph.D.

Prohlášení autora o původnosti díla

Jméno a příjmení autora: Ing. Jakub Hejč
VUT ID autora: 106152
Typ práce: Disertační práce
Akademický rok: 2022/23
Téma závěrečné práce: Metody hlubokého učení pro segmentaci elektrokardiografických signálů

Prohlašuji, že svou závěrečnou práci jsem vypracoval 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 autor uvedené závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom 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 autora*

*Autor podepisuje pouze v tištěné verzi.

PODĚKOVÁNÍ

Rád bych poděkoval všem, kdo mně podporovali v průběhu doktorského studia. Zvláštní poděkování patří mé školitelce doc. Ing. Janě Kolářové, Ph.D. a vedoucímu výzkumného týmu ICE doc. MUDr. Zdeňku Stárkovi, Ph.D. za poskytnuté konzultace, nezměrnou trpělivost a vytváření pracovních podmínek, které umožnili dokončení práce. Vřelé poděkování patří také MUDr. Martinu Pešlovi, Ph.D. za zájem a cenné poznatky k dizertační práci. Dále děkuji kolegům Ing. Ředinovi, Ing. Rakové, Ing. Širůčkové a Ing. Kubalové za čas a ochotu věnovanou anotaci elektrofyziologických dat. Nejsrdečnější poděkování patří mé manželce Kláře. . . za vše, čím pro mě nejen při studiu byla.

Obsah

Úvod	12
Cíle práce	14
1 Zpracování klinických dat	17
1.1 Úvod	17
1.2 Intervenční elektrofyziologické vyšetření	19
1.3 Dílčí cíle	20
1.4 Zpětná analýza dat	20
1.4.1 Základní popis	20
1.4.2 Výsledky	22
1.5 Softwarové balíčky pro manipulaci s EFV daty	23
1.6 Učební datové sady	25
1.6.1 Charakteristika zařazených pacientů	25
1.6.2 Selektce zájmových úseků	25
1.6.3 Expertně anotovaná datová sada a (DS-A)	26
1.6.4 Automaticky anotovaná datová sada B (DS-B)	28
1.7 Diskuze a limitace	31
2 Segmentace síňových elektrogramů pomocí modelů hlubokého učení	33
2.1 Současný stav	33
2.2 Dílčí cíle	36
2.3 Formální definice problému	37
2.4 Architektura modelu	37
2.4.1 Enkodér s reziduálním spojením	37
2.4.2 Blok dilatovaných konvolucí	39
2.4.3 Pyramidový dekodér	41
2.5 Kriteriaální funkce	42
2.6 Optimalizační algoritmus	43
2.6.1 Stochastická gradientní optimalizace	43
2.6.2 Plánování kroku učení	44
2.7 Stratifikace DS-A	45
2.8 Předzpracování dat	46
2.9 Bayesovská optimalizace hyperparametrů	47
2.9.1 Optimalizace M_{aopt} a M_{lopt}	48
2.10 Hodnocení kvality modelu	50
2.10.1 Statistické metody	52
2.11 Výsledky	52
2.11.1 Optimalizace architektury	52

2.11.2	Optimalizace kriteriální funkce	55
2.11.3	Srovnání modelů M_{aopt} a M_{lopt}	56
2.11.4	Segmentace fibrilačních záznamů	59
2.11.5	Hodnocení vlivu šumu	60
2.11.6	Vliv augmentace dat	65
2.11.7	Vliv stratifikace datové sady	66
2.12	Ukázky segmentace	66
2.13	Diskuze	74
2.13.1	Kvalita segmentace pomocí DNN modelů	74
2.13.2	Vliv fibrilačních záznamů	75
2.13.3	Vlivu šumu	76
2.13.4	Vliv nastavení penalizační funkce	77
2.13.5	Vliv stratifikace datové sady	77
2.13.6	Limitace	78
3	Slabě dozorovaná segmentace P vln	79
3.1	Současný stav	79
3.1.1	Metody založené na prohledávání lokální oblasti	80
3.1.2	Metody založené na potlačení komorové aktivity	81
3.1.3	Metody využívající hluboké učení	81
3.1.4	Shrnutí	82
3.2	Dílčí cíle	83
3.3	Definice problému segmentace s neúplnou informací	84
3.4	Architektura modelu	85
3.4.1	Hyperparametry	85
3.4.2	Regionální příznakové mapy	85
3.5	Normalizovaná křížová entropie pro kontrastní učení	86
3.5.1	Kontrastní učení v lokalizované oblasti signálu	87
3.6	MIL vrstva	90
3.7	Popis experimentů	90
3.8	Trénovací data	91
3.9	Předzpracování dat	92
3.10	Nastavení procesu učení	93
3.11	Hodnocení kvality segmentace	94
3.12	Výsledky	94
3.13	Ukázky segmentace	97
3.14	Diskuze	101
3.14.1	Jak věrohodná je segmentace P vlny hlubokým modelem?	101
3.14.2	Využití intrakardiálních referencí	102
3.14.3	Limitace	102

Literatura	105
Seznam příloh	124
A Přílohy	125
A.1 Vybrané implementační detaily	125
A.1.1 Parametry augmentačních technik v Kapitole 2	125
A.1.2 Parametry augmentačních technik v Kapitole 3	126
A.1.3 Definice zahřívací fáze	126
A.1.4 Softwarové a hardwarové prostředky	127
B Curriculum vitae	128

Seznam obrázků

1	Blokové znázornění provázanosti kapitol disertační práce	15
1.1	Standardní anatomického postavení diagnostických katétrů při EFV	19
1.2	Schéma archivačního datového formátu WorkMate – EKG data	22
1.3	Schéma archivačního datového formátu WorkMate – klinické anotace	23
1.4	Schéma systému pro manipulaci s elektrofyziologickými daty	24
1.5	Vylučovací kritéria pro selekci EFV záznamů	25
1.6	Zastoupení diagnostických tříd v DS-A	27
1.7	Relativní časové zastoupení síňových stahů v DS-A	28
1.8	Zastoupení obecných diagnostických tříd v DS-B.	29
1.9	Ukázka typických záznamů v DS-A a DS-B	30
2.1	Morfologická a frekvenční variabilita lokálních intrakardiálních elektrogramů	35
2.2	Obecný princip sémantické segmentace EGM.	37
2.3	Architektura 1D reziduální jednotky s předaktivací	38
2.4	Schéma 1D konvolučního enkodéru založeného na architektuře ResNet	39
2.5	Schématické znázornění dilatovaného konvolučního operátoru.	40
2.6	Schéma bloku s dilatovanými konvolučními jádry.	40
2.7	Blokové schéma 1D dekodéru s pyramidovým zpracováním příznaků.	41
2.8	Závislost FTL na exponenciálním váhování	42
2.9	Schématické znázornění implementovaného algoritmu pro řízení kroku učení.	45
2.10	Histogram relativního rozložení diagnostických tříd v DS-A _T a DS-A _V	45
2.11	Blokové schéma řetězce pro předzpracování signálů	47
2.12	Relativní skóre významnosti optimalizovaných hyperparametrů architektury	53
2.13	Závislost F ₁ skóre na hyperparametrech modelu M _{aopt}	54
2.14	Vztah mezi F ₁ skóre a hyperparametry FTL	56
2.15	Vztah mezi Se a P ⁺ a hyperparametry FTL	56
2.16	F ₁ , Se a P ⁺ na DS-A _V pro modely M _{aopt} a M _{lopt}	57
2.17	PE na DS-A _V pro modely.	57
2.18	Srovnání přesnosti stanovení počátku a konce u stahů SB, AB a ST.	58
2.19	F ₁ , Se a P ⁺ v závislosti na hladině bílého šumu	60
2.20	Vliv bílého šumu na přesnost nalezení počátku a konce síňové aktivity	61
2.21	Vliv bílého šumu na PE	62
2.22	F ₁ , Se a P ⁺ v závislosti na hladině síťového brumu	62
2.23	Vliv síťového brumu na přesnost nalezení počátku a konce síňové aktivity.	63
2.24	Vliv bílého šumu na PE	63
2.25	F ₁ , Se a P ⁺ v závislosti na počtu dostupných kanálů	64
2.26	Časová odchylka aktivačních časů v závislosti na počtu kanálů	64
2.27	Vliv augmentačních technik na metriky kvality	65
2.28	Krabicové grafy F ₁ hodnot M _{aopt} pro různé scénáře stratifikace.	66
2.29	Ukázka automatické segmentace EGM č. 1.	67
2.30	Ukázka automatické segmentace EGM č. 2.	67

2.31	Ukázka automatické segmentace EGM č. 3	68
2.32	Ukázka automatické segmentace EGM č. 4	68
2.33	Ukázka automatické segmentace EGM č. 5	69
2.34	Ukázka automatické segmentace EGM č. 6	69
2.35	Ukázka automatické segmentace EGM č. 6	70
2.36	Ukázka automatické segmentace EGM č. 7	70
2.37	Ukázka automatické segmentace EGM č. 8	70
2.38	Ukázka automatické segmentace EGM č. 9	71
2.39	Ukázka automatické segmentace EGM č. 10	71
2.40	Ukázka automatické segmentace EGM č. 11	72
2.41	Ukázka automatické segmentace EGM č. 12	73
2.42	Ukázka automatické segmentace EGM č. 13	73
2.43	Ukázka automatické segmentace EGM č. 14	74
3.1	Obecný princip učení s neúplnou informací v případě segmentace P vlny. . .	84
3.3	Princip kontrastní penalizace pro dílčí úseky příznakových map	89
3.4	Testované scénáře s různým typem penalizace modelu	91
3.5	Histogram relativního rozložení diagnostických tříd v DS-B _T a DS-B _V . . .	92
3.6	Distribuce napěťových hodnot EKG v DS-B	93
3.7	Ukázka segmentace P vlny č. 1	97
3.8	Ukázka segmentace P vlny č. 2	98
3.9	Ukázka segmentace P vlny č. 3	99
3.10	Ukázka segmentace P vlny č. 4	100

Seznam tabulek

1.1	Přehled standardních databází EKG záznamů s referenční anotací P vlny	17
1.2	Základní charakteristika datové sady DS-A.	27
1.3	Základní charakteristika datové sady DS-B	29
2.1	Nastavení hyperparametrů optimalizačního algoritmu.	43
2.2	Parametry řízení kroku učení a ukončení optimalizace.	44
2.3	Specifikace prostoru hyperparametrů modelu pro parametry architektury	49
2.4	Specifikace prostoru hyperparametrů pro kritériální funkci	50
2.5	Výsledky Bayesovské optimalizace architektury – metriky kvality	52
2.6	Výsledky Bayesovské optimalizace architektury – hyperparametry	53
2.7	Výsledky Bayesovské optimalizace penalizační funkce	55
2.8	Srovnání přesnosti stanovení počátku a konce síňové u stahů SB, AB a ST.	58
2.9	Výsledky ablační studie u modelu M_{lopt} pro různé AF scénáře	59
3.1	Specifikace prostoru hyperparametrů modelu pro parametry architektury	85
3.2	Nastavení hyperparametrů procesu učení	93
3.3	Výsledky segmentace v MIL vrstvě	94
3.4	Výsledky Se a P^+ významných bodů P vlny	95
3.5	Výsledky Se a P^+ významných bodů P vlny u NSR	96
A.1	Specifikace parametrů augmentačních technik.	125
A.2	Specifikace parametrů augmentačních technik.	126

Úvod

Standardní 12-svodový elektrokardiogram (EKG) je pro svou neinvazivnost, ekonomickou efektivitu a diagnostickou výtěžnost nejvýznamnější technikou pro monitoraci srdeční aktivity [1]. Základní podstatou je snímání časově-prostorových změn elektromagnetického pole, které vzniká díky šíření změn transmembránového napětí srdečních buněk [2]. Snímaný časový průběh EKG signálu reprezentuje elektrofyziologické parametry jednotlivých fází srdečního cyklu a umožňuje diagnostiku i stanovení rizik široké škály srdečních onemocnění.

Aktivita srdce typicky zahrnuje elektrickou aktivaci, kontrakci a relaxaci srdečních síní a srdečních komor. V EKG signálu snímaném z povrchu těla je elektrická aktivace a kontrakce reprezentována vlnou P v případě síní a komplexem QRS v případě komor. Elektrická restituce komor je pak označována jako vlna T. Časové uspořádání, morfologie a amplituda těchto vln patří mezi významné klinické indikátory. Jejich kvantitativní hodnocení je základním krokem při diagnostické interpretaci EKG záznamu [3]. Časovou lokalizaci, stanovení počátku a konce zmíněných útvarů označujeme obecně jako *segmentaci* EKG signálu. V případě dalších parametrů, jakými mohou být např. pozice vrcholů, pak širším pojmem *extrakce významných bodů*.

EKG signál vykazuje vysokou míru variability. To je dáno širokým spektrem možných abnormalit a jejich individuální elektrokardiografickou manifestací u různých demografických kohort. Vizualní interpretace proto klade vysoké nároky na klinickou odbornost a čas hodnotitele. I přes zavedené standardy [3] je vizualní interpretace zatížena lidskou chybou, která může v krajním případě vést k mylnému závěru a nastavení nevhodné terapie.

Tyto důvody vedly k rozvoji nových výpočetních metod, jejichž cílem je automatizace analýzy EKG a extrakce klinicky konzistentních parametrů bez ohledu na komplexnost signálu. Obecně lze tyto algoritmy zařadit do dvou skupin: a) metody pro segmentaci EKG signálu a b) metody pro klasifikaci EKG do jedné či více diagnostických tříd. Obě skupiny metod se přitom z logických důvodů v různé míře vzájemně doplňují. Konvenční algoritmy sestávají z několika základních kroků [4], ve stručnosti: a) odstranění šumu pomocí filtračních technik; b) transformace dat za účelem dalšího zvýraznění vln; c) rozpoznání (detekce) jednotlivých vln; d) segmentace vln a extrakce klinických příznaků; d) přiřazení diagnostické třídy. Bod d) pak může být realizován jak pomocí heuristických pravidel, tak s využitím statistických metod a modelů [5].

U konvenčních metod (uvedených výše) je návrh celého řetězce založen na expertních znalostech o vlastnostech EKG signálu. Extrahované příznaky tak obvykle mají snadno interpretovatelný vztah s klinickými projevy srdeční činnosti. Neúplné lidské poznání v oblasti elektrokardiografie však v důsledku tuto tzv. manuální extrakci příznaků limituje, zanáší do analýzy další zkraslení a potenciálně také ztrátu informace. Výsledkem je vyšší míra nepřesností než u lidských hodnotitelů a tendence k podučení systému vlivem přílišné selektivity příznaků.

Oproti zmíněným přístupům jsou metody hlubokého učení [6] schopny získat (optimální) reprezentaci vysoce komplexních surových dat, a to na různých úrovních abstrakce. Slovo optimální uvádíme v závorce, jelikož se nejedná o optimalitu klinickou, nýbrž matematicky definovanou. Hluboké učení zahrnuje optimalizační algoritmy a komplexní nelineární modely, jejichž cílem je imitovat informační procesy probíhající v lidském mozku. Hluboké neuronální modely (DNN¹) patří mezi potenciální univerzální aproximátory [7] a v poslední dekádě dosáhly významných úspěchů v oblastech rozpoznání obrazu [8, 9], rozpoznávání řeči [10], predikce genové exprese [11], in-silico modelování léčiv [12], a řadě dalších. Z toho důvodu našlo hluboké učení cestu i do oblasti analýzy EKG signálů [13]. Metody umožňují automaticky extrahovat relevantní příznaky, čímž lze snížit vliv nevhodného předzpracování na kvalitu segmentace/klasifikace a případně identifikovat zcela nové vztahy mezi daty a klinickým obrazem pacienta. Současné DNN modely reprezentují rozhodovací proces založený řádově na milionech až miliardách vnitřních parametrů. Přes dosavadní pokrok v technikách interpretace [14] je porozumění komplexním vztahům uvnitř hlubokých modelů stále limitováno, čímž je limitována i důvěra v tyto systémy. Přeparametrizování modelu dále souvisí s problematikou přeučení modelu (aproximace šumu) a dosažení robustní generalizace. Současné metody hlubokého učení proto zpravidla vyžadují dostupnost adekvátního množství dostatečně reprezentativních dat [15]. V oblasti segmentace a klasifikace EKG jsou za tímto účelem obvykle využívány standardní veřejně dostupné databáze signálů. Tyto databáze však trpí některými významnými nedostatky jako je nedostatečný počet subjektů, silně asymetrické zastoupení diagnostickým tříd, chyby v klinicky stanovených referenčních hodnotách či referenční hodnoty stanovené jiným automatickým algoritmem s nízkou věrohodností.

¹Z anglického Deep Neural Networks

Cíle práce

Dizertační práce se zabývá návrhem metod hlubokého učení pro segmentaci elektrokardiografických záznamů. Pro nedostatečný stav řešení v této oblasti je segmentace zaměřena na úseky s výskytem elektrické aktivity srdečních síní v intrakardiálních elektrogramech (EGM) a standardním povrchovém EKG. V běžném ambulantním EKG záznamu je aktivita síní v množství elektrofyziologických scénářů prakticky neidentifikovatelná. EGM naopak poskytuje věrohodnou informaci o časové aktivaci lokalizované oblasti síně. Do DNN modelů je možné tuto znalostní bázi zabudovat ve formě vektorové reprezentace. Model lze následně přenést do oblasti úloh zaměřených na analýzu standardního EKG signálu snímaného z povrchu těla. Díky přímému vztahu vnořené vektorové reprezentace a elektrofyziologického obrazu můžeme předpokládat zlepšení interpretovatelnosti příznaků extrahovaných modelem na vyšších úrovních abstrakce.

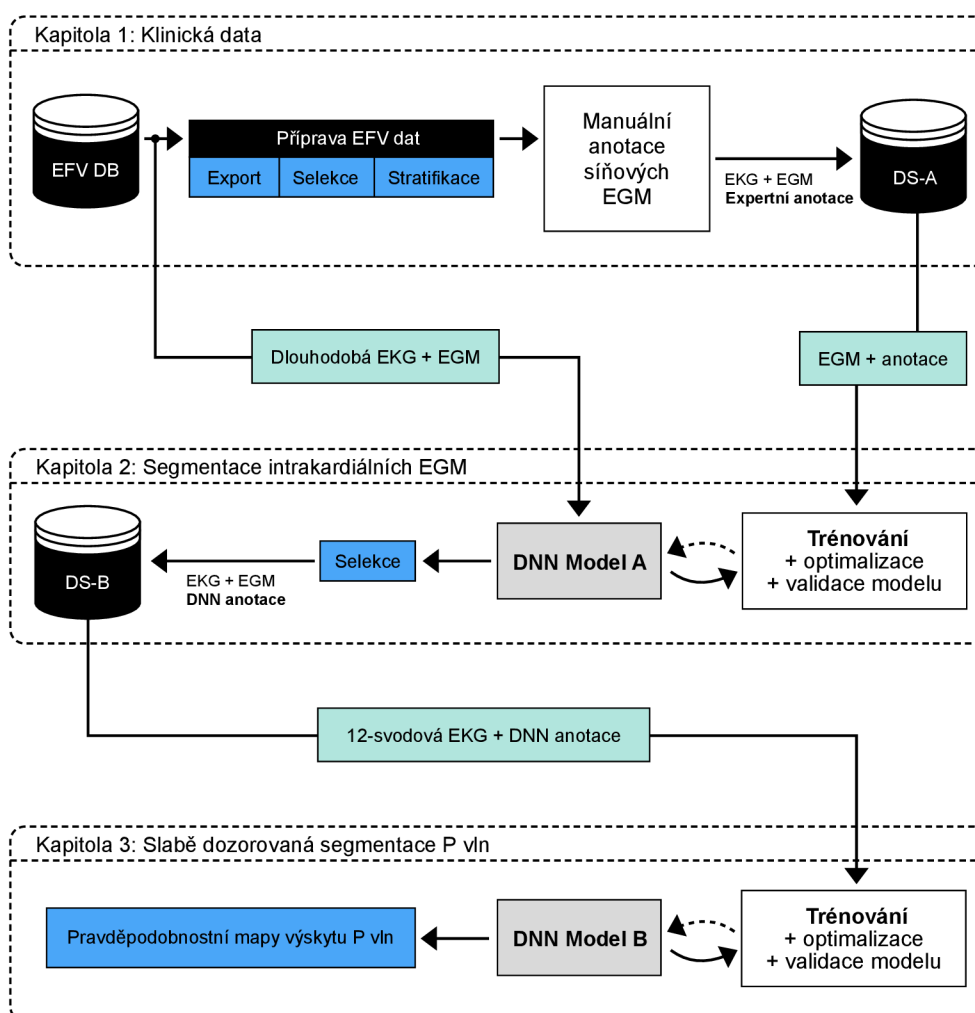
Za tímto účelem navrhujeme využít data z multimodálního snímání elektrické aktivity srdce v průběhu invazivního elektrofyziologického vyšetření (EFV). EFV absolvuje široké spektrum pacientů s poruchami srdečního rytmu a jeho součástí je jak snímání 12-svodového EKG, tak intrakardiálních elektrogramů (EGM). Tato data významně rozšiřují znalostní bázi a oproti existujícím EKG databázím pro vývoj DNN modelů poskytují zlatý standard při určení typu arytmogenního substrátu. Nízká míra nejistoty u takto stanovené diagnózy významně navyšuje kvalitu trénovacích dat a následně i DNN modelů. Kvalita klinických anotací v dozorovaném učení a schopnost vysvětlit výstup DNN modelu přitom patří mezi aktuální problémy prakticky ve všech oblastech hlubokého učení.

Cíle dizertační práce jsou založeny na automatizaci řetězce pro zpracování elektrokardiografických dat s cílem přispět k řešení výše zmíněných problémů, a jsou specifikovány následovně:

1. Návrh a implementace procesů pro manipulaci s EFV daty. Sestavení trénovacích databází s referenční anotací síňové aktivity, kvalitativní analýza databáze a stanovení limitací pro její využití.
2. Návrh a implementace metod hlubokého učení pro segmentaci síňové aktivity v intrakardiálních EGM s cílem automatizovat proces stanovení referenčních anotací v rozsáhlých datech. Optimalizace algoritmů pro zajištění neinferiority vůči současným metodám. Testování modelů vůči expertně stanoveným referenčním hodnotám.
3. Návrh a realizace metod hlubokého učení pro zabudování vektorové reprezentace síňové aktivity do DNN modelu na základě učebních dat s neúplnou informací. Kvantitativní a kvalitativní ověření DNN modelem generovaných příznakových map na EKG záznamech se sinusovým rytmem a za různých patofyziologických scénářů.

Organizace textu

Dizertační práce je rozdělena do tří kapitol, jejichž logická provázanost je znázorněna na Obrázku 1. V Kapitole 1 je popsán návrh softwarových nástrojů, které byly nezbytné pro zpracování, popis a další manipulaci s EFV daty. Následně je specifikována metodologie selekce pacientů, elektrofyziologických segmentů a jejich stratifikace do trénovacích databází DS-A a DS-B. Pro obě databáze byla provedena kvalitativní a kvantitativní analýza, ze které vyplývají některá nastavení algoritmů a limitace využití. DS-A obsahuje záznamy standardního 12-svodového EKG a intrakardiálních EGM z vybraných lokalit endokardu. V EGM záznamech bylo provedeno vizuální hodnocení a expertní stanovení počátku a konce síňové aktivace. Druhá datová sada obsahuje taktéž záznamy 12-svodového EKG a intrakardiálních EGM, avšak s významně vyšším zastoupením subjektů. Referenční intervaly síňové aktivity (DNN anotace) byly v tomto případě získány automatickou segmentací EGM pomocí DNN modelu A.



Obr. 1: Blokové znázornění pracovního postupu dizertační práce a provázanosti kapitol.

Kapitola 2 je věnována realizaci DNN modelu A pro segmentaci síňové aktivity v EGM s ohledem na různé elektrofyziologické scénáře. Hlavním cílem je automatizace řetězce pro interpretaci velkobjemových EGM dat z EFV. K naplnění cílů jsou využity EGM záznamy z expertně anotované sady DS-A. Začátek kapitoly je věnován stručnému přehledu současných metod zaměřených na segmentaci a detekci aktivačního času lokálních EGM. Kapitola zahrnuje návrh architektury DNN modelu A, jeho hyperparametrů a vybraných bloků optimalizačního procesu. Pomocí Bayesovské optimalizace a ablačních technik jsou prozkoumány vlastnosti procesu učení pro různá nastavení. Cílem analýzy je zajistit optimální generalizační schopnost DNN modelu A s ohledem na velikost a vlastnosti DS-A. Na závěr je provedeno hodnocení segmentace pro různé scénáře pomocí zavedených metrik kvality. DNN model s nejvyšší kvalitou segmentace je posléze využit pro automatické stanovení referenčních intervalů (DNN anotace) v početně rozsáhlejší DS-B.

Kapitola 3 je věnována segmentaci P vln v povrchových EKG signálech. Kapitola se zabývá problematikou učení s neúplnou informací a návrhem metodologie pro zabudování vektorové reprezentace síňové aktivity do DNN modelu B. Pro učení modelu byly použity povrchové svody a DNN anotace z datové sady DS-B. V této části byly navrženy modifikace vybraných technik hlubokého učení, primárně se zaměřením na penalizační funkce. Pro sérii experimentů bylo provedeno kvantitativní a kvalitativní hodnocení na jehož základě byla stanovena optimální strategie učení. Následně byly diskutovány možné přínosy DNN modelů se zabudovanou informací z EFV záznamů v automatické diagnostice poruch srdečního rytmu. Pro široký rozsah problematiky byl rozbor současného stavu poznání proveden pro každou z dílčích kapitol na začátku odpovídající sekce.

1 Zpracování klinických dat

1.1 Úvod

Míra generalizace, robustnost a efektivita dozorovaných modelů strojového učení je do značné míry omezena množstvím [15] a kvalitou [16] dostupných trénovacích dat. Oba parametry jsou přirozeně spojeny s vysokou časově-ekonomickou náročností; v medicínské oblasti pak navíc se zákonnými požadavky na soukromí pacienta [17]. Sběr dat může být tím pádem otázkou týdnů až měsíců s obdobně dlouhým časovým horizontem potřebným pro jejich následný klinický popis (anotaci) odborným hodnotitelem. Převážná většina prací [18, 19, 20, 21, 22, 23, 24, 25] zaměřených na automatizovanou interpretaci EKG proto pro strojové učení využívá některou ze standardních databází. Mezi nejznámější lze zařadit například *MIT-BIH Arrhythmia Database* [26] obsahující přes 4000 dlouhodobých holterovských měření, *PTB Diagnostic ECG Database* [27] zahrnující s rozšířenou verzí [28] dohromady přes 20 000 pacientů z různých populačních kohort, či některé další [29, 30, 31]. Zmíněné databáze obsahují řádově desítky tisíc referenčních záznamů kategorizovaných dle základních demografických ukazatelů a typu přítomné patologie. Obvykle jsou zahrnuty poruchy srdečního rytmu, strukturální i funkční nemoci myokardu. Méně početnou skupinu tvoří databáze, které poskytují referenční hodnoty klinicky významných intervalů vhodných pro vývoj segmentačních modelů. Mezi datové sady popisující síňovou aktivitu tak lze zařadit pouze *QT Database (QT-DB)* [32], *The Common Standards for Electrocardiography database (CSE)* [33], *MIT-BIH Arrhythmia Database P-Wave Annotations (MIT-BIH-ADPA)* [34], *Brno University of Technology ECG Signal Database (BUT-PDB)* [35] a *Lobachevsky University Electrocardiography Database (LUE-DB)* [36, 37]. Základní deksriptivní parametry těchto databází uvádíme v Tabulce 1.1.

Tab. 1.1: Přehled standardních databází EKG záznamů s referenční anotací P vlny. *Týká se pouze části EKG cyklů, většina anotována 14 různými automatickými algoritmy.

Název	Subjektů	Kanálů	Délka záznamu [s]	Hodnotitelů/Kardiologů
QT-DB [32]	105	2	900	2/2
CSE-DB [33]	125	15	10	*5/5
BUT-PDB [35]	50	2	120	2/0
MIT-BIH-ADP [34]	12	2	1800	2/0
LUE-DB [36, 37]	200	12	10	2/2

Standardní databáze poskytují možnost vzájemného srovnání segmentačních algoritmů na obdobných kohortách pacientů. Současně však disponují celou řadou problémů. Obecně lze zmínit především limitovaný počet subjektů a tím i malou reprezentativnost dat vzhledem k celkové populaci. Do QT-DB [32], CSE-DB [33] a LUE-DB [36, 37] byly zařazeny primárně fyziologické záznamy nebo záznamy, které mají minimální vliv na morfologii

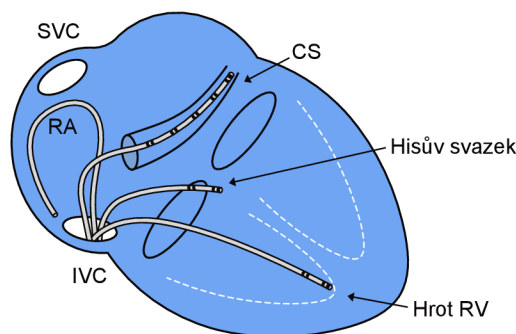
a časové uspořádání P vln. Potenciálně závažné choroby s relativně nízkou prevalencí, jako jsou např. tachykardie s širokým QRS komplexem, jsou reprezentovány spíše sporadicky. Tato extrémní nerovnoměrnost zastoupení diagnostických kategorií přináší řadu problémů z pohledu optimalizace modelu [38]. Důsledkem je snížená generalizační schopnost a významně nekonzistentní výsledky napříč různými kohortami pacientů, jak ukázaly rozsáhlé experimentální studie v [39]. Jen část databází je anotována zkušeným kardiologem [32, 33, 36, 37]. Zcela sporadicky je provedena nezávislá křížová validace několika hodnotitelů. Ani mezi zkušenými kardiology přitom nepanuje shoda na přesné lokalizaci hranic EKG segmentů či systému pravidel pro stanovení diagnostické třídy [40].

Kvalitu anotací může ovlivnit řada faktorů mezi nimiž lze zmínit například dosaženou odbornost, konfirmační zkreslení, únavu a chyby v percepci hodnotitele, nedostatek informační výtěžnosti dat či přítomnost atypických a vzácných případů. Proces anotace je tudíž přirozeným zdrojem šumu zaneseným do datové sady. Konzervativní odhad chybovosti diagnostických tříd stanovený pomocí autokorekčních algoritmů činí ve standardních databázích kolem 13 % [41]. U anotace síňové aktivity lze přitom předpokládat chybovost významně vyšší. Jedná se totiž o segment signálu s nízkým poměrem signál-šum, který je v řadě arytmiických scénářů ve standardním povrchovém EKG záznamu téměř neidentifikovatelný. Modely strojového učení jsou přitom známy dobrou schopností aproximovat výskyt náhodných chyb u anotací v trénovací sadě [42]. Negativními důsledky jsou vyšší nároky na množství trénovacích dat [43], snížení kvality predikce způsobené přeucením modelu, narušení stability extrahovaných příznaků a zkreslení rozložení pravděpodobnosti klasifikačních tříd [44, 16]. Anotační šum je přitom považován za potenciálně vyšší riziko pro model než konvenčně uvažovaný šum na straně surových dat či extrahovaných příznaků [45].

Rizika zmíněná ve spojitosti s chybnými anotacemi je možné snížit několika základními přístupy. Prvním z nich je již zmíněná identifikace a následná korekce potenciálně chybných anotací. Autokorekční algoritmy [46, 47, 41] obvykle využívají nedozorované metody učení pro výběr těch instancí (trénovacích vzorů) datové sady, které vykazují podobné pravděpodobnostní charakteristiky extrahovaných příznaků, avšak mají přiděleny sémanticky odlišné anotace. Instance jsou poté z procesu učení buď zcela vyloučeny nebo je jim přiřazena anotace ze skupiny trénovacích vzorů s nejvyšší mírou vzájemné shody příznaků. Problémem autokorekce je nízká kontrola nad procesem oprav a potenciální vyloučení méně obvyklých, nicméně validních případů. Další možností je využití syntetických EKG z virtuálních kohort s definovanou klinickou a demografickou charakteristikou. Genezi EKG záznamů je možné provést pomocí konvenčních parametrických metod [48] i s využitím pokročilých modelů jako jsou generativní kontradiktorní sítě [49] a numerické simulace elektrické aktivity srdečního orgánu [50, 51]. Přes slibné výsledky dosahují virtuální kohorty jen limitované variability generovaných příznaků elektrické aktivity ve vztahu k celopopulační charakteristice. Virtuální kohorty proto v současnosti nemohou plně nahradit sběr klinických dat.

1.2 Intervenční elektrofyziologické vyšetření

EFV je standardní procedurou u pacientů indikovaných k ablační terapii pro přítomnost poruch srdečního rytmu. V průběhu EFV jsou zavedeny multielektrodové katétry do předem definovaných lokalizací v pravé či levé části srdce. Mezi standardní lokalizaci patří koronární splav (CS) obíhající posteriorní povrch srdce mezi levou síní a levou komorou ústící v pravé síní v blízkosti dolní duté žíly. Další lokalizace pak mohou zahrnovat např. oblast Hisova svazku, strop pravé síně a další oblasti. Standardní postavení katétrů je uvedeno na Obrázku 1.1.



Obr. 1.1: Schématické znázornění standardního anatomického postavení diagnostických katétrů při EFV. CS – koronární splav; RA – pravá síní; SVC – horní dutá žíla; IVC – dolní dutá žíla; RV – pravá komora.

V průběhu vyšetření je snímáno nejen povrchové 12-svodové EKG, ale také intrakardiální elektrogramy (EGM) z uvedených katétrů. EGM poskytují informaci o časově-prostorovém uspořádání elektrické aktivity lokálních úseků myokardu, a to za různých elektrofyziologických podmínek. Jedná se o významné parametry pro charakterizaci elektrických vlastností tkáně a jejich interpretace umožňuje ve většině případů přímé klinické potvrzení původce obtíží. Katetrizační ablaci navíc podstupují jinak zdraví jedinci čistě s přítomností lokalizovaného arytmogenního substrátu i pacienti s funkčním postižením myokardu například ischemického původu, přítomností kardiomyopatie a souvisejícími komorbiditami. Jedná se tedy o širokou skupinu pacientů jejichž elektrokardiografická data by mohla významnou měrou přispět ke zvýšení kvality současných modelů strojového učení určených pro automatickou interpretaci EKG. EFV data byla dosud u zmíněných úloh využívána zcela minimálně. Dle znalosti autora byla dosud publikována jen jedna práce [52], která pracuje se záznamy jednoho subjektu.

1.3 Dílčí cíle

Tato část disertační práce zahrnuje dílčí cíle, které souvisejí s problematikou exportu, agregace a klinickým hodnocením elektrofyziologických signálů. Patří mezi ně primárně následující:

- Zpětná analýza souborových formátů s elektrofyziologickými daty.
- Implementace softwarových balíčků pro konverzi dat do otevřeného formátu.
- Selektce kohorty pacientů a sestavení datových sad.
- Kvalitativní analýza a klinická specifikace datových sad.
- Manuální anotace segmentů síňové aktivity v EGM záznamech.

1.4 Zpětná analýza dat

12-svodová EKG a intrakardiální EGM byly snímány za pomoci komerčního systému WorkMate 4.2 a 4.3 (Abbott). Systém WorkMate nativně ukládá data do proprietárních binárních formátů s nedokumentovanou strukturou. Dedikované nástroje pro export neumožňují dávkové zpracování a interně limitují rozsah exportovaných dat i přímý přístup k jejich surové verzi. Aby bylo možné plnohodnotně využít elektrofyziologické záznamy z celé procedury, byla provedena kvalitativní vysoko a nízkourovňová analýza proprietárního formátu s cílem identifikovat jeho strukturu a implementovat softwarové nástroje pro konverzi dat do otevřených formátů. Navržené nástroje jsou veřejně dostupné a kromě naplnění cílů této práce doufáme, že přispějí k širšímu využití EFV záznamů vědeckou komunitou.

1.4.1 Základní popis

Dle dokumentace jsou signály zaznamenány při vzorkovací frekvenci 2000 Hz a s rozlišením 78 nV/LSB ¹. Intrakardiální EGM jsou snímány unipolárně vůči Wilsonově svorce a bipolární inter-elektrodová montáž je provedena softwarově ex-post až při zobrazení svodů. Standardní interní předzpracování zahrnuje hardwarový filtr typu horní propust s mezní frekvencí ve volitelném rozmezí 0.01–1.0 Hz a pásmovou propust (50 Hz) pro potlačení síťového brumu. Charakteristiky těchto filtrů nejsou veřejně dokumentovány. Systém dále obsahuje volitelné softwarové filtry určené primárně pro zobrazení a export elektrokardiografických signálů. Nastavení těchto filtrů nemá vliv na uložená surová data.

Z dostupné dokumentace je patrné, že data jsou ukládána v relacích, tj. Událostí nebo uživatelsky řízeným spuštěním a ukončením záznamu. Každé relaci pak odpovídá právě jeden datový soubor. Minimální délka záznamu a tedy i velikost souboru je softwarově nastavitelná. Systém je schopen současně zaznamenat 256 a zobrazit až 448 kanálů rozdělených do 7 skupin po 64 kanálech.

¹Konvenční označení nejméně významného bitu z anglického Least Significant Bit

Délka slova pro uložení jednoho vzorku dat je 32 bitů. Základní souborová struktura elektrofyziologické studie standardně obsahuje jeden a více *.log souborů s úseky elektrofyziologických dat s doplňujícími informacemi v souborech *MASTER* a *ENTRIES*. Zmíněné informace umožňují alespoň semisémantickou analýzu datové struktury.

Pro sestavení minimální akceptovatelné dokumentace struktury byly využity vybrané postupy reverzního inženýrství. Ty lze obecně formulovat jako posloupnost kroků vedoucích k identifikaci hlavní struktury souborového formátu, rozdělení struktur dle typu rozsahu na statické a dynamické, přiřazení významů sekvencím ASCII² znaků, přiřazení významu a rozsahu adres dílčím strukturám na základě diferenční analýzy pro různá systémová nastavení. Výchozím krokem bylo stanovení množiny parametrů $P = \{p_1, p_2, \dots, p_i\}$, u nichž byl předpokládán funkční vztah k elektrofyziologickým datům, a na jejichž základě by bylo následně možné provést diferenční analýzu. Vzhledem k celkovému množství nastavitelných parametrů uvádíme pouze jejich obecnou kategorizaci:

1. Demografické a procedurální údaje.
2. Hardwarové parametry akvizice dat.
3. Počet a parametry nastavení snímaných kanálů.
4. Minimální délka záznamu.
5. Parametry klinických anotací.

Pro každý parametr byl specifikován uživatelsky nastavitelný rozsah a krok změny. Výsledkem byla sada experimentů, která umožnila provést diferenční analýzu mezi jednotlivými pokusy s následnou identifikací rozsahu adres analyzovaných parametrů v datové struktuře. Pro upřesnění demonstrujeme obecný postup analýzy pomocí Pseudokódu 2.

Algorithm 1 Postup zpětné analýzy

Vstup: Množina parametrů $P = \{p_1, p_2, \dots, p_i\}$

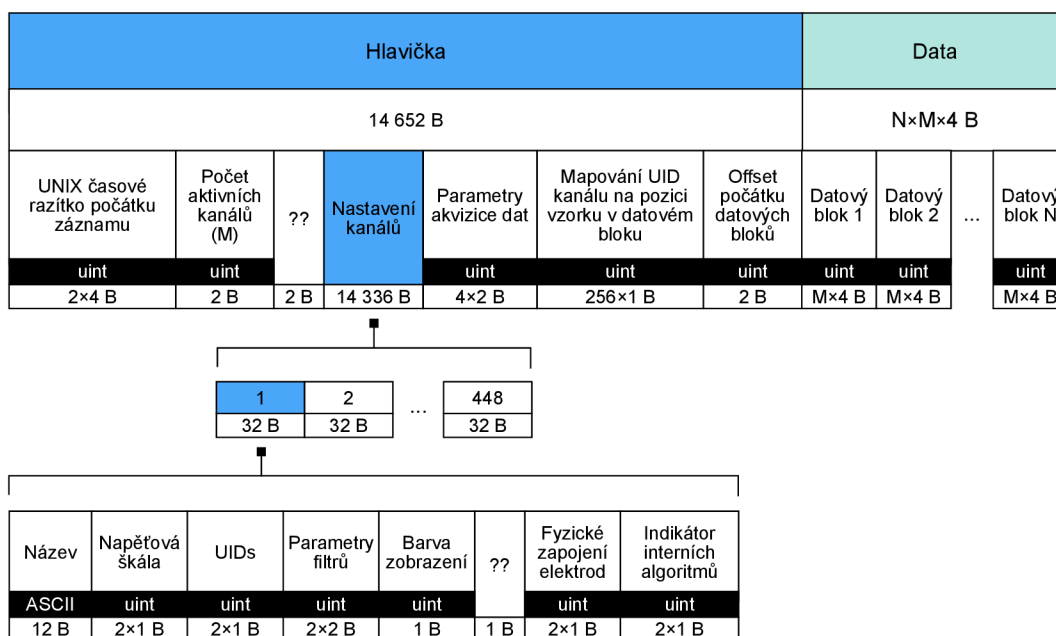
Výstup: Množina dvojic bit adres $A = \{\{a_1, b_1\}, \{a_2, b_2\}, \dots, \{a_i, b_i\}\}$ pro každý p_i

- 1: **for** každý parametr $p_i \in P$ **do**
 - 2: Inkrementuj hodnotu p_i
 - 3: Proveď akvizici dat
 - 4: Porovnej datovou strukturu souborů v kroku p_i a p_{i-1}
 - 5: **if** existuje rozdíl mezi p_i a p_{i-1} **then**
 - 6: Identifikuj rozsah adres $\{a_i, b_i\}$ odpovídající nalezenému rozdílu
 - 7: $A \leftarrow \{a_i, b_i\}$
 - 8: **end if**
 - 9: **end for**
 - 10: Proveď agregaci potenciálních adres pro každý parametr
 - 11: Ověř správnost stanoveného na retrospektivních datech
-

²Z anglického American Standard Code for Information Interchange

1.4.2 Výsledky

Na Obrázcích 1.2 a 1.3 jsou uvedena schémata identifikovaných bloků s rozsahy bitových adres pro nalezené parametry P . V prvním zmíněném schématu je uvedena specifikace datového souboru **.log* se záznamem elektrokardiografických signálů. Samotný datový blok je jediným dynamickým segmentem, jehož struktura závisí na počtu a elektrické montáži zaznamenávaných kanálů. U kanálů s bipolární montáží je typicky k dispozici dvojice unikátních identifikátorů (UID³), které mapují relativní adresu odpovídající dvojice unipolárně snímaných surových vzorků signálu v datovém bloku (*chunk*). Každý vzorek signálu je uložen jako 32 bitové celé číslo. Pro konverzi do reálného signálu je třeba datový blok dekódovat pomocí dvojkového doplňku. Z pohledu parametrů signálu je informačně významným segmentem část hlavičky obsahující základní nastavení akviziční jednotky (parametry akvizice dat), konkrétně vzorkovací frekvence, napěťové rozlišení signálu, mezní frekvence hardwarového filtru typu horní propust a typ pásmové zadržky pro potlačení síťového brumu.

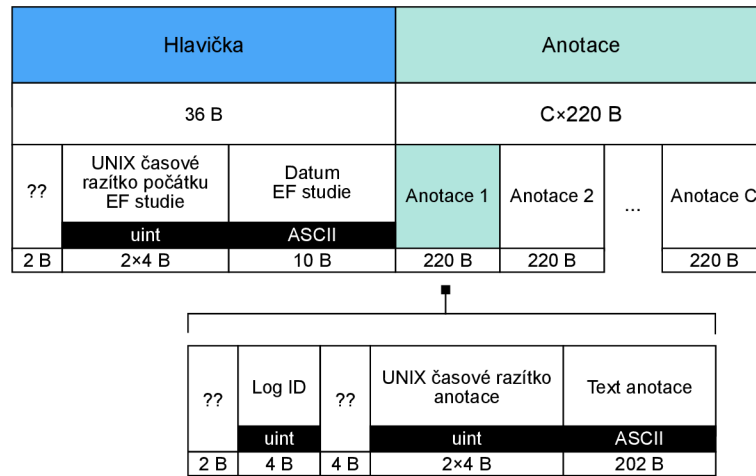


Obr. 1.2: Schéma struktury nativního datového formátu pro archivaci signálů **.log* na platformě WorkMate.

Na Obrázku 1.3 je dokumentováno schéma identifikovaných bloků binárního souboru *ENTRIES*. Do souboru jsou zaznamenány procedurální informace technického a klinického charakteru jako jsou parametry stimulačních protokolů, parametry ablační terapie, automaticky identifikované události, parametry elektrofyziologických projevů srdce a diagnostické informace stanovené expertním týmem EFV pracoviště.

³Z anglického Unique Identifier

Každá anotace má přiřazeno časové razítko a unikátní identifikátor **.log* souboru, ke kterému se vztahuje. Jedná se tedy o významný zdroj informací pro retrospektivní analýzu a interpretaci EKG a EGM záznamů.



Obr. 1.3: Schéma struktury binárního souboru pro archivaci anotací průběhu a klinických nálezů v průběhu EFV.

1.5 Softwarové balíčky pro manipulaci s EFV daty

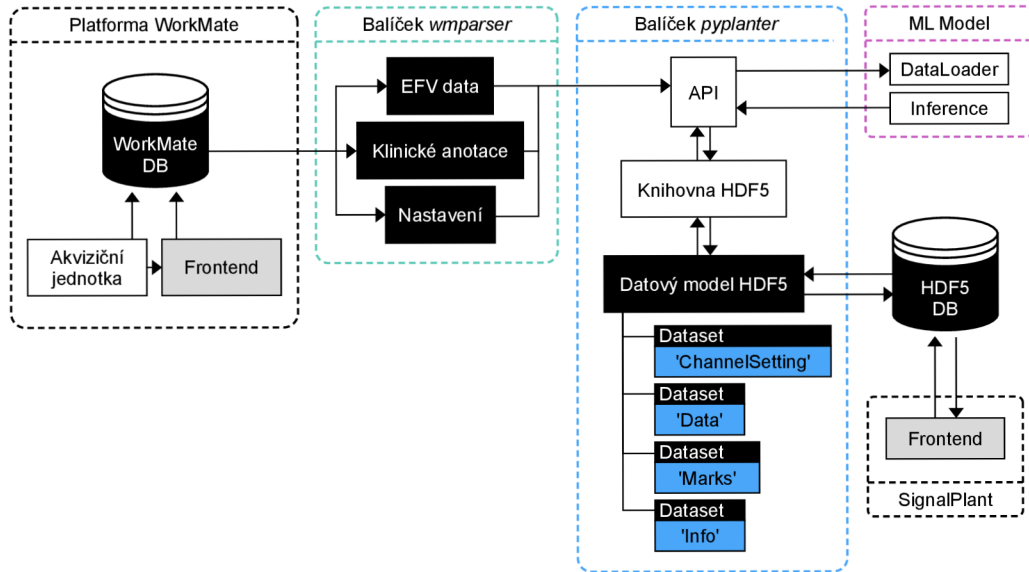
Tato sekce stručně pojednává o architektuře systému pro zpracování EFV dat. Systém vznikl pro potřeby dosažení cílů této disertační práce a zahrnuje dvojici aplikačních (API⁴) rozhraní *wmparser* a *pyplanter*. Hlavním účelem navrženého systému je podpora při tvorbě a klinické anotaci učebních datových sad, uvedených v navazujících sekcích této kapitoly.

Balíček *wmparser* poskytuje API pro dávkový export kompletních surových elektrofyziologických dat z platformy WorkMate. Balíček obsahuje moduly pro parsování dedikovaných souborových formátů, validaci a anonymizaci dat, či manipulaci s kanály včetně uživatelem definované elektrické montáže. Modul pro parsování vychází z výsledků uvedených v kap. 1.4, a v tuto chvíli umožňuje export EKG a EGM záznamů, klinických anotací, parametrů akvizice a základních procedurálních údajů.

Balíček *pyplanter* poskytuje API pro konverzi dat do otevřeného datového formátu a jejich následnou manipulaci. API umožňuje programovou manipulaci s datovými poli, editaci kanálů a klinických anotací, anotacemi podmíněnou selekci datových segmentů a další operace. Datový model byl postaven na hierarchickém souborovém formátu a knihovně *Hierarchical Data Format* verze 5 (HDF5®) [53], který splňuje základní požadavky na rychlost čtení a zápisu pro úlohy strojového učení, širokou podporu napříč programovacími jazyky, existenci volně dostupných prohlížečů a možnost blokového čtení a zápisu dat.

⁴Z anglického Application Programming Interface

Jako výchozí rozhraní je využit balíček *h5py* verze 3.1, nad které bylo vystaveno vlastní aplikační rozhraní, které standardizuje datový model do režimu kompatibility s volně šiřitelnou platformou pro vizuální analýzu biologických signálů SignalPlant (AV ČR) [54]. Platforma umožňuje standardizovanou vizualizaci a manuální značení, což je nezbytný požadavek ze strany klinických hodnotitelů. Blokové schéma komponent pro manipulaci s EFV daty a jejich provázanosti je uvedeno na Obrázku 1.4.



Obr. 1.4: Blokové schéma komponent pro manipulaci s EFV daty. ML Model – rozhraní procesů strojového učení pro práci s daty.

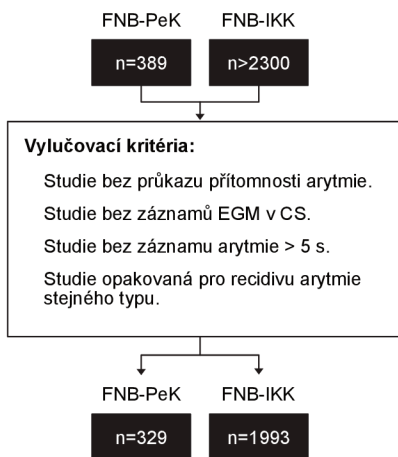
1.6 Učební datové sady

Následující text pojednává o metodologii selekce pacientů, výběru zájmových oblastí EKG a EGM signálů a následného stanovení klinických referenčních hodnot (anotaci) pro potřeby strojového učení. Dále byla provedena kvalitativní analýza datových sad a diskuze jejich možného využití. Data byla dávkově získána pomocí nástrojů uvedených v kap. 1.5.

1.6.1 Charakteristika zařazených pacientů

Jedná se o retrospektivní elektrokardiografická data, která byla měřena u pacientů indikovaných k EFV ve FN Brno. Všichni pacienti podepsali Informovaný souhlas se zpracováním klinických dat a využití dat pro účely této práce bylo schváleno etickou komisí FN Brno 05-090920/EK.

Do datové sady byly zařazeny 2 kohorty pacientů. Kohorta dětí a mladých dospělých obsahuje konsektivní záznamy 329 pacientů ve věku 6–19 let (48.2 % žen), kteří podstoupily EFV mezi roky 2012–2021 na pracovišti Pediatrické kliniky (FNB-PeK). Jedná se primárně o zdravé jedince bez přidružených kardiovaskulárních onemocnění. Druhá kohorta obsahuje 1993 konsektivních záznamů pacientů (z toho 32 % žen) ve středním věku v rozsahu 18–83 let a 25 mladých dospělých do 18 let, kteří podstoupily EFV v rozmezí let 2014–2022 na pracovišti Interní kardiologické kliniky (FNB-IKK) FN Brno. Úvodní selekce pacientů byla provedena pomocí semi-automatického a vizuálního hodnocení na základě vylučovacích kritérií uvedených na Obrázku 1.5.



Obr. 1.5: Vylučovací kritéria pro selekci EFV záznamů.

1.6.2 Selekcce zájmových úseků

U všech zařazených studií byla provedena vizuální inspekce EKG a EGM záznamů s následným výběrem zájmových úseků. Na tomto procesu se podílel autor této práce a kliničtí

inženýři z týmu Intervenční srdeční elektrofyziologie Nemocnice u Sv. Anny. Všichni hodnotitelé mají 2 a více let praxe s hodnocením EKG a EGM záznamů v elektrofyziologické laboratoři. Klinická interpretace zájmových úseků vycházela z elektrofyziologických nálezů potvrzených operujícím kardiologem.

Pro zařazenou studii byl proveden výběr 1–5 nepřekrývajících se segmentů. Přesná délka segmentu nebyla pro variabilitu v délce záznamů stanovena, orientační doporučená byla 10 s. Cílem výběru bylo maximalizovat klinické zastoupení variabilních typů poruch srdečního rytmu. Priorita výběru byla určena následovně:

1. Záznam normálního sinusového rytmu, pokud je přítomen.
2. Záznam klinické tachykardie, pro kterou byl pacient indikován k EFV.
3. Záznamy s atypickou aktivací síní.
4. Procedurálně indukované poruchy srdečního rytmu nesouvisející s primárním substrátem.
5. Záznam s řízenou stimulací síní či komor.

Obecně delší segmenty byly voleny pro záznamy s běžící fibrilací síní kvůli rozsáhlé variabilitě v morfologii EGM signálů. V případě dostupnosti byly preferovány segmenty přechodu mezi arytmií a sinusovým rytmem, případně mezi dvěma a více typy arytmií. Zahrnuty byly také segmenty s řízenou stimulací síní a komor, které mohou stran přítomnosti stimulačních artefaktů a posloupnosti vln imitovat záznamy u pacientů s implantovaným kardiostimulátorem.

Každý segment byl exportován jako samostatný HDF5® soubor obsahující záznam 12-svodového EKG, 5-svodového bipolárního EGM z katétru v CS a 2-svodový bipolární záznam z katétru umístěného v oblasti Hisova svazku, pokud byl přítomen. Klinická interpretace segmentu byla kódována dle standardu SNOMED CT®.

1.6.3 Expertně anotovaná datová sada a (DS-A)

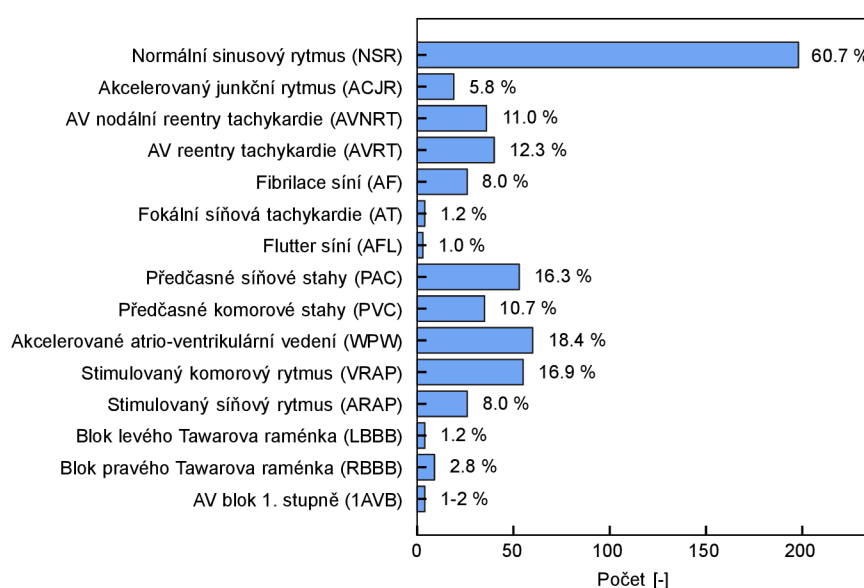
Účelem DS-A je podpora při učení systémů pro rozpoznání síňové aktivity v EGM záznamech. Sestavení DS-A slouží k dosažení hlavního cíle disertační práce č. 2. Do DS-A byly zařazeny EFV záznamy od 100 po sobě jdoucích pacientů z kohorty dětí a mladých dospělých na pracovišti FNB-PeK. Počet byl limitován časovou náročností manuální anotace. Z důvodu nedostupnosti dat v době tvorby DS-A nebyl zařazen žádný pacient středního věku, což je jednou z hlavních limitací této datové sady. Jak již bylo zmíněno, jedná se primárně o skupinu zdravých jedinců bez přidružených kardiovaskulárních onemocnění s přibližně rovnoměrným zastoupením mužů a žen. Celkem bylo vybráno 326 zájmových segmentů. Základní charakteristiku DS-A uvádíme v Tabulce 1.2.

Na Obrázku 1.6 je uvedeno relativní zastoupení diagnostických tříd zařazených segmentů. Každý segment může obsahovat jednu a více tříd. Z toho důvodu může více segmentů od stejného pacienta obsahovat stejnou diagnostickou třídu. Nejčastěji (60.7 %) jsou zastoupeny záznamy s normálním sinusovým rytmem (NSR). Z klinicky relevantních poruch srdečního rytmu jsou zastoupeny primárně tachykardie typické pro tuto kohortu pacientu

Tab. 1.2: Základní charakteristika datové sady DS-A.

Název parametru	$n = 100$
Věk [roky]	14.0 (12.0–17.0)
Pohlaví: ženy	48 (48 %)
Počet záznamů	326
Délka záznamu [s]	8.5 (6.4–12.2)

– AV nodální reentry tachikardie (AVNRT) a AV reentry tachykardie (AVRT) s případným obrazem akcelerovaného atrio-ventrikulárního vedení (WPW⁵). Zbývající poruchy rytmu obvykle souvisejí s procedurálními úkony a nejsou primárním důvodem indikace k EFV.



Obr. 1.6: Zastoupení diagnostických tříd v DS-A.

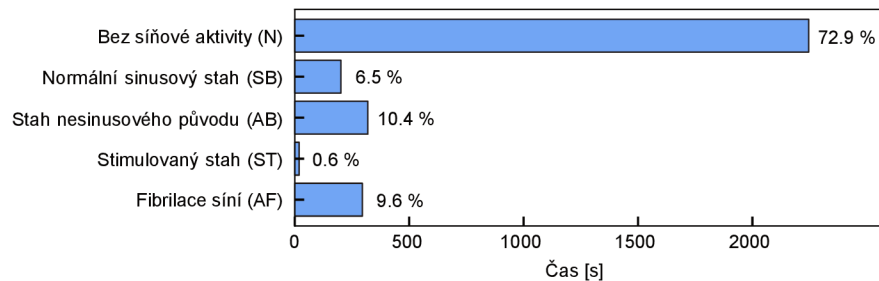
V aplikaci SignalPlant v. 1.2.7.8 byla následně provedena vizualizace a manuální anotace aktivačního času síňové aktivity v 5-kanálových bipolárních signálech snímaných v CS. EGM signály byly zobrazeny v jednotném napětovém měřítku s předzpracováním pomocí filtru s konečnou impulzní charakteristikou (FIR⁶) typu pásmová propust. Mezní frekvence filtru byly nastaveny dle standardních doporučených hodnot pro vizuální hodnocení EGM v rozmezí 30 a 250 Hz [55]. Za počátek síňové aktivity byl považován nejčasnější lokální aktivační čas v rámci všech 5 kanálů. Za konec síňové aktivity byl považován konec elektrické aktivity časově nejvzdálenějšího lokálního elektrogramu. Výjimku tvořili záznamy s fibrilací síní, které vykazují vysokou míru náhodnosti a přesný počátek a konec v rámci více svodů obvykle nelze stanovit. Z tohoto důvodu byl jako počátek a konec síňové aktivity

⁵Wolff-Parkinson-White syndrom

⁶Z anglického Finite Impulse Response

u signálu s fibrilací síní označen počátek a konec samotné arytmie. Záměrem tohoto anotčního přístupu je mj. Využití referenčních pozic pro přesné stanovení vzniku a terminace fibrilace síní z povrchového EKG signálu. V případě sloučených síňových a komorových stahů byly hranice síňové aktivity odhadnuty na základě morfologie elektrogramů při sinusovém rytmu.

Každý manuálně označený úsek síňové aktivity byl zařazen do jedné ze 4 klasifikačních tříd: a) normální sinusový síňový stah (SB); b) abnormální síňový stah (AB); c) fibrilace síní (AF) a d) stimulovaný síňový stah (ST). Do kategorie AB patří všechny síňové stahy, které s největší pravděpodobností vznikly mimo sino-atriální uzel. Do kategorie ST byly zařazeny ty síňové stahy, které vznikly jako reakce na řízenou stimulaci síní. Absolutní a relativní časové zastoupení jednotlivých typů síňových elektrogramů ve vztahu k celkové době záznamů uvádíme na Obrázku 1.7.



Obr. 1.7: Časové zastoupení jednotlivých typů síňových stahů v DS-A. Hodnoty vychází z referenčních úseků stanovených manuální anotací.

V případě kategorie SB, AB a AF je časové zastoupení v datasetu z pohledu optimalizačních úloh přibližně stejné a pohybuje se v rozmezí 6.5–10.4 % celkové délky záznamů. Zbývajících 72.9 % náleží úseků bez elektrické aktivity, či v malém rozsahu aktivitě komor.

Při vyloučení AF segmentů trvá jeden síňový stah průměrně 59 ± 17 ms. Nejkratší a nejdelší délka trvání odpovídá hodnotám 19, respektive 173 ms. V obou případech se jedná o významně odlehlé hodnoty, které odpovídají atypické časové aktivaci myokardu napříč elektrodami katétru. Průměrná maximální amplituda SB, AB a SF stahů činí 1.85 ± 1.94 mV. 90 % vzorků síňových signálů však spadá do rozmezí -0.789–0.790 mV. U stimulačního artefaktu indukovaného při stimulaci do některého z CS kanálů může tato hodnota dosahovat až 200 mV.

1.6.4 Automaticky anotovaná datová sada B (DS-B)

Sada DS-B vznikla a kontinuálně se rozvíjí jako rozsáhlá databáze arytmiických záznamů s anotovanou elektrickou aktivitou síní. DS-B slouží k dosažení hlavního cíle disertační práce č. 3. Selektce pacientů a zájmových segmentů proběhla stejným způsobem jako v Sekci 1.6.3. Databáze rovněž obsahuje 12-svodové povrchové EKG záznamy a intrakardiální EGM z katétru v CS. Na rozdíl od DS-A jsou segmenty síňové aktivity označeny

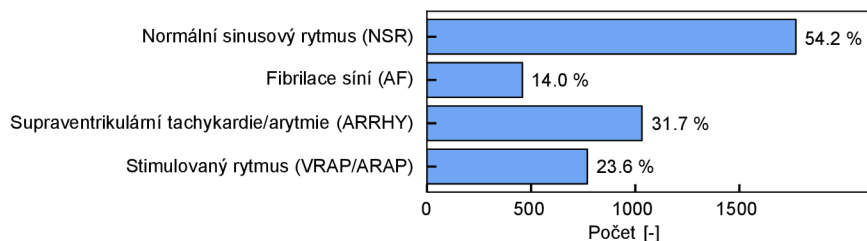
automaticky pomocí neuronálního modelu, který je navržen a popsán v Kapitole 2. Učení modelu bylo provedeno na DS-A. Důvodem k tomuto kroku je sestavení takové datové sady, která obsáhne pokud možno všechny dostupné EFV záznamy. Manuální anotace takového množství záznamů je z ekonomických důvodů prakticky neuskutečnitelná. Automatické anotace byly podrobeny zevrubné vizuální kontrole při výběru zájmových segmentů a následně také post-hoc analýzou pomocí semi-automatických metod založených na variabilitě rytmu.

Z celkového počtu 2322 studií se podařilo do současné doby do DS-B zařadit celkem 708 pacientů a 3265 individuálních EKG a EGM segmentů. Zařazení pacienti reprezentují jak kohortu mladých dospělých, tak kohortu pacientů středního věku. Jedná se tedy o průřez populací, která byla indikována k EFV. Základní charakteristiku DS-B uvádíme v Tabulce 1.3.

Tab. 1.3: Základní charakteristika datové sady DS-B

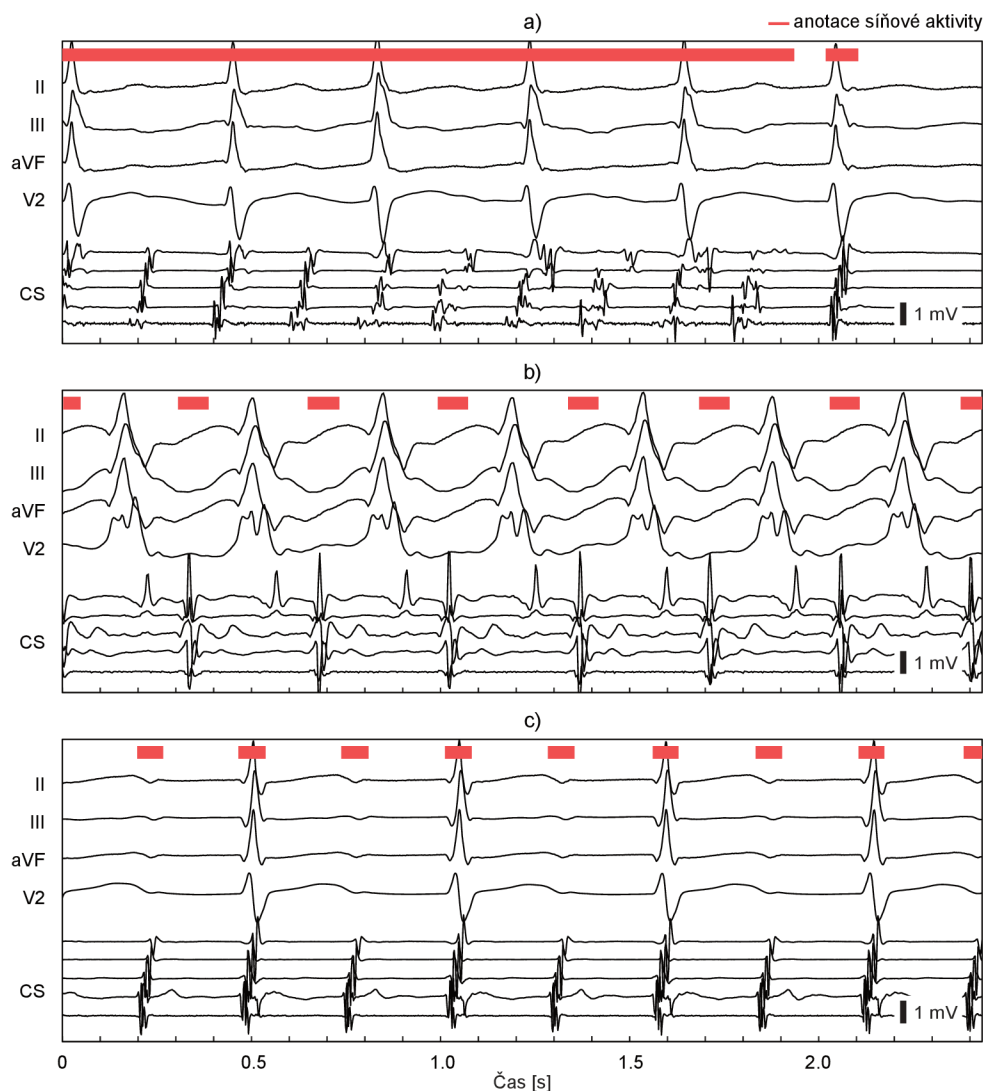
Název parametru	$n = 708$
Věk [roky]	36.6 (15.0–58.0)
Pohlaví: ženy	295 (41.7 %)
Počet záznamů	3265
Délka záznamu [s]	11.0 (8.7–15.4)

Z celkového počtu 3265 zájmových segmentů obsahuje 1771 záznamů sinusový rytmus (NSR), 1036 záznamů některý z typů supraventrikulárních tachykardií (ARRHY) včetně síňových tachykardií, 458 záznamů fibrilaci síní (AF) a 770 záznamů řízenou stimulaci síní či komor. Záznamy NSR mohou obsahovat další časově lokalizované poruchy rytmu jako jsou síňové a komorové extrasystoly, které však nejsou zatím v DS-B popsány. Zastoupení diagnostických tříd zařazených segmentů je uvedeno na Obrázku 1.8. Oproti DS-A jsou významně více (14.0 %) přítomny záznamy s AF, které bývají obvyklým zdrojem selhání algoritmů pro detekci P vlny. Záznamy s atypickou síňovou aktivitou, a síňovou aktivitou skrytou v QRST komplexu, jsou reprezentovány primárně skupinou ARRHY. Ta zahrnuje širokou škálu arytmií obdobně jako v Obrázku 1.6 s distribucí diagnostických tříd vychýlenou směrem k poruchám typickým pro dospělou populaci.



Obr. 1.8: Zastoupení obecných diagnostických tříd v DS-B.

Na Obrázku 1.9 ilustrujeme typické arytmiické záznamy v DS-A a DS-B. Z 12-svodového EKG je pro názornost zobrazen výběr svodů II, III, aVF a V2. Ve všech ukázkách je patrná variabilita morfologie signálů jak s ohledem na klinickou diagnózu, tak časové uspořádání síňové aktivity. V horní části je ukázka záznamu s fibrilací síní, která má relativně pravidelný převod na komory a občasně se vyskytuje organizovaná síňová aktivita. Ve střední části obrázku je záznam s širokokomplexou tachykardií, která může být snadno zaměněna za komorovou tachykardii. P vlny na povrchovém EKG jsou v tomto případě skryty v T vlně. Ve spodní části obrázku je záznam AVNRT, u které dochází k atypickému poměru převodu síňového vzruchu na komory. Na povrchovém EKG lze proto s jistotou tachykardii diagnostikovat jen v případě správné identifikace síňové aktivity. Pro automatické algoritmy je interpretace obdobných záznamů výzvou a často příčinou selhání.



Obr. 1.9: Ukázka typických záznamů v DS-A a DS-B s anotací segmentů síňové aktivace: a) fibrilace síní; b) antidromní AVRT; c) AVNRT.

1.7 Diskuze a limitace

Obě navržené datové sady obsahují záznamy s širokou škálou poruch srdečního rytmu a relativně vysoký počet individuálních subjektů. Obdobná klinická rozmanitost záznamů není v žádné ze současných standardních EKG databází k dispozici, viz Tabulka 1.1. Dle znalosti autora práce se jedná o vůbec první databázi, která obsahuje anotace P vln skrytých v QRST komplexu. Referenční pozice síňové aktivity jsou stanoveny s vysokou mírou spolehlivosti oproti vizuální interpretaci 12-svodového EKG. To platí jak pro manuálně, tak pro automaticky stanovené pozice (viz výsledky v Kapitole 3). Záznamy v DS-A i DS-B jsou kategorizovány do nadstandardního množství diagnostických tříd na základě elektrofyziologickou studií potvrzeného arytmogenního substrátu. Takto získané klinické údaje lze považovat za více relevantní než interpretaci povrchového EKG, na které jsou založeny všechny standardní EKG databáze.

Přes nespornou informační výtěžnost trpí data několika nedostatky, které je při jejich využití nutno vzít v potaz. Vzhledem k tomu, že byli zařazeni pouze pacienti indikovaní k EFV, se jedná o selektivní kohortu pacientů. Další selektivní zkreslení bylo zavedeno způsobem výběru zájmových segmentů, který se zaměřoval primárně na arytmiické záznamy. Databáze tudíž nereprezentují skutečné zastoupení diagnóz v populaci a algoritmy na nich učené mohou nadhodnocovat skutečné riziko přítomnosti poruchy srdečního rytmu u zdravé populace. Za zdravou populaci lze sice považovat kohortu dětských pacientů, ta je však zastoupena v menší míře. Morfologie 12-svodového EKG u dětí bývá navíc pro jiné anatomické poměry odlišná od dospělých jedinců [56]. Některé dokumentované nálezy poruch srdečního rytmu (např. fibrilace síní a blokády Tawarových ramének u dětských pacientů) nemají původ ve strukturálním postižení myokardu, ale mohou být důsledkem procedurálních postupů EFV. Morfologické příznaky (i dosud neznámé) EKG i EGM signálu z toho důvodu nemusí zcela odpovídat nálezům u pacientů, u nichž je daný typ arytmie primární indikací k EFV. Zařazení pacienti pocházejí ze dvou spádových center. Do budoucna by proto bylo vhodné datové sady rozšířit o pacienty z dalších evropských regionů.

Dalším významným problémem může být přenositelnost algoritmů učených na datech z EFV na jiné, zpravidla ambulantní platformy. Z technických důvodů se v průběhu EFV využívá odlišné anatomické umístění povrchových elektrod pro snímání EKG než u ambulatorních záznamníků. Odlišné umístění však může být zdrojem morfologických změn v signálu s dopadem na některé typické elektrokardiografické markery srdečních onemocnění [57]. U DNN modelů je možné tento efekt do určité míry eliminovat vhodnou augmentací EKG signálů či doučením modelu na datech s odlišnou charakteristikou.

Z pohledu návrhu automatických algoritmů mají datové sady několik obecných nedostatků, které mohou negativně ovlivnit proces optimalizace i interpretaci výsledků. V první řadě je třeba zmínit, že anotace síňových EGM nereprezentují celou dobu trvání elektrické aktivace obou síní, což je patrné i v deskriptivním popisu DS-A. EGM z oblasti CS představuje v závislosti na umístění katétru přibližně časovou aktivaci levé síně. Při normálním sinusovém rytmu se jedná o druhou polovinu časového trvání P vlny. Hrubý odhad počátku

P vlny je možné provést jen na základě znalosti její průměrné délky. V případě ostatních rytmů může být anotace umístěna jak v počáteční, tak konečné fázi síňového cyklu. Z pohledu detekce síňové aktivity je proto obvykle nutné stanovit časové okno, v rámci kterého bude predikovaná pozice ve vztahu k referenční anotaci akceptována. V případě segmentačních technik se podobným problémem zabývají metody učení s neúplnou informací. s ohledem na zastoupení diagnostických tříd a jednotlivých typů síňových rytmů je zřejmé, že jsou obě datové sady zatíženy nerovnoměrným zastoupením vstupní a výstupní informace, tzv. datovou imbalance. Vstupní imbalance je dána krátkými časovými intervaly, které odpovídají aktivaci síní, ve vztahu k ostatním EKG komplexům a segmentům bez elektrické aktivity. Výstupní imbalance je způsobena silně asymetrickým zastoupením jednotlivých diagnostických tříd, a tím i různě obtížných EKG záznamů pro interpretaci. Oba jevy mají negativní dopad na současné optimalizační metody a je třeba je zohlednit jak při návrhu i validaci automatických algoritmů.

Na závěr je třeba zmínit, že síňové anotace ani přiřazené diagnostické třídy zatím neprošly křížovou validací. U referenčních segmentů v DS-A proto není možné stanovit interindividuální variabilitu expertního stanovení jejich počátku a konce. Z tohoto důvodu není pro DS-A možné určit ani specifickou hranici optimality rozměření. U diagnostických tříd byly dodrženy standardní postupy pro hodnocení dat z EFV studií. Proto lze předpokládat i vysokou míru spolehlivosti jejich stanovení. Do DS-B bylo z časových důvodů zařazeno jen 708 pacientů a klasifikace EKG záznamů byla provedena oproti DS-A do 4 základních diagnostických tříd. Zvolené kategorie by však měly být dostatečné pro snížení vlivu výstupní imbalance zmíněné výše. DS-B se navíc dále rozšiřuje o nové subjekty i přiřazené diagnostické třídy.

Přes uvedené nedostatky je při vhodně zvolené metodologii možné využít DS-A i DS-B pro učení či validaci automatických algoritmů pro interpretaci 12-svodových EKG záznamů. Anotované intrakardiální signály mohou sloužit k vývoji algoritmů pro segmentaci EGM a metod pro potlačení signálu vzdáleného elektromagnetického pole. Řešení obou zmíněných problémů nachází uplatnění při vytváření 3D modelů elektro-anatomických map u komplexních ablačních procedur. Referenční pozice síňové aktivity mohou potenciálně nalézt uplatnění při vývoji slabě dozorovaných metod separace zdrojů pro vícesvodové EKG záznamy.

2 Segmentace síňových elektrogramů pomocí modelů hlubokého učení

2.1 Současný stav

Katétrová ablace patří mezi vysoce efektivní terapii supraventrikulárních a komorových tachykardií s trvajícím klinickým efektem a zlepšením indikátorů kvality života [58]. Nedílnou součástí procedury jsou specifické protokoly programované srdeční stimulace a snímání lokálních EGM pomocí multielektrokových katétrů zavedených v srdečních dutinách. Analýza časově-prostorové distribuce EGM v průběhu spontánního šíření arytmie, v době vzniku a terminace či jako odpověď na programovanou stimulaci umožňuje charakterizaci elektrofyziologických vlastností tkáně, mechanismu arytmie a potenciální lokace substrátu pro ablaci [59, 60].

Mezi základní techniky, umožňující identifikaci arytmogenního substrátu, patří měření lokální časové aktivace (LAT¹). Ta reprezentuje časový okamžik elektrické aktivace tkáně v oblasti snímací elektrody v rámci síňové (A) či komorové (V) aktivity. Analýza LAT u multipolárních signálů umožňuje odvodit další parametry jako jsou rychlost a směr šíření aktivační vlny, post-stimulační intervaly, restituční křivky tkáně či frekvence a časová distribuce srdečních cyklů. Tyto klinické parametry nesou kvantitativní informaci o vlastnostech myokardu a jsou základním vodítkem pro lokalizaci re-entry okruhů a fokálních ložisek u supraventrikulárních tachykardií [61, 62, 63, 64, 65]. Reprodukovatelné stanovení přesného počátku LAT je základním požadavkem při vytváření aktivačních map pomocí elektroanatomického mapování arytmí.

Při unipolárním snímání EGM je počátek LAT tkáně spojován s časem minima první derivace signálu lokálního elektrogramu. Tento vztah byl odvozen z extracelulárního měření akčního napětí na buněčných modelech [66, 67]. Zásadní limitací unipolárního záznamu je citlivost na vzdálené elektromagnetické pole generované odlehlou srdeční svalovinou, tzv. "far-field". Bipolární snímání sice umožňuje tento efekt do značné míry potlačit, avšak za cenu zkreslení EGM vlivem prostorové geometrie elektrod a jejich orientaci vůči aktivační vlně [68]. Metoda založená na derivaci signálu je přesto využívána i v případě bipolárních EGM.

Odlišný koncept anotace LAT volí metody založené na morfologických vlastnostech lokálního EGM. Základním principem je nalezení významného a snadno identifikovatelného vzorku signálu. Mezi nejjednodušší a konvenčně využívané přístupy patří detekce maxima surového EGM a odhad barycentra hodnot napětí [69]. Tyto metody opakovaně dosáhli nižší průměrné odchylky a dlouhodobé reprodukovatelnosti ve srovnání s expertními anotacemi než LAT určený na základě minima první derivace [69, 70, 71]. V obou případech je nutné předchozí vymezení segmentu s cílem eliminovat signály a artefakty, které nejsou předmětem zájmu. Semi-automatické přístupy spočívají v manuálním stanovení časového

¹Z anglického Local Activation Time

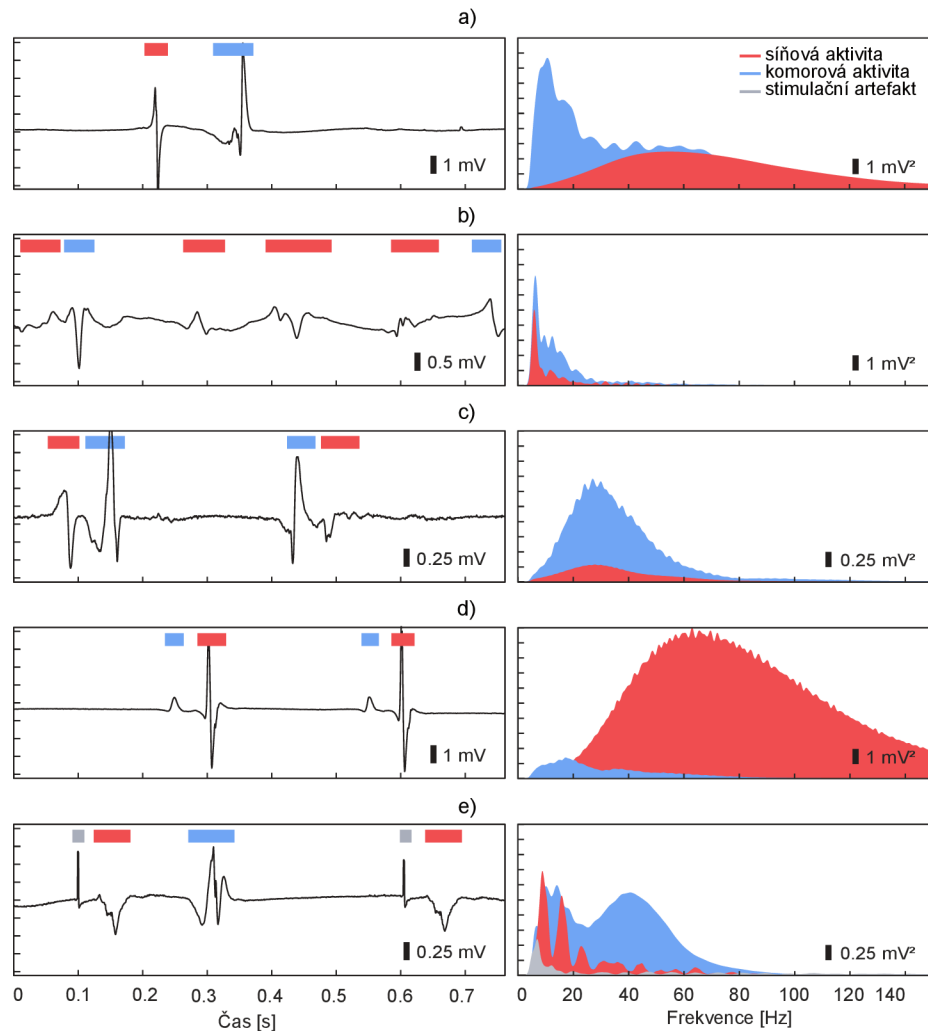
okna navázaného na referenční bod, kterým může být např. komorový komplex v povrchovém EKG, případně stimulační (S) artefakt [72]. Méně časté je využití adaptivního okna stanoveného z jiného – referenčního – kanálu [69]. Plně automatické metody pro identifikaci segmentů EGM popsané v literatuře zahrnují primárně základní filtrační a transformační techniky jako jsou korelace signálu s bankou vzorů [73], nelineární [74] operátor pro odhad okamžité energie signálu [75, 72], vlnková transformace [76], derivace a umocnění signálu [69] či Hilbertova transformace [77]. Anotace LAT v transformovaném signálu je následně provedena pomocí uživatelem předdefinovaným nebo adaptabilním [69] napětovým prahem.

Hlavním zdrojem selhání automatické segmentace při mapování supraventrikulárních tachykardií bývá obvykle přítomnost artefaktů z programované stimulace a elektrická aktivita komorového myokardu. Za normálních elektrofyziologických podmínek lze vymezení okna zájmu v kombinaci s konvenční lineární filtrací považovat za dostatečný nástroj pro eliminaci vlivu komorové aktivity. Stimulační artefakty je možné, jak ukázali autoři v [72], efektivně potlačit i v blízkosti síňového EGM pomocí přizpůsobeného filtru. Volba stimulačního protokolu či změna srdečního rytmu vedou ke změnám S-S, S-A a A-V intervalů. Důsledkem může být časové překrytí některé z dvojic S, A a V komplexů. Síňová a komorová složka signálu navíc v závislosti na vlastnostech tkáně a lokalizaci vykazuje vysokou míru morfologické variability. Ta může kolísat v obou případech od fragmentovaného průběhu s nízkou amplitudou až po ostře vymezené komplexy s velmi podobným frekvenčním obsahem. Tento fakt ilustrujeme na vybraných příkladech z DS-A na Obrázku 2.1. Výsledkem výše uvedených metod může v takovém případě být buď nadměrné odstranění síňové složky nebo neschopnost časově diferencovat sdružený A-V komplex se všemi negativními důsledky na detekci a segmentaci.

Přes pokrok v automatické segmentaci EGM je z výše uvedených příčin manuální akvizice i vizuální hodnocení LAT stále nezbytnou součástí EFS. S ohledem na počet snímaných kanálů a celkový akviziční čas se jedná o časově náročnou činnost zatíženou subjektivní chybou, a tím i sníženou reprodukovatelností. Automatická segmentace dlouhodobých záznamů z intrakardiálního snímání není jen výsadou elektroanatomického mapování. Uplatnění je možné nalézt v klinicky zaměřených studiích i v monitoraci pacientů s implantabilními stimulátory a defibrilátory [78].

Hlavní motivací této části disertační práce je návrh metody pro robustní sémantickou segmentaci síňových EGM ve vícerozměrných signálech snímaných multipolárním diagnostickým katétrem v koronárním sinu (CS). Jedná se o standardní a dobře reprodukovatelnou referenční lokalizaci v posteriorní oblasti srdce mezi levou síní a levou komorou. Zde zaznamenané EGM obvykle obsahují signály odpovídající síňové i komorové lokální aktivaci. Časově-prostorová distribuce síňových elektrogramů v CS je významným parametrem pro diferenciaci supraventrikulárních tachykardií [64, 65, 62], což je nutný předpoklad úspěšné ablační terapie.

Informace o časové aktivaci síní v kontextu arytmiických záznamů je v kombinaci se standardním povrchovým EKG záznamem cenným zdrojem pro semi- a plně supervizované úlohy hlubokého učení. Tato data by proto mohla nalézt uplatnění v úlohách zaměřených na klasifikaci srdečního rytmu s potenciálem zvýšit generalizační schopnosti hlubokých modelů a interpretovatelnost predikovaných výsledků.



Obr. 2.1: Ukázka morfologické variability lokálních EGM (vlevo) s odhadem výkonových spekter jednotlivých složek (vpravo) pomocí Welshova periodogramu (realizace z úseků o délce 20 s, frekvenční rozlišení 2 Hz). Separace S, A a V komplexu byla provedena váhováním oknem typu "Flat top"s manuálně stanovenými hranicemi. SR – sinusový rytmus; WPW – Wolff-Parkinson-White syndrom; KES – komorová extrasystola; AVRT – atrio-ventrikulární reentry tachykardie.

2.2 Dílčí cíle

Dílčí cíle vychází z požadavků kladených na kvalitu segmentace a vlastností trénovací datové sady. Požadavky na segmentaci síňových EGM jsou stanoveny následovně: a) nezávislost modelu na a-priorní znalosti procedurálních parametrů; b) přiřazení klasifikační třídy každému vzorku signálu s ohledem na možnou přítomnost polymorfních fibrilačních vln; c) potlačení vlivu překrytí stimulačními artefakty a elektrickou aktivitou komor; d) robustnost vůči běžným typům rušení.

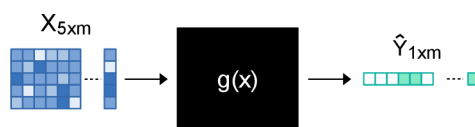
Vlastnosti datové sady potenciálně limitující proces návrhu a evaluace modelu definujeme jako: i) vstupní nerovnoměrné zastoupení morfologických typů síňových EGM; ii) významně vychýlené časové zastoupení síňových EGM oproti úsekům bez aktivity, které zasahují celkem 72.9 % všech vzorků signálu; iii) potenciálně malý rozsah datové sady. Cíle kapitoly jsou definovány následovně:

1. Návrh a implementace modelu pro sémantickou segmentaci síňových EGM. Primárním účelem modelu je nalezení globálního aktivačního času síňové aktivity ve více-kanálovém intrakardiálním záznamu z CS. V případě fibrilace síní je účelem nalezení hranice určující počátek a terminaci arytmie.
2. Optimalizace modelu z pohledu definovaných metrik kvality.
3. Kvantitativní hodnocení modelu proti expertně stanoveným referenčním bodům.
4. Implementace a kvantitativní hodnocení metod pro snížení vlivu nerovnoměrně zastoupených dat.
5. Implementace augmentačních technik s hodnocením jejich vlivu na segmentaci.
6. Inference modelu na dlouhodobých záznamech z EFS v rámci realizace DS-B.

2.3 Formální definice problému

Sémantická segmentace je proces zobrazení vstupních dat pomocí obecné transformační funkce $g(x)$ do takového prostoru proměnných, který umožní každému vzorku vstupního signálu x_i přiřadit hodnotu \hat{y}_i tak, aby představovala jednu z předem definovaných kategorií (Obrázek 3.1). Sémantickou segmentaci lze tedy považovat za problém klasifikace do jedné či více klasifikačních tříd. V našem případě je cílem stanovit, zda-li vzorek intrakardiálního signálu reprezentuje elektrickou aktivitu síní, což problém zjednodušuje na binární klasifikaci, a tedy $\hat{y}_i \in \{0, 1\}$.

Na základě formální definice je základním požadavkem na transformační funkce $g(x)$ zachování stejného časového rozlišení (vzorkovací frekvence) u vstupních a výstupních dat. Vstupní matice dat $5 \times m$ je v našem případě tvořena 5-kanálovým EGM signálem z CS katétru. Výstupní matice o rozměru $1 \times m$ je tvořena binární sekvencí, kde $\hat{y}_i = 1$ představuje vzorek odpovídající síňové aktivitě.



Obr. 2.2: Obecný princip sémantické segmentace EGM.

Následující text pojednává o hledání takové transformační funkce $g(x)$ a jejich parametrů, pro které segmentace EGM dosahuje na základě definovaných metrik nejvyšší kvality. Jedná se o explorativní studii zaměřenou na využití metod hlubokého učení. Studie je provedena na interním datasetu DS-A, který byl pro tyto účely vytvořen, viz sekce 1.6.3.

2.4 Architektura modelu

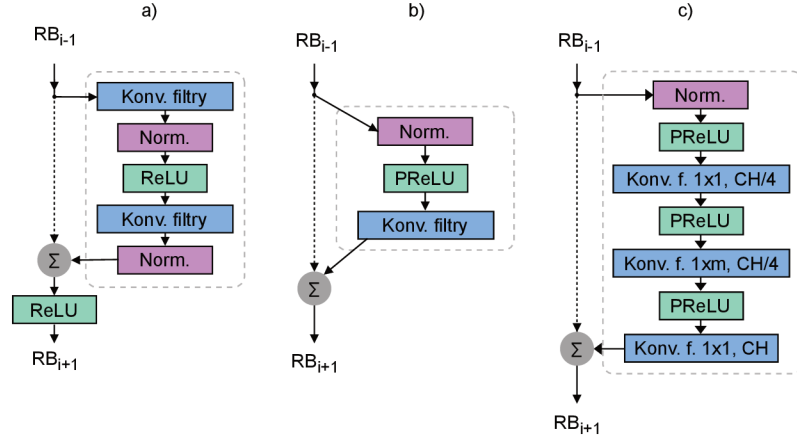
2.4.1 Enkodér s reziduálním spojením

Enkodérová část modelu zajišťuje extrakci příznaků na několika úrovních časové škály a je založena na vlastní implementaci plně konvoluční neuronové sítě (FCN²) typu ResNet, která je typická reziduálním spojením mezi jednotlivými konvolučními bloky [79, 80]. Tato paralelní větev rozšiřuje konvoluční blok $c(x)$ o identické zobrazení vstupních dat ve formě $c(x) + x$. Identické zobrazení umožňuje významným způsobem zvětšit hloubku sítě a zachovat přitom tok gradientů přes jednotlivé vrstvy v průběhu optimalizace [80].

V rámci této práce jsme na základě rešeršních poznatků a empirického testování aplikovali některé dílčí modifikace originálního reziduálního bloku RB. Vzhledem k počtu dimenzí prostoru vstupních dat byly původní dvourozměrné konvoluční masky nahrazeny jejich 1D ekvivalentem. Posun konvoluční masky je proto dán pouze ve směru časové osy a jednotlivé EGM kanály tvoří nezávislé příznaky vstupující do konvolučního filtru. Oproti

²Z anglického Fully Convolutional Network

původnímu řazení vrstev v RB [79] bylo využito schéma s tzv. předaktivací [80]. V tomto uspořádání jsou data nejdříve normalizována a následně prochází aktivační vrstvou. Až poté jsou zařazeny samotné konvoluční masky. Za identickým zobrazením je odstraněna výstupní aktivační vrstva, což umožňuje další snížení rizika mizejícího gradientu. Dle série experimentů provedené autory He et al. [79] přispívá předaktivovaný RB k lepší regularizaci modelu a v průměru nižší dosažené chyby při nezávislém testování.



Obr. 2.3: Bloková schémata: a) originální reziduální jednotky [79]; b) předaktivované jednotky [80]; c) předaktivované jednotky s úzkým hrdlem. Zkratky ReLU a PReLU znázorňují typy po částech lineárních aktivačních funkcí. Zkratka CH označuje počet příznaků vstupní aktivační mapy.

Na Obrázku 2.3 jsou uvedena bloková schémata originálního (a) a předaktivovaného RB (b). Třetí z architektonických bloků (c) reprezentuje předaktivovanou jednotku s úzkým hrdlem, která je využita v této práci. Úzké hrdlo je tvořeno kompresní konvoluční vrstvou s maskou velikosti 1×1 , která snižuje celkový počet vstupních příznaků CH na, v našem případě, $\frac{1}{2}$ původní hodnoty. Úzké hrdlo efektivně snižuje počet parametrů. Tím je snížena i výpočetní náročnost konvolučních filtrů pro receptivní snímání. Počet originálních příznaků je restaurován výstupními dekompresními filtry o rozměru 1×1 .

Jako aktivační funkce byla použita neohraničená po částech lineární funkce (PReLU³) s adaptabilním parametrem a_i , který moduluje sklon její záporné části [81]:

$$f(x) = \begin{cases} y_i & \text{pro } y_i > 0 \\ a_i y_i & \text{pro } y_i \leq 0 \end{cases} \quad (2.1)$$

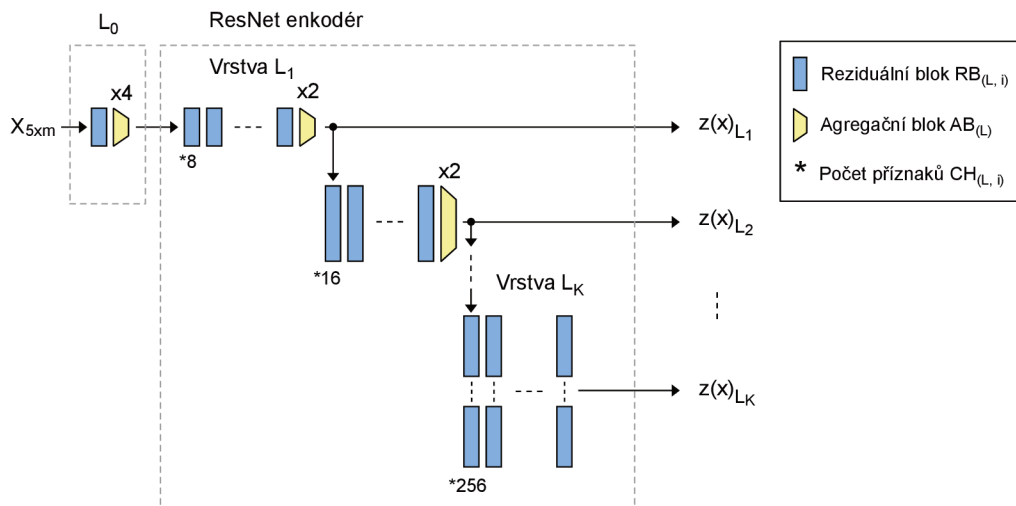
kde index i značí, že sklon je modulován napříč kanály a zůstává konstantní pro všechny prvky jednoho kanálu. Jedná se o zobecnění původní [82] rektifikační jednotky s $a_i = 0$, která se z logického důvodu potýkala s problémem mizejícího gradientu a v extrémních

³Z anglického Parametric Rectified Linear Unit

případech s úplným vyhasnutím zpětného šíření gradientů. Učenlivost parametru a_i navíc odstraňuje jeden z hyperparametrů s nutností manuálního nastavení a_i . Dle autorů Kaiming et al. [81] zlepšuje PReLU aktivace, oproti konstantnímu nastavení a_i , klasifikační schopnosti modelu za cenu zanedbatelného navýšení počtu jeho parametrů. Obdobný trend, přestože nebyl rozsáhle ověřen, byl pozorován i na EGM záznamech.

Na Obrázku 2.4 uvádíme zjednodušené blokové schéma ResNet enkodéru. Vstupní vrstva L_0 je tvořena jedním RB doplněným o průměrující agregační operátor s cílem redukovat vzorkovací frekvenci vstupního signálu na 500 Hz. Enkodér obsahuje volitelný počet vrstev L_k . Jednotlivé vrstvy mohou obsahovat jeden a více reziduálních bloků $RB_{L,i}$. Za účelem rozšíření receptivního pole a snížení výpočetní komplexity je každá vrstva typicky zakončena agregačním operátorem, který snižuje vzorkovací frekvenci na $\frac{1}{2}$ původní hodnoty. V našem případě je agregace provedena konvenčním průměrováním s oknem velikosti 1×2 s doplňkovou 1×1 konvoluční maskou.

Počet generovaných příznaků roste s každou následující vrstvou. Jeho růst byl dán konstantním faktorem 2^k pro $k \in \{x \in \mathbb{Z} \mid x \geq 3\}$. Výstup enkodéru je tvořen n -tíci příznakových map $z(X)_{L,k}$ ze všech úrovní (vrstev). Příznakové mapy jsou následně využity v segmentační (dekodérové) části modelu.

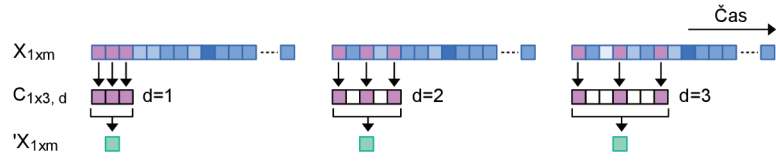


Obr. 2.4: Blokové schéma 1D konvolučního enkodéru založeného na architektuře ResNet.

2.4.2 Blok dilatovaných konvolucí

Signál síňové aktivace je za běžných okolností relativně úzce časově vymezen. Výjimkou je signál snímaný v průběhu fibrilace síní, který vykazuje vysokou míru neuspořádanosti a ve vícesvodových záznamech se projevuje jako polymorfní signál s variabilní frekvencí a dlouhou dobou trvání. Pro segmentaci celého úseku s výskytem fibrilačních vln [83] vyžaduje klasifikátor dostatečně široké receptivní pole pro zachycení dlouhodobých variací signálu. Ty se pohybují o několik řádů výše než je průměrná délka trvání normální síňové

depolarizace. Proto se nejví jako efektivní běžné řešení v podobě rozšíření konvolučních jader nebo prohloubení enkodérové části modelu (zvýšení stupně podvzorkování signálu), kdy oba způsoby značně navyšují počet parametrů modelu. Rozšíření receptivní oblasti je proto realizováno zařazením reziduálních bloků s dilatovanými konvolučními maskami RB_D za poslední vrstvu enkodéru L_K . Ty umožňují exponenciálně navýšit receptivní oblast bez ztráty rozlišení či zvýšení počtu parametrů modelu. Současně poskytují informace o okolním kontextu [84] a umožňují zachytit dlouhodobé variace např. u řečových signálů [85]. Základní princip rozšíření receptivního pole pomocí dilatované konvoluční masky je uveden na Obrázku 2.5.

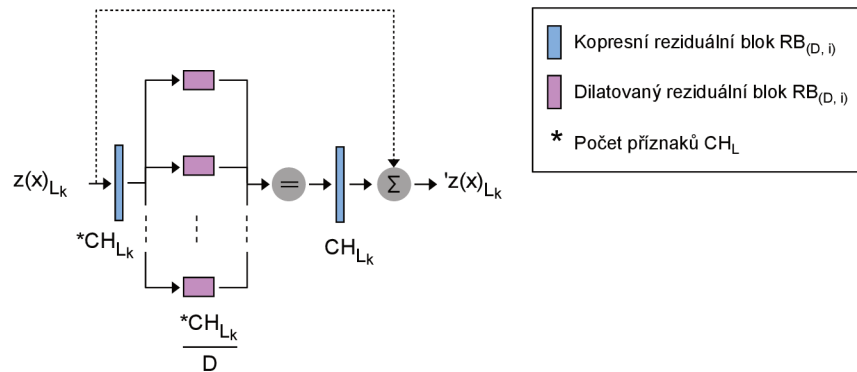


Obr. 2.5: Schématické znázornění dilatovaného konvolučního operátoru.

Dilatační faktor d určuje diskrétní vzdálenost mezi po sobě jdoucími prvky konvolučního jádra. Vynechané vzorky signálu se tedy následně nepodílejí na konvoluční operaci. Přibližnou velikost receptivní oblasti filtru v ms lze definovat jako:

$$t_{rec} = \frac{1000}{f_s} \cdot (1 + d_i \cdot (f_{RB} - 1)), \quad (2.2)$$

kde d_i je dilatační faktor ve vrstvě $RB_{D,i}$, f_s je vzorkovací frekvence vstupní příznakové mapy a f_{RB} je počet aktivních prvků konvoluční masky. Schéma paralelního řazení v bloku je uvedeno na Obrázku 2.6.

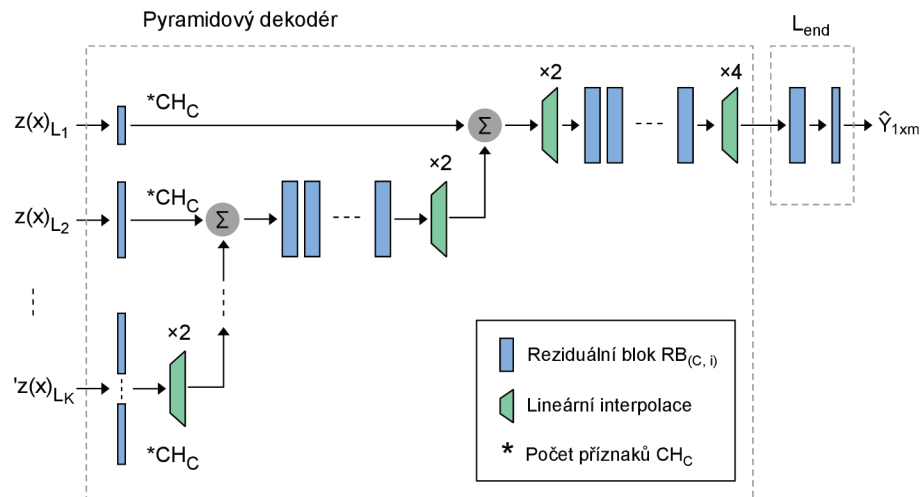


Obr. 2.6: Schéma bloku s dilatovanými konvolučními jádry. D – počet paralelních větví a kompresní faktor vstupních konvolučních masek; CH_{L_k} – počet příznaků vstupní mapy.

Aby bylo možné extrahovat příznaky v různém časovém měřítku, je v dekodérové části použito více paralelních větví. Z nich každá obsahuje konvoluční jádra 1×3 s variabilní mírou dilatace d . Vstupní příznaková mapa je nejdříve komprimována s faktorem D , kde D odpovídá počtu paralelních větví s dilatovanými RB_D . Po průchodu vstupních dat dilatovanými maskami jsou příznakové mapy z každé větve obdobně jako v [84] sdruženy do jednoho výsledného tensoru. Z pohledu výpočetní náročnosti obsahují všechny bloky RB_D stejný počet parametrů, avšak díky rozdílným dilatačním faktorům široký časový rozsah signálu. Růst parametru d je dán konstantním faktorem 2^k pro $k \in \{x \in \mathbb{Z} \mid x \geq 1\}$.

2.4.3 Pyramidový dekodér

Účelem dekodérové části je agregace extrahovaných příznakových map do časového měřítka, které odpovídá vstupním EGM signálům. V rámci této práce byla použita hybridní struktura modelu vycházející z architektury U-net [86, 87] a architektury pyramidových příznaků [88, 89]. Schéma architektury je uvedeno na Obrázku 2.7. Do enkodéru vstupují příznakové mapy $z(X)_{L,k}$ ze všech vrstev enkodéru. Vzhledem k tomu, že se počty příznaků v každé mapě liší, je v prvním kroku provedena komprese pomocí masek 1×1 do prostoru o stejném počtu dimenzí. Převzorkování časové škály neprobíhá pomocí dekonvolučních filtrů [90] jako v případě U-Net [86], ale s využitím prosté lineární interpolace. Ta nevyžaduje dodatečné učení a umožňuje snížit výpočetní náročnost modelu. Z agregovaných příznakových map je na výstupu získán signál o rozměru $1 \times m$ se stejnou vzorkovací frekvencí jakou měla vstupní EGM data. Na rozdíl od autorů Lin et al. [88] nevyužíváme v této části práce nezávislé učení jednotlivých vrstev dekodéru, a to s ohledem na řídký výskyt referenčních segmentů síňové aktivity v nejnižších úrovních hierarchie modelu.



Obr. 2.7: Blokové schéma 1D dekodéru s pyramidovým zpracováním příznaků. CH_C – parametr určující počet příznaků po kompresi vstupních příznakových map.

2.5 Kriteriační funkce

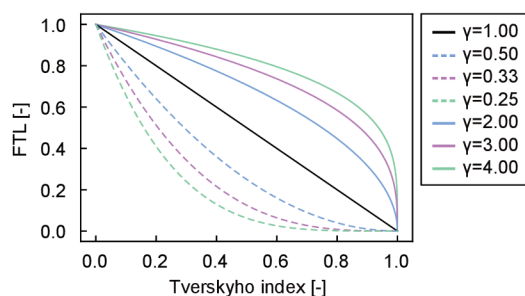
Kriteriační funkce je exaktně definovaný cíl matematické optimalizace parametrů neuronové sítě. Zatímco pro klasifikační úlohy je typickou volbou křížová entropie, v úlohách sémantické segmentace medicínských dat dominuje využití spojitě aproximace Sørensen–Dice koeficientu (SDC) [91]. Sørensen–Dice koeficient je založen na relativním vyjádření míry překryvu predikovaného a referenčního segmentu. Z principu tato metrika nabývá totožné hodnoty pro stejnou míru překryvu v případě rozměrných i malých objektů. Z tohoto vyplývá i její zřejmá schopnost snížit vliv nerovnoměrného zastoupení pozitivních a negativních vzorků signálu.

Autoři v [92] dále rozšířili kriteriační funkci založenou na SDC o exponent, který zvyšuje příspěvek obtížně rozlišitelných objektů k celkové penalizaci modelu. Zobecněním je pak kriteriační funkce aproximující Tverskyho index (FTL) [93], která umožňuje pomocí dvou hyperparametrů nastavit kompromis mezi senzitivitou (Se) a pozitivní prediktivní hodnotou (P^+) a je využita v této části práce:

$$FTL = \sum_n \left\{ 1 - \frac{\hat{Y} \odot T}{\hat{Y} \odot T + \alpha[(1 - \hat{Y}) \odot T] + \beta[\hat{Y} \odot (1 - T)]} \right\}^\gamma, \quad (2.3)$$

kde \odot značí Hadamardův součin tenzorů; \hat{Y} a T jsou tenzory představující surová skóre vrácená modelem, respektive binární referenční hodnoty; α a β jsou váhy upravující vztah mezi Se a P^+ ; exponent γ moduluje příspěvek snadno a obtížně trénovatelných vzorů k celkové hodnotě kriteriační funkce. Pro $\alpha = \beta = 0.5$ and $\gamma = 1$ se kriteriační funkce zjednodušuje na 1-SDC. To umožňuje nalezení optimálního nastavení v široké skupině kriteriačních funkcí běžně užívaných v oblasti sémantické segmentace.

Na Obrázku 2.8 je znázorněn vztah mezi původní hodnotou Tverskyho indexu a výslednou chybou, kterou generuje FTL pro různá nastavení exponenciální váhy γ . V optimálním případě nabývá Tverskyho index hodnoty 1 a FTL hodnoty 0.



Obr. 2.8: Průběh závislosti FTL na Tverskyho indexu pro vybrané varianty váhy γ .

2.6 Optimalizační algoritmus

2.6.1 Stochastická gradientní optimalizace

Optimalizace modelu byla provedena pomocí algoritmu AdamW [94] s L_2 regularizací separovanou od kroku učení [95]. Algoritmus patří do rodiny stochastických metod gradientního sestupu s adaptivním krokem učení a průměrováním lokálních gradientů napříč jednotlivými iteracemi učení s cílem zamezit uvíznutí v lokálním minimu. Základním principem je úprava vah modelu nikoliv na základě aktuální hodnoty gradientu, ale s využitím exponenciálního klouzavého průměru gradientů a jejich umocněných variant přes více po sobě jdoucích iterací. Metoda zavádí do optimalizace dva nové hyperparametry $\beta_1, \beta_2 \in [0, 1)$, které regulují strmost poklesu exponenciálních vah obou zmíněných klouzavých průměrů. Vztah mezi hyperparametry a počtem iterací i , přes které je proveden odhad klouzavého průměru lze zjednodušeně formulovat jako [94]:

$$i = \frac{1}{1 - \beta_x}. \quad (2.4)$$

Metoda by měla ve srovnání se svými protějšky poskytovat rychlejší konvergenci při zachování výchozích hodnot hyperparametrů β_1, β_2 a ϵ , což je výhodné zejména v rozsáhlých experimentálních studiích. Některé případové studie zabývající se segmentací obrazových dat [96, 97] kritizují horší generalizační schopnosti na neznámých datech ve srovnání s konvenčními neadaptivními metodami stochastického gradientního sestupu (SGD). Rozsáhlé empirické studie publikované v [98, 99] však poukazují na to, že při adekvátní optimalizaci hyperparametrů nejsou adaptivní metody významně horší než SGD. A to i v případě, že se optimalizace týká jen úvodního kroku učení a pravidel pro jeho plánování v navazujících iteracích (tzv. *Scheduling*). Vzhledem k tomu, že ani při hrubém prohledávání stavového prostoru nebyly nalezeny hodnoty, které by významně zlepšily kvalitu segmentace, byly parametry β_1, β_2 a ϵ ponechány na hodnotách doporučených autory v [94].

Optimální výchozí hodnota kroku učení η_0 byla získána pomocí Bayesovské optimalizace (blíže viz dále). Hodnoty hyperparametrů optimalizačního algoritmu jsou uvedeny v Tabulce 2.1.

Tab. 2.1: Nastavení hyperparametrů optimalizačního algoritmu.

Hyperparametr	η_0	β_1	β_2	ϵ
Hodnota	0.005	0.9	0.999	0

2.6.2 Plánování kroku učení

Pro průběžné snižování kroku učení byla implementována metoda multiplikativního poklesu s faktorem w_d řízená mírou ustálení Sørensen–Dice koeficientu na validační sadě DS- A_V (tzv. metoda *Reduce On Plateau*). Adekvátní plánování velikosti kroku učení obecně snižuje pravděpodobnost uvíznutí v lokálním minimu v časných fázích učení. V závěru učení naopak umožňuje dosáhnout minima optimalizované funkce. Dosažení plateau fáze s následným snížením kroku učení bylo provedeno za předpokladu, že v N po sobě jdoucích epochách nedošlo ke zlepšení SDC alespoň na hodnotu:

$$\text{median}(\text{SDC}_{N_i}, \text{SDC}_{N_{i-1}}, \dots, \text{SDC}_{N_{i-n}}) + \delta_{V_1}. \quad (2.5)$$

Prahová hodnota δ_{V_1} je uvedena v Tabulce 2.2 a reprezentuje očekávané procentuální zlepšení SDC mezi dvěma epochami. Jedná se o jedno z konvenčních pravidel pro řízení kroku učení, které bylo rozšířeno o výpočet mediánové hodnoty z posloupnosti hodnot metriky kvality. Tím byla navýšena robustnost vůči náhlým výkyvům metriky, které jsou u kriteriální funkce SDC běžné a mohou způsobit předčasné ukončení procesu optimalizace. Optimalizace modelu byla ukončena po dosažení alespoň jednoho z následujících endpointů:

1. Konvergence SDC skóre na validační sadě. Za konvergenci byla považována situace, kdy po třetím snížení kroku učení a dosažení fáze plateau nedošlo k očekávanému nárůstu metriky kvality alespoň o prahovou hodnotu δ_{V_3} .
2. Divergence SDC mezi trénovací a validační sadou. Definována jako rozdíl v N po sobě jdoucích epochách větší než prahová hodnota δ_{V_2} .

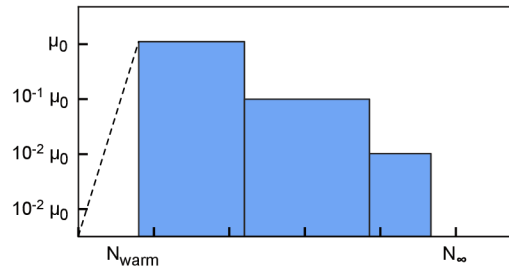
V počáteční fázi trénování byla aplikována tzv. zahřívací fáze, definována počtem epoch N_{warm} . Během té dochází k lineárnímu nárůstu kroku učení v každé iteraci i o předem stanovený inkrement až do cílové hodnoty η_0 , viz blíže v Příloze v sekci A.1.3. Cílem této fáze je hypotetické snížení vlivu náhodné selekce trénovacích vzorů v počátku učení na směr a strmost posunu v prostoru optimalizované funkce [100]. Parametry řízení kroku učení a procesu optimalizace jsou uvedeny v Tabulce 2.2.

Tab. 2.2: Parametry řízení kroku učení a ukončení optimalizace.

Hyperparametr	N_{warm}	w_d	δ_{V_1}	δ_{V_2}	δ_{V_3}
Hodnota	8	0.1	0.01	0.1	0.005

Hodnoty v Tabulce 2.2 byly získány pomocí heuristické analýzy a zůstaly konstantní pro všechny provedené experimenty. Hodnota N_{warm} byla zvolena experimentálně a to přesto, že existují heuristicky odvozená pravidla [101] zohledňující parametry optimalizačního

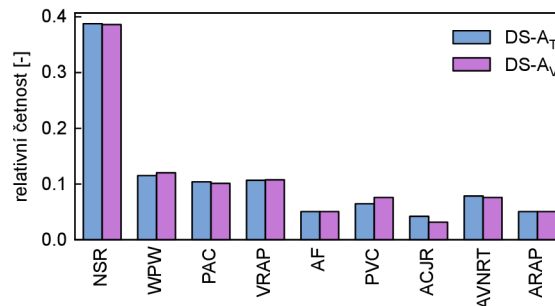
algoritmu. I konzervativně stanovený odhad však v podmínkách naší trénovací sady několikanásobně překračoval počet epoch potřebných k úplnému naučení modelu. Princip metody je znázorněn na Obrázku 2.9.



Obr. 2.9: Schématické znázornění implementovaného algoritmu pro řízení kroku učení.

2.7 Stratifikace DS-A

Dataset byl pomocí metody hold-out rozdělen na trénovací DS- A_T a validační DS- A_V množinu v poměru 70:30. Data byla rozdělena pomocí stratifikované randomizace tak, aby každá sada obsahovala pouze unikátní pacienty a nedošlo k zanesení závislosti mezi trénovací a validační sadou. Cílem stratifikace bylo dosažení maximální homogenity obou sad z pohledu zastoupení vybraných diagnostických tříd. Výběr se řídil dvěma základními kritérii: a) podobnost časového uspořádání síňové aktivity; b) potenciálně významný vliv na kvalitu segmentace (obtížně segmentovatelné úseky). Aby byla zajištěna reprodukovatelnost experimentů, byla provedena randomizace pacientů pomocí pseudo-náhodného generátoru s předdefinovaným semínkem. Výběr semínka byl dán optimální stratifikací pacientů. Kvalita stratifikace byla stanovena na základě hodnoty relativní entropie a statistické shody diskrétního rozdělení. Shoda byla testována pomocí Pearsonova χ^2 testu. Zastoupení jednotlivých diagnostických tříd v DS- A_T a validační DS- A_V po stratifikaci je uvedeno na Obrázku 2.10. Obecný princip metody je popsán Algoritmem 2.



Obr. 2.10: Relativní rozložení diagnostických tříd v DS- A_T a DS- A_V po stratifikaci.

Algorithm 2 Náhodné prohledávání množiny semínek

Vstup: Množina náhodně generovaných semínek $S = \{s_1, s_2, \dots, s_i\}$.

Výstup: Semínko s_{best}

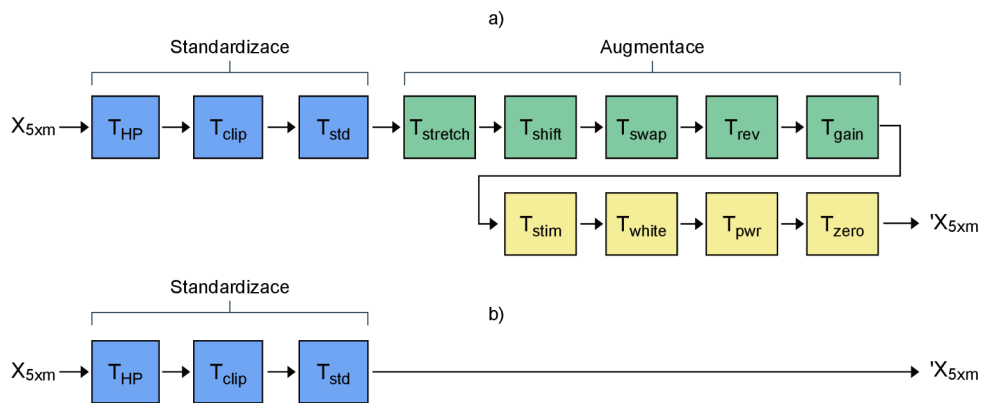
```
1: for každé semínko  $s_i$  do
2:   Promíchej pořadí v DS-A pomocí pseudo-náhodného generátoru se semínkem  $s_i$ .
3:   Zařaď prvních 70 % záznamů do DS- $A_T$ .
4:   Zařaď zbylých 30 % záznamů do DS- $A_V$ .
5:   Urči pravděpodobnostního rozložení  $P(x_T)$  a  $Q(x_V)$  diag. tříd v DS- $A_T$  a DS- $A_V$ .
6:   Urči  $\chi^2, p$  pro  $P(x_T)$  a  $Q(x_V)$ .
7:   if  $p > 0.05$  then
8:     Urči relativní entropii  $d(P||Q)_i$ .
9:     if  $d(P||Q)_i < d(P||Q)_{best}$  then
10:        $s_{best} \leftarrow s_i$ .
11:        $d(P||Q)_{best} \leftarrow d(P||Q)_i$ .
12:     end if
13:   end if
14: end for
```

2.8 Předzpracování dat

Předzpracování dat zahrnuje dvě skupiny transformačních funkcí. První skupina transformací slouží k aplikaci základních filtrů a normalizaci signálů do škály zpracovatelné modelem. Tyto transformace byly aplikovány v průběhu učení i při vybavení modelu po jeho nasazení. Do řetězce vstupuje matice surových bipolárních EGM signálů s originální vzorkovací frekvencí 2000 Hz. Snížení objemu dat podvzorkováním je plně ponecháno na vstupní vrstvě hlubokého modelu. V prvním kroku je aplikován FIR filtr typu horní propust (T_{HP}) s mezní frekvencí 5 Hz. Hodnota mezní frekvence byla volena striktněji na základě standardů pro unipolární záznamy [55], aby došlo k potlačení nízkofrekvenčního rušení, ale bylo zachováno maximum užitečné informace. Signál je následně ořezán (T_{clip}) na hladině napětí 12.46 mV. Posledním krokem je standardizace (T_{std}) škály napětí pomocí do jednotkového rozpětí v rámci 10-90 percentilu, což odpovídá ± 0.79 mV.

Druhá skupina transformací byla aplikována jen v průběhu trénování. Jedná se o řetězec morfologických transformací/augmentací trénovací sady DS- A_T . Jejich cílem je modulace vstupních dat se zachováním základního charakteru síňových signálů. Jde o jednu ze základních strategií pro regularizaci modelu zanesením větší intra-individuální variability a vystavením modelu náročnějším podmínkám učení [102, 103]. Morfologická augmentace zahrnuje široké spektrum metod, jejichž studium je mimo hlavní záměr disertační práce. Z toho důvodu byly implementovány dva typy obecně užívaných technik [104] pro modulaci morfologie a zanesení aditivního šumu. V bloku $T_{stretch}$ dochází k roztažení signálu v časové ose o $\pm 0 - 20\%$, a tím změně napěťové dynamiky signálu. Transformace je zalo-

žena na lineární interpolaci. Následně jsou aplikovány náhodné transformace jako je časový posun o i vzorků (T_{shift}), záměna dvou po sobě jdoucích kanálů (T_{swap}), změna polarity signálu (T_{rev}) simulující změnu směru šíření aktivační vlny či chybné zapojení elektrod a změna zesílení (T_{gain}) o $\pm 0 - 20\%$ aplikovaná nezávisle na každý kanál. Dále následují transformace založené na simulaci běžných typů rušení v elektrokardiografii. Vstupní EGM signál je rozšířen o aditivní šum Gaussovského typu (T_{white}) a síťové rušení (T_{pwr}) s náhodně generovanou úrovní energie tak, aby bylo dosaženo poměru signál-šum od 20 do 0 dB. V bloku T_{stim} byl implementován generátor stimulačních artefaktů s náhodně generovanou amplitudou, šířkou a délkou cyklu v rozsahu 0.5–5.0 mV, 2–10 ms, respektive 200–2000 ms. Blok T_{zero} pak simuluje přítomnost odpojených elektrod pomocí nulování kanálu s přednastavenou pravděpodobností aktivace 0.1. Celý řetězec aplikovaných metod je ukázán na Obrázku 2.11. Nastavení parametrů jednotlivých transformačních funkcí bylo provedeno na základě heuristické analýzy a zůstalo pro všechny navazující experimenty konstantní. Souhrn parametrů uvádíme v Příloze A.1.1. Vliv jednotlivých bloků na metriky kvality byl ověřen pomocí ablační studie. Vliv nastavení parametrů nebyl pro široké možnosti nastavení hlouběji zkoumán.



Obr. 2.11: Blokové schéma řetězce pro předzpracování signálů: a) řetězec předzpracování při učení; b) řetězec předzpracování při vybavování modelu.

2.9 Baysovská optimalizace hyperparametrů

Baysovská optimalizace [105] patří mezi sekvenční optimalizační techniky (SMBO⁴) využívající Baysovskou logiku pro informovaný odhad nejlepší kombinace hyperparametrů účelové funkce vycházející ze série po sobě jdoucích experimentů. Účelová funkce f má zpravidla neznámé analytické vyjádření [106] a nemusí být ani konvexní ani spojitá, což prakticky vylučuje použití gradientních metod. Účelová funkce je proto v případě SMBO aproximována pomocí pravděpodobnostního modelu a úvodního odhadu apriorního rozdělení. V každé iteraci je na základě apriorního rozdělení proveden výběr kandidátních

⁴Z anglického Sequential Model-Based Optimization

hodnot hyperparametrů a výpočet funkční hodnoty objektivní funkce. Na jejich základě je poté stanoveno posteriorní rozdělení, které poskytuje zpřesněný odhad pravděpodobnostního modelu a novou apriorní znalost pro následující pozorování. Mezi nejčastější metody pro konstrukci pravděpodobnostního modelu lze zařadit zejména Gaussovské procesy (GP) [107] a metody založené na odhadu Parzenových oken (TPE⁵) [106]. Na rozdíl od GP, které modelují pravděpodobnostní vztah $p(y|x)$, vychází metoda TPE z odhadu podmíněné pravděpodobnosti hyperparametrů (x) vzhledem k hodnotě účelové funkce (y), respektive hodnotě metriky kvality hlubokého modelu, tedy $p(x|y)$. TPE místo apriorního rozdělení definuje dva hierarchické procesy:

$$p(x|y) = \begin{cases} l(x) & \text{pro } y < y^* \\ g(x) & \text{pro } y \geq y^* \end{cases} \quad (2.6)$$

kde $l(x)$ a $g(x)$ jsou generativní modely prostoru hyperparametrů, které pravděpodobně přispějí ke zlepšení či zhoršení metriky kvality ve srovnání s prahovou hodnotou y .

V kontextu této práce je cílem nalézt takovou kombinaci hyperparametrů modelu x^* , pro kterou metrika kvality segmentace založená na F_1 skóre na validační sadě dat dosahuje maximální hodnoty:

$$x^* = \operatorname{argmax}_{x \in \mathcal{X}} f(x), \quad (2.7)$$

kde, \mathcal{X} je kompaktní množina hyperparametrů definována v následujícím textu. Pro optimalizaci byl využit pravděpodobnostní model TPE (Optuna, [108]). Ten umožňuje pozorování kategoriálních hyperparametrů, a dle autorů [106] dosahuje lepších výsledků než v případě konvenčního prohledávání hrubou silou i Bayesovovské optimalizace s GP modely

2.9.1 Optimalizace M_{aopt} a M_{lopt}

Optimalizace byla provedena nezávisle pro dvě skupiny hyperparametrů. První sada experimentů byla zaměřena na optimalizaci vybraných hyperparametrů architektury a procesu učení. Parametry kritériální funkce byly nastaveny konstantně na hodnotu $\alpha_{\text{FTL}} = \beta_{\text{FTL}} = 0.5$ a $\gamma_{\text{FTL}} = 1.0$, což odpovídá aproximaci Sørensen–Dice koeficientu. Výsledkem je segmentační model označený jako M_{aopt} . Druhá sada experimentů měla za úkol najít nejhodnější parametry kritériální funkce z rovnice 2.3 pro optimalizovaný model M_{aopt} . Výsledkem druhého kola optimalizace parametrů je model M_{lopt} . Kritérium optimality bylo stanoveno jako dosažení nejvyššího SDC na validační sadě DS- A_V . Jedná se tedy o jednocílovou optimalizaci. Na výpočetní efektivitu modelu byl kladen důraz primárně při návrhu architektonických bloků a tento parametr není při optimalizaci zohledněn. Abychom snížili míru vlivu stochastické inicializace, byla v každé iteraci provedena 3 nezávislá opakování se stejnou sadou hyperparametrů. Jako vstup do algoritmu TPE bylo

⁵Z anglického Tree Parzen Estimator

vybráno skóre s nejvyšší hodnotou. Směrodatná odchylka sady měření byla pro nízký počet opakování zanedbána.

Prostor parametrů pro optimalizaci modelu M_{aopt} je uveden v Tabulce 2.3. Horní hranice definičního oboru byla u většiny případů zvolena tak, aby nedošlo k překročení dostupných paměťových prostředků grafického akcelérátoru. Přestože implementace modelu umožňuje variabilní nastavení na úrovni jednotlivých vrstev i reziduálních bloků, byly parametry voleny tak, aby bylo dosaženo interpretovatelných výsledků v rozumném výpočetním čase. Šířka konvolučních masek f_{RB} je proto shodná pro všechny reziduální jednotky v modelu, vyjma 1×1 filtrů a vstupní vrstvy L_0 , která má za úkol zajistit čtyřnásobné podvzorkování signálu bez významné ztráty informace. Počty konvolučních filtrů, které definují dimenzionalitu prostoru příznakové mapy v každé vrstvě, jsou nastaveny konstantně. Výchozí hodnota i strmost růstu v každé vrstvě byla definována v sekci 2.4.1. Maximální šířka modelu je proto modulována současně s jeho hloubkou, a to na základě parametru K_e .

Tab. 2.3: Specifikace prostoru hyperparametrů modelu pro parametry architektury

Název hyperparametru	Symbol	Definiční obor
Výchozí krok učení	η_0	$\{1^{-5}, 5^{-5}, 1^{-4}, \dots, 1^{-1}\}$
Dávka trénovacích vzorů	BS_T	$\{8, 16, 24, 32\}$
Šířka konvolučních masek	f_{RB}	$\{3, 5, \dots, 11\}$
Šířka konvoluční masky v L_0	f_{L_0}	$\{3, 5, \dots, 25\}$
Počet vrstev v enkodéru	K_e	$\{x \in \mathbb{N} \mid 2 \leq x \leq 6\}$
Počet RB v jedné vrstvě	N_{RB}	$\{x \in \mathbb{N} \mid 2 \leq x \leq 8\}$
Počet kompresních filtrů dekodéru	CH_C	$\{32, 40, \dots, 128\}$
Počet vrstev dilatovaných konvolucí	D_d	$\{x \in \mathbb{N} \mid 2 \leq x \leq 5\}$

Obdobně je modulována receptivní oblast bloku s dilatovanými konvolučními filtry, viz sekce 2.4.2. Maximální hodnota parametr D_d určuje společně s hloubkou modelu maximální receptivní oblast dilatovaných konvolucí. Pro zdůvodnění volby uvádíme ilustrační příklad. Pro $K_e = 6$ a $D_d = 5$ je vzorkovací frekvence příznakové mapy v časové ose přibližně 15 Hz a dilatační faktor masky 32. Pro šířku masky 1×3 odpovídá zachycený časový rozsah dle rovnice 2.2 4333 ms. Taková oblast je dostatečná pro zachycení rozdílů mezi jednotlivými síňovými stahy i střednědobých změn, které mohou pomoci rozlišit relativně pravidelné arytmie od prakticky náhodných EGM u fibrilace síní.

Cílem druhé sady experimentů bylo nalezení optimálních parametrů penalizační funkce z Rovnice 2.3 pro architekturu M_{aopt} . Výstupem je optimalizovaný model M_{lopt} . Separovaná studie parametrů kritériální funkce byla zvolena kvůli snazší interpretaci jejich vlivu na proces učení ve specifických podmínkách DS-A, a případné možnosti manuálního ladění s ohledem na další, klinicky významné metriky. Proměnné α_{FTL} a β_{FTL} jsou parametry váhující vliv falešně negativních a falešně pozitivních vzorků. Exponenciální váha γ_{FTL} moduluje velikosti Tverskyho indexu na gradient chyby, který bude v průběhu učení propagován modelem. Definiční obor hyperparametrů kritériální funkce je uveden v Tabulce 2.4.

Tab. 2.4: Specifikace prostoru hyperparametrů pro kritériální funkci

Název hyperparametru	Symbol	Definiční obor
Váhování falešně negativních vzorků	α_{FTL}	$\{x \in \mathbb{R} \mid 0.1 \leq x \leq 0.9\}$
Váhování falešně pozitivních vzorků	β_{FTL}	$x = 1 - \alpha_{FTL}$
Exponenciální váha celkové chyby	γ_{FTL}	$\{\frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4\}$

2.10 Hodnocení kvality modelu

Problematikou hodnocení modelů pro segmentaci elektrokardiografických signálů se detailně zabývá explorativní studie od Butkuvienė et al. [109]. Z jejich výsledků vychází jako nejvíce relevantní metrika založená na Matthewsově korelačním koeficientu (MCC), která přikládá rovnoměrnou váhu všem hodnotám v matici záměn. Naopak některé standardní metriky jako senzitivita, specifická a plocha pod ROC křivkou jsou, zejména v případě segmentace řídké se vyskytujícími objekty, značně zkresleny. MCC byl sledován v průběhu všech experimentů. I přesto byl v rámci hodnocení kvality modelu reportován široce užívaný Sørensen–Dice koeficientu (SDC), a to na základě těchto důvodů: a) použitá kritériální funkce je generalizací SDC a většina optimalizačních kroků je založena na hodnocení SDC; b) snazší interpretace výsledků s ohledem na předchozí bod; c) v průběhu všech experimentů vykazoval trend MCC a SDC extrémní míru shody.

Pro významové odlišení spojitě aproximace SDC využitě při optimalizaci modelu a samotné metriky kvality bude v rámci sekce výsledků uváděn alternativní název F_1 skóre. F_1 skóre lze pro matici záměn definovat jako:

$$F_1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}, \quad (2.8)$$

kde TP označuje správně identifikovaný pozitivní vzorek, FP falešně pozitivní vzorek a FN falešně negativní vzorek signálu.

Z pohledu binární klasifikace lze F_1 skóre interpretovat jako harmonický průměr senzitivity (Se) a pozitivní prediktivní hodnoty (P^+). Díky tomu může jednu hodnotu F_1

reprezentovat široká škála kombinací Se a P⁺. Z tohoto důvodu byly ve vybraných případech reportovány také tyto dvě metriky. Se a P⁺ lze na základě správnosti zařazení vzorků definovat následovně:

$$\text{Se} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (2.9)$$

$$\text{P}^+ = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (2.10)$$

Pro hodnocení lokálních aktivačních časů byl každý spojitý segment stanovený modelem převeden na individuální instanci. Instanci tvoří n-tice pozic reprezentující počátek a konec síňové aktivity v EGM z CS. Přesnost stanovení pozic byla hodnocena na základě absolutní (AE) a střední kvadratické chyby (RMSE) definované následovně:

$$\text{AE} = |\hat{y}_i - y_i|, \quad (2.11)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}, \quad (2.12)$$

kde \hat{y}_i je odhad pozice počátku či konce provedený modelem a y_i je referenční anotace. Dle standardu ANSI/AAMI-EC57:1998 byly brány v potaz pouze ty hodnoty, které se vyskytli v tolerančním intervalu τ_{tol} v okolí referenční pozice. Hodnota tolerance byla stanovena na $\tau_{\text{tol}} = 35$ ms, což odpovídá dvojnásobku směrodatné odchylky doby trvání síňových EGM v DS-A.

Jedním z nedostatků sémantické segmentace může být vznik dvou a více nespojitých oblastí pro reprezentaci jednoho objektu. Model i v takovém případě obvykle dosahuje vysokých hodnot metrik kvality na úrovni jednotlivých vzorků signálu. Použití konvenčních metrik proto může být zavádějící, a to především v situaci, kdy požadujeme jak klasifikaci izolovaných síňových stahů, tak dlouhých úseků se záznamem fibrilace síní. Z tohoto důvodu byla zavedena doplňková metrika založená na celkové procentuální odchylce (PE) mezi počtem referenčních anotací a modelem identifikovaných instancí:

$$\text{PE} = 100 \cdot \frac{(\text{P}_t - \text{A}_t)}{\text{A}_t}, \quad (2.13)$$

kde PE je průměrná procentuální odchylka v %, P_t je počet pozorovaných instancí síňové aktivity a A_t je aktuální počet referencí v signálu stanovených expertem. Optimální segmentace nastává pro hodnotu PE blížící se nule. Jelikož rovnice nezahrnuje absolutní hodnotu reziduí, je možné sledovat také směr případného systematického zkreslení. Pro hodnotu $\text{PE} > 0$ je možné metriku interpretovat jako míru diskontinuity, pro $\text{PE} < 0$ potom jako míru slučování detekovaných oblastí.

2.10.1 Statistické metody

Výsledky vztahující se k hodnocení kvality modelu jsou uvedeny jako nejvyšší dosažené skóre v případě, že počet experimentů pro jednu množinu parametrů byl menší než deset. V opačném případě jsou hodnoty uvedeny jako průměr \pm směrodatná odchylka u normálně rozložených dat a jako medián (mezikvartilové rozpětí) u dat s jiným než normálním rozdělením. Vzhledem k tomu, že je srovnání jednotlivých variant modelu provedeno vždy na stejné sadě dat a primárním cílem je ověřit, zdali specifické nastavení vede k systematickému zvýšení či snížení metriky kvality, je významnost rozdílů mezi variantami testována pomocí párového Studentova t-testu, případně jeho neparametrické varianty u dat s jiným než normálním rozdělením. Rozdíly v rozptylu hodnot reziduí byly testovány párově oproti výchozímu modelu pomocí Levenova testu. Významnost hyperparametrů z pohledu dosažení nejvyšší kvality modelu byla stanovena pomocí funkcionální varianty analýzy rozptylu (fANOVA) [110].

2.11 Výsledky

2.11.1 Optimalizace architektury

Počet iterací Bayesovské optimalizace byl nastaven na minimální doporučenou hodnotu 200, což lze vzhledem k nalezení optima v prvních 100 iteracích považovat za dostatečné. Aby byl snížen vliv náhodné inicializace modelu, proběhly v každé iteraci celkem tři experimenty, z nichž maximální hodnota byla využita pro potřeby TPE algoritmu. Celkový výpočetní čas činil cca 240 hodin.

Výsledky pro trojici nastavení, při kterém bylo dosaženo nejlepší hodnoty metriky F_1 , jsou uvedeny v Tabulkách 2.5 a 2.6. Všechny tři varianty dosahují obdobné hodnoty F_1 skóre okolo 0.887, a to pro relativně široký rozsah nastavení a komplexnost modelu, viz Tabulka 2.6. Další dílčí výsledky týkající se významnosti jednotlivých hyperparametrů jsou znázorněny na Obrázcích 2.12 a 2.13. Na Obrázku 2.12 jsou uvedena relativní skóre významnosti pro (a) skupinu hyperparametrů reprezentující architekturu modelu a (b) skupinu všech optimalizovaných hyperparametrů. Na Obrázku 2.13 jsou pak vyobrazeny obrysové grafy závislostí mezi dvojicí hyperparametrů a výsledného F_1 skóre.

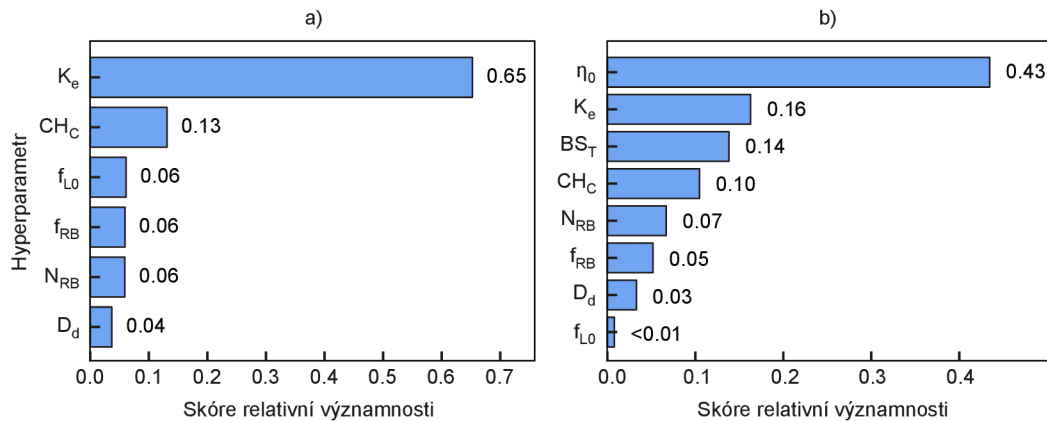
Tab. 2.5: Modely s nejvyšším dosaženým F_1 skóre po Bayesovské optimalizaci architektury. DS- A_T – trénovací sada; DS- A_V – validační sada; Exp. – pořadí optimalizačního kroku.

Exp.	DS- A_T			DS- A_V		
	F	Se	P ⁺	F	Se	P ⁺
15	0.899	0.895	0.904	0.887	0.917	0.861
70	0.904	0.896	0.912	0.887	0.910	0.864
98	0.869	0.816	0.921	0.884	0.908	0.860

Tab. 2.6: Hyperparametry modelů s nejvyšším dosaženým F_1 skóre po Bayesovské optimalizaci architektury. Ve sloupci zcela vpravo je uveden celkový počet trénovatelných koeficientů jednotlivých variant. Exp. – pořadí optimalizačního kroku.

Exp.	Hyperparametry								Počet p. modelu
	η_0	BS_T	f_{RB}	f_{L0}	K_e	N_{RB}	CH_C	D_d	
15	0.005	16	5	7	6	6	32	5	1298928
70	0.005	24	5	17	5	7	48	4	413740
98	0.005	16	7	17	5	6	40	4	414389

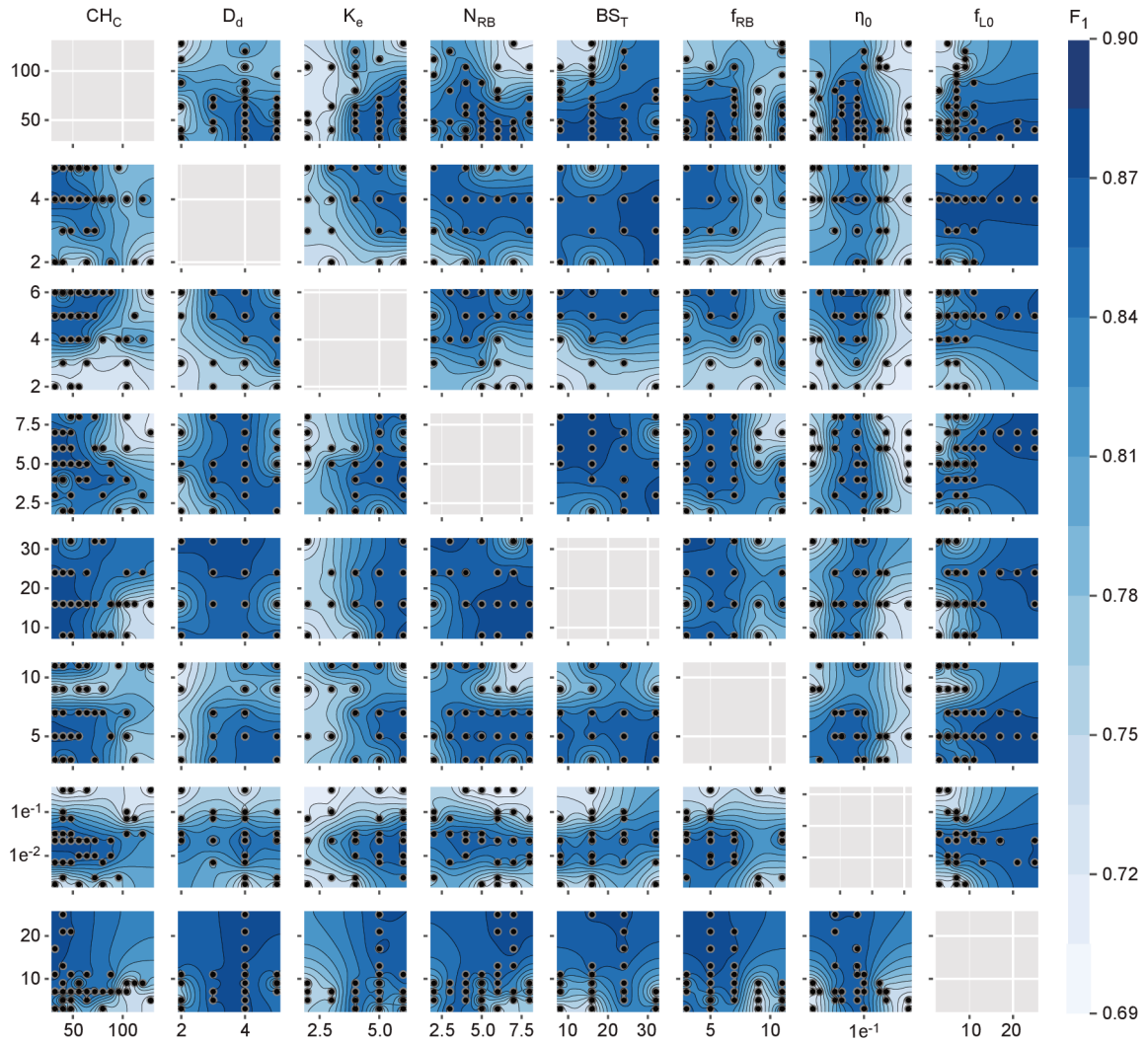
Dle očekávání má největší vliv na dosažené F_1 skóre výchozí krok učení η_0 s celkovým skóre relativní významnosti (fANOVA) 0.43. Při zohlednění pouze těch hyperparametrů, které souvisejí s architekturou modelu, se nejvýznamněji (relativní skóre 0.65) na F_1 skóre podílel parametr K_e . Ten reprezentuje maximální hloubku a šířku modelu, resp. počet v sérii zařazených vrstev enkodéru. Jedná se o parametr, který v rámci navržené architektury nejvíce ovlivňoval kapacitu modelu extrahovat relevantní příznaky. Současně však navýšil počet trénovatelných koeficientů a tím i výpočetní náročnost modelu. To je patrné při srovnání variant s $K_e = 5$ a $K_e = 6$ vrstvami v Tabulce 2.6. Za optimální model M_{aopt} byla vybrána varianta z experimentu č. 70, která dosahuje stejného F_1 skóre, ovšem s třetinovým počtem trénovatelných koeficientů. Přidání poslední vrstvy s 256 konvolučními filtry nepřineslo z pohledu výsledků segmentace žádný další benefit.



Obr. 2.12: Relativní skóre významnosti jednotlivých hyperparametrů z pohledu cílů optimalizace pro: a) hyperparametry architektury modelu; b) všechny optimalizované hyperparametry.

Na Obrázku 2.12 je uveden přehled jednotlivých hyperparametrů a jejich významnost z pohledu optimalizace modelu. Z těch, které mají dopad na celkovou kapacitu modelu, patří mezi nejvýznamnější (relativní skóre 0.13) počet kompresních filtrů CH_C v dekodérové části. Zajímavostí je, že při navýšení nad hodnotu cca 100 filtrů (viz Obrázek 2.13) došlo

k systematickému poklesu F_1 skóre. Model tedy více těžil z komprimované reprezentace příznakových map než z jejich přeparametrizování. Optimální hodnota se dle výsledků pohybovala v oblasti 32–48 filtrů, což pro navržený enkodér představovalo zhruba průměr počtu filtrů v jednotlivých vrstvách. Z pohledu konstrukce reziduálních bloků byly pro řešení zadaného problému preferovány varianty s užší receptivní oblastí filtrů f_{RB} v rozsahu 5–7 vzorků, avšak s navýšením počtu jejich sériového řazení N_{RB} . Jak v případě receptivní šířky vstupního filtru, tak velikosti trénovací dávky bylo možné dosáhnout vysokého F_1 skóre pro různé varianty nastavení.



Obr. 2.13: Obrysové grafy znázorňující vztah mezi dvojicí hyperparametrů modelu M_{aopt} a F_1 skóre na DS-A_V. BS_T - velikost trénovací dávky; f_{RB} - šířka konvolučního filtru v reziduálních blocích; D_d - počet vrstev s dilatovanými konvolucemi; K_e - počet vrstev v enkodéru; N_{RB} - počet reziduálních bloků v jedné vrstvě; CH_C - počet kompresních filtrů dekodéru; f_{L0} - šířka vstupního konvolučního filtru; η_0 - úvodní krok učení. Pořadí bloků na vertikální a horizontální ose je totožné.

2.11.2 Optimalizace kritériální funkce

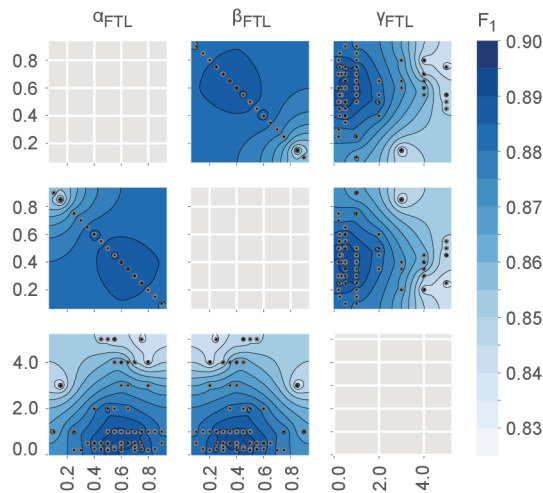
V rámci hyperoptimalizace parametrů kritériální funkce FTL (rovnice 2.3) bylo provedeno celkem 80 experimentů. V každém experimentu byla realizována tři nezávislá měření se stejným nastavením, z nichž byla pro optimalizační algoritmus vybrána hodnota s nejvyšším F_1 skóre. V Tabulce 2.7 uvádíme výsledky 3 nejlepších modelů. Jako M_{lopt} je dále označen model z experimentu č. 21, který z pohledu cílů optimalizace dosáhl nejvyššího skóre.

Tab. 2.7: Hyperparametry kritériální funkce FTL získané pomocí Beysovské optimalizace. Uvedeny první 3 varianty s nejvyšším F_1 na DS- A_V .

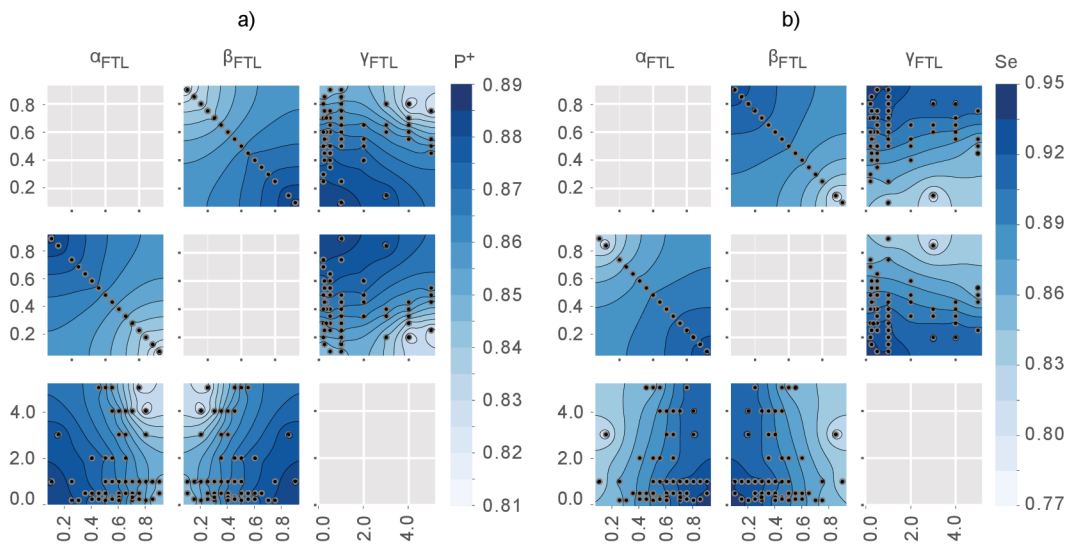
Exp.	DS- A_T			DS- A_V			Hyperparametry		
	F_1	Se	P^+	F_1	Se	P^+	α_{FTL}	β_{FTL}	γ_{FTL}
21	0.898	0.900	0.895	0.889	0.883	0.894	0.4	0.6	0.5
48	0.885	0.906	0.866	0.888	0.902	0.876	0.5	0.5	1.0
41	0.873	0.876	0.871	0.888	0.863	0.915	0.4	0.6	0.5

V případě F_1 se pohybovalo nejlepší dosažené skóre v okolí hodnoty 0.888, a to i pro hyperparametry $\alpha_{\text{FTL}} = \beta_{\text{FTL}} = 0.5$ a $\gamma_{\text{FTL}} = 1$. To odpovídá nastavení penalizace u modelu M_{aopt} . S tím souvisí i nižší relativní významnost obou parametrů na metriku F_1 s hodnotou 0.251. Zásadní vliv na dosažené skóre se projevil u exponenciální váhy γ_{FTL} s relativní významností 0.749. Z Obrázku 2.14 je patrné, že se optimální nastavení γ_{FTL} pohybovalo v rozmezí 0.25–1.00. V relativním srovnání byly tedy více penalizovány obtížně klasifikovatelné vzorky signálu. V průměrném případě naopak model dosahoval dostatečné míry jistoty na to, aby koeficient γ_{FTL} mohl snížit jejich příspěvek k penalizaci.

Přestože hodnoty Se a P^+ nebyly při optimalizaci zohledněny, znázorňujeme jejich vztah k nastavení hyperparametrů na Obrázku 2.15. Pro výchozí hodnoty $\alpha_{\text{FTL}} = \beta_{\text{FTL}} = 0.5$ dosahoval model významně vyšší hodnoty Se (0.902) na úkor chybné klasifikace negativní třídy. Zvýšením parametru β_{FTL} došlo ke změně poměru mezi oběma metrikami směrem k vyšším hodnotám P^+ . U modelu s nejvyšším F_1 v Tabulce 2.7 dosahuje P^+ hodnoty 0.894 se současným snížením počtu správně zařazených vzorků v případě pozitivní klasifikační třídy. Tato závislost odpovídá Rovnici 2.3, která při zvýšení β_{FTL} model více trestá za falešnou pozitivitu. Konvenčně je podobných cílů dosahováno pomocí změny rozhodovacího prahu na normalizovaných výstupních hodnotách modelu. Zajímavostí je, že změna prahu u modelů M_{aopt} i M_{lopt} v rozmezí hodnot 0.1–0.9 neměla zásadní vliv ani na jednu z metrik. V případě Se i P^+ dochází v obou případech k posunu maximálně o 0.004, respektive 0.007 bodů.



Obr. 2.14: Obrysové grafy znázorňující vztah mezi hyperparametry kritériální funkce α_{FTL} , β_{FTL} a γ_{FTL} a F_1 skóre na DS- A_V . Pořadí řádků v matici grafů odpovídá pořadí sloupců.

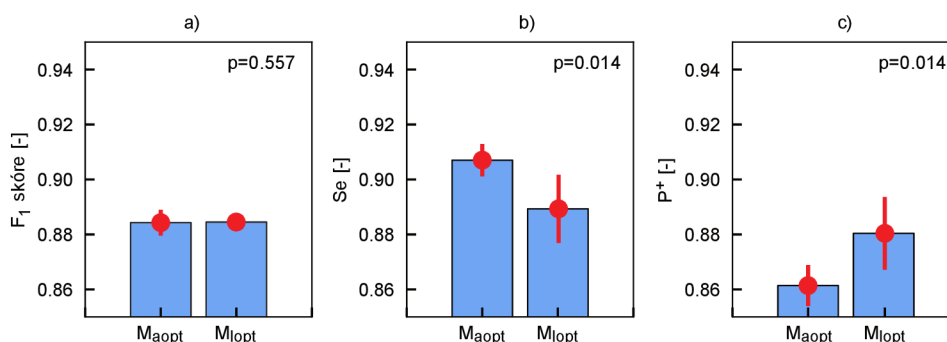


Obr. 2.15: Obrysové grafy znázorňující vztah mezi (a) senzitivitou, (b) pozitivní predikční hodnotou a hyperparametry kritériální funkce na DS- A_V . Pořadí řádků v matici grafů odpovídá pořadí sloupců.

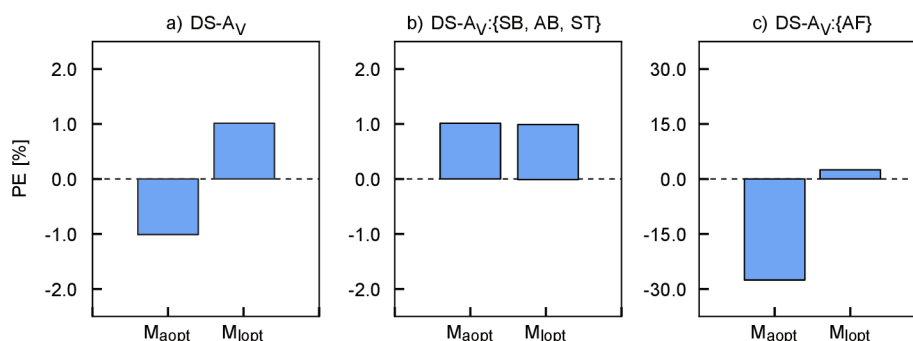
2.11.3 Srovnání modelů M_{aopt} a M_{lopt}

Aby byl snížen vliv náhodné inicializace parametrů na metriky kvality, byla pro každý z modelů provedena studie zahrnující 10 nezávislých učebních experimentů. Výsledky jsou shrnuty na Obrázku 2.16. Z výsledků lze potvrdit předpoklad stanovený v předchozí sekci. Optimalizace kritériální funkce neměla významný vliv na hodnotu F_1 ($p=0.557$). Naopak dosažené skóre v ostatních metrikách se pro oba modely významně lišilo. M_{aopt} dosaho-

val průměrně o 0.021 bodů vyšší Se ($p=0.014$), avšak za cenu systematického snížení P^+ v průměru o 0.019 bodů ($p=0.014$). Centrální tendence obou metrik byly v případě učení pomocí optimalizované kriteriální funkce u M_{lopt} téměř totožné. Poskytují proto lépe vyvážený poměr mezi Se a P^+ . Za pozornost stojí zvýšená míra rozptylu při učení modelu M_{lopt} jak u Se, tak P^+ . V tomto kontextu je nutné zmínit implementační specifika kriteriální funkce. Pro hodnotu $\gamma_{FTL} = 1$ u modelu M_{aopt} není z logických důvodů exponent v implementaci zahrnut, což mohlo ovlivnit numerickou stabilitu výpočtu.



Obr. 2.16: Srovnání metrik kvality (a) F_1 (b) Se a (c) P^+ na DS- A_V pro dvojici optimalizovaných modelů M_{aopt} a M_{lopt} . Bodový a intervalový odhad proveden z 10 nezávislých experimentů.



Obr. 2.17: Srovnání PE na (a) kompletní DS- A_V , (b) záznamech bez fibrilace síní a (c) pouze záznamech s fibrilací síní pro dvojici optimalizovaných modelů M_{aopt} a M_{lopt} . Bodové odhady provedeny pro nejlepší modely z obou typů.

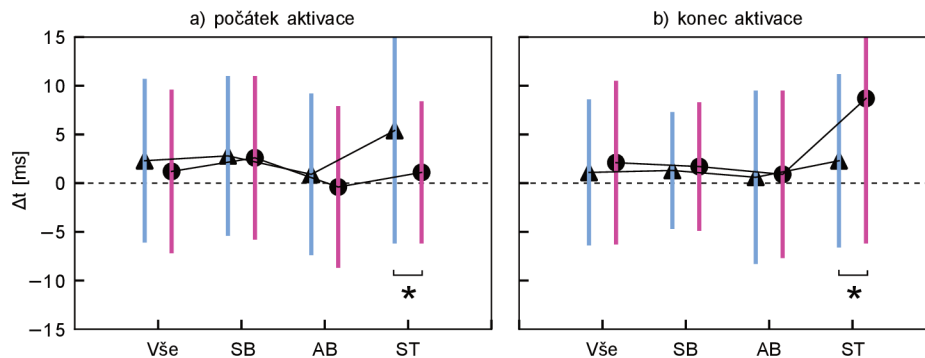
Na Obrázku 2.17 je srovnána kontinuita klasifikovaných segmentů pro oba modely pomocí metriky PE (blíže nule=lepší). V případě celé datové sady a segmentů reprezentující izolované síňové stahy sinusového (SB), nesinusového (AB) a stimulovaného (ST) původu, dosahovali oba modely prakticky shodné chyby. Ta zhruba o 1 % převyšovala počet expertem stanovených úseků. Významné zhoršení kvality segmentace bylo možné spatřit v případě záznamů s fibrilací síní, viz Obrázek 2.17 (c). Přestože u modelu M_{lopt} došlo k

nárůstu o 1.5 % oproti skupině (b), jednalo se stále o výsledek, který reprezentuje malou míru diskontinuity v klasifikovaných segmentech. U modelu M_{aopt} byl naopak přítomen významný pokles identifikovaných segmentů, a to o celých 27.3 %. V případě AF záznamů tak zřejmě u M_{aopt} docházelo ke spojení sémanticky nesouvisejících objektů.

Chyba stanovení pozice počátku a konce síňové aktivity byla provedena jak pro celou datovou sadu, tak pro jednotlivé typy síňových stahů. Z tohoto hodnocení byly vyjmuty segmenty označující fibrilaci síní, protože se jedná o koncepčně odlišný typ detekce s přesností vychýlenou o několik řádů. Výsledky analýzy jsou uvedeny na Obrázku 2.18 a v Tabulce 2.8.

Tab. 2.8: Srovnání přesnosti stanovení počátku a konce síňové u stahů SB, AB a ST.

Model	RMSE počátku [ms]				RMSE konce [ms]			
	Vše	SB	AB	ST	Vše	SB	AB	ST
M_{aopt}	8.4	8.4	8.3	7.3	8.4	6.6	8.6	14.9
M_{lopt}	8.4	8.2	8.3	11.6	7.5	6.0	8.9	8.9



Obr. 2.18: Srovnání přesnosti stanovení počátku a konce síňové aktivity typu SB, AB a ST pro dvojici optimalizovaných modelů M_{aopt} (plný kruh s purpurovými vousy) a M_{lopt} (plný trojúhelník s modrými vousy). Symbol * značí statisticky významný rozdíl mezi skupinami.

Při stanovení počátku síňové aktivity dosáhly oba modely shodné hodnoty RMSE 8.4 ms. U konce síňové aktivity si lépe vedl model M_{lopt} s hodnotou 7.5 ms ve srovnání s 8.4 ms u M_{aopt} . V případě síňových stahů sinusového (SB) a abnormálního původu (AB) dosahují oba modely téměř stejné průměrné odchylky od referenčních hodnot. Kategorie AB přitom zahrnuje i fúzané stahy síňové a komorové aktivity. Významných rozdílů (Levenův test) u obou modelů bylo pozorováno v případě kategorie ST, tedy u stimulovaných rytmů. V případě modelu M_{lopt} došlo ke zhoršení ($p=0.01$) RMSE až na 11.6 ms oproti 7.3 ms u M_{aopt} . M_{lopt} v tomto případě systematicky nadhodnocoval počátek aktivity s průměrnou hodnotou reziduí 5.4 ms. Opačná situace nastala u stanovení konce síňového stahu.

V tomto případě významně ($p < 0.01$) zaostával model M_{aopt} s RMSE 14.9 ms oproti 8.9 ms u M_{lopt} . Současně došlo k nadhodnocení konce s průměrem reziduí 8.7 ms.

Z pohledu distribuce absolutních chyb bylo v případě M_{lopt} celkově 54.0 % a 81.0 % odchylek ≤ 5 ms, respektive ≤ 10 ms. U modelu M_{aopt} bylo pozorováno mírné zhoršení s 50.2 % absolutních odchylek v kategorii ≤ 5 ms a 80.6 % odchylek bylo menších než 10 ms. Vzhledem k výsledkům srovnání obou modelů je další hodnocení kvality v následujících sekcích reportováno pro model M_{lopt} .

2.11.4 Segmentace fibrilačních záznamů

U kvality segmentace fibrilačních záznamů (AF) bylo předmětem zájmu primárně celkové F_1 skóre a kontinuita detekovaných segmentů PE. Testování bylo provedeno pro model M_{lopt} , který na základě předchozích výsledků dosáhl významně lepší hodnoty PE v případě AF než M_{aopt} . Testování bylo založeno na jednoduché ablační studii se zařazením dvou scénářů, u kterých byl předpokládán největší dopad na výsledky segmentace:

- Model bez dilatovaných konvolučních masek (experiment M_{lopt} w/o dil). Pro eliminaci vlivu počtu parametrů modelu byl blok zachován, ale dilatační faktor byl nastaven na hodnotu 1.
- Model učený i testovaný bez přítomnosti anotací z kategorie AF (experiment M_{lopt} w/o AF).

Výsledky pro podmnožinu $DS-A_V:\{SB, AB, ST\}$ a podmnožinu fibrilačních záznamů $DS-A_V:\{AF\}$ jsou uvedeny v Tabulce 2.9. První řádek tabulky odpovídá referenčnímu modelu bez aplikovaných změn. V případě M_{lopt} w/o dil. bylo při učení i testování zachováno zařazení fibrilačních záznamů do pozitivní třídy. Pro M_{lopt} w/o AF byly fibrilační úseky zařazeny do negativní třídy, tj. anotace jsou zcela vyřazeny z učebního i testovacího procesu.

Tab. 2.9: Výsledky ablační studie u modelu M_{lopt} pro různé AF scénáře.

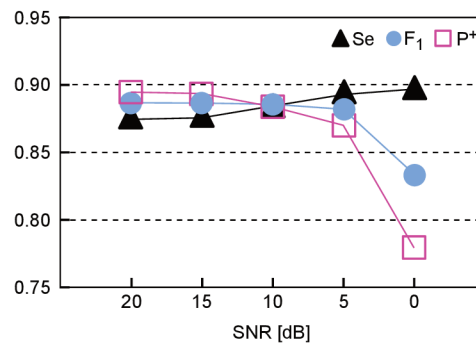
Exp.	DS- $A_V:\{SB, AB, ST\}$				DS- $A_V:\{AF\}$			
	F_1	Se	P^+	PE [%]	F_1	Se	P^+	PE [%]
M_{lopt} - ref. model	0.889	0.883	0.894	0.97	0.886	0.877	0.894	2.48
M_{lopt} w/o dil.	0.884	0.865	0.903	1.03	0.866	0.913	0.823	-33.06
M_{lopt} w/o AF	0.878	0.872	0.885	2.33	0.618	0.671	0.573	42.11

Při inaktivaci bloku dilatovaných konvolucí došlo ke snížení F_1 skóre o 0.5 % na podmnožině s SB, AB a ST síňovými stahy, resp. o 2.0 % na výběru záznamů s AF. Současně model zaznamenal zhoršení PE na $DS-A_V:\{AF\}$. Celkově v tomto případě došlo o snížení počtu detekovaných segmentů o více než 30 % oproti počtu referencí. Nárůst falešně pozitivních vzorů indikovaný snížením P^+ v rámci $DS-A_V:\{AF\}$ na hodnotu 0.823 nemusí

v kontextu datasetu znamenat zhoršení schopnosti identifikace AF. Vzhledem k tomu, že se v záznamech nezdá objevují úseky vzniku a terminace AF, může být toto snížení důsledkem relativně vysoké míry nejistoty stanovení referenčních pozic v těchto oblastech. Model M_{lopt} trénovaný op úplném odstranění AF anotací vykazoval systematické zhoršení ve všech sledovaných metrikách. A to překvapivě i v případě záznamů bez přítomné AF (Tabulka 2.9). Zde počet detekovaných segmentů narostl z původních 0.97 % na 2.33 % Na podmnožině $DS-A_V:\{AF\}$ klesla hodnota F_1 , Se i P^+ pod kritickou hranicí 0.7, což značí velmi špatnou schopnost zařazení AF úseků do – v tomto specifickém případě – negativní třídy. Tomu odpovídal i nárůst hodnoty PE na 42.11 %.

2.11.5 Hodnocení vlivu šumu

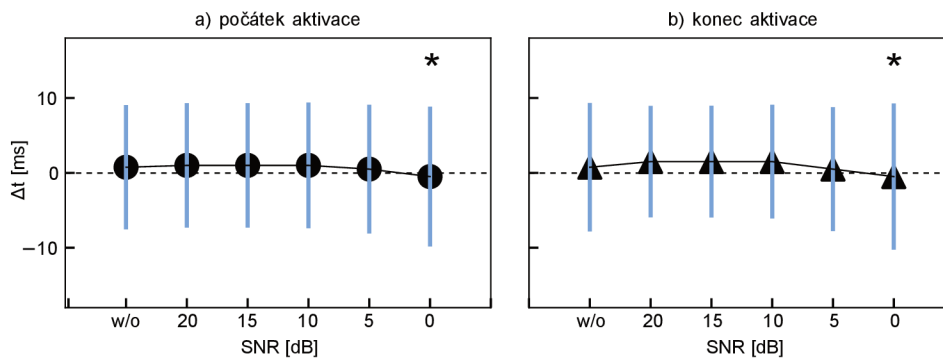
V této sekci jsou prezentovány výsledky segmentace pro tři různé typy uměle generovaného rušení. V prvním případě se jedná o aditivní bílý šum Gaussovského typu, který má širokospektrální charakter a může tak imitovat elektrické chování tkáně při fibrilaci síní. Do hodnocení byly zahrnuty hladiny poměru signál–šum (SNR) v rozmezí 0–20 dB. Šum byl pro každou hladinu aplikován do všech EGM kanálů. Na Obrázku 2.19 je uveden vliv přítomnosti bílého šumu na metriky F_1 , Se a P^+ . Vliv na přesnost stanovení počátku a konce síňové aktivity je vyobrazen na Obrázku 2.20.



Obr. 2.19: Srovnání F_1 , Se a P^+ v závislosti na hladině bílého šumu u $DS-A_V$.

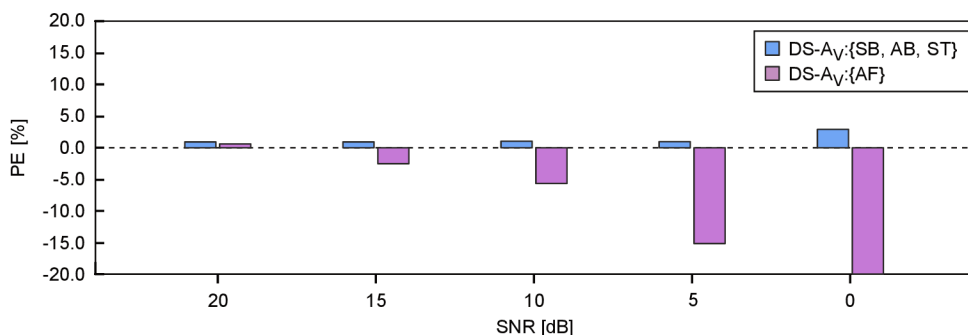
Jak u metrik odvozených z matice záměn, tak v případě přesnosti stanovení hranic síňové aktivity bylo kritickou hladinou SNR 5 dB. Hodnota F_1 skóre klesla při snížení SNR z 5 dB na 0 dB z hodnoty 0.882 na 0.835. Při pohledu na zbylé dvě metriky je zřejmým důvodem nárůst falešné pozitivivity, čemuž odpovídal rozdíl v P^+ kolem 0.09 bodů při snížení SNR z 5 na 0 dB. Příčinou zvýšené falešné pozitivivity byla primárně tendence modelu interpretovat bílý šum jako fibrilaci síní. Vizuální demonstrace je uvedena v následující sekci. Hodnota PE zaznamenala prakticky zanedbatelný nárůst v rozsahu 0.20–0.29 při všech úrovních šumu.

Pokles kvality segmentace za hranicí 5 dB byl patrný i na přesnosti stanovení konce síňových segmentů. Na Obrázku 2.20 je znázorněna závislost průměrné chyby a RMSE pro jednotlivé hladiny SNR. Zatímco do hranice 5 dB byly rozdíly v reziduích mezi detekovanou a referenční pozicí prakticky minimální, na úrovni 0 dB se RMSE zvýšilo v případě stanovení počátku i konce segmentu. Oproti výchozímu stavu došlo v případě počátku aktivace k nárůstu RMSE z 8.4 ms na 8.7 ms. Ten je však na hranici statistické významnosti ($p=0.02$). Stanovení konce aktivace při SNR 0 dB doprovázel nárůst z 7.5 na 9.6 ms ($p<0.0001$). Přesnost určení konce segmentu byla na rozdíl od počátku ovlivněna i při hladině SNR 5 dB. Zde došlo ke zvýšení ($p=0.02$) na 8.3 ms oproti variantě w/o. Z grafů je dále patrný trend směrem od systematického nadhodnocení pozic s kladnou průměrnou chybou směrem ke striktněji určeným hranicím. Rozdíly pro jednotlivé úrovně šumu byly testovány párově proti výchozímu stavu (w/o) bez aditivního zarušení s hladinou $\alpha = 0.01$ po Bonferonniho korekci.



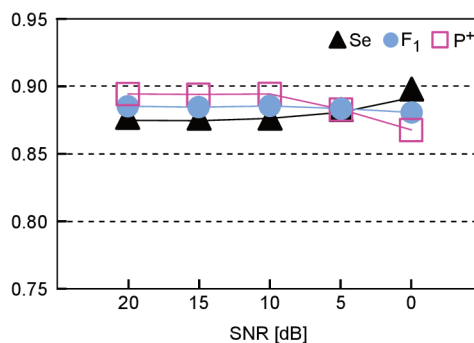
Obr. 2.20: Vliv hladiny bílého šumu na přesnost nalezení počátku a konce síňové aktivity v celé DS-A_V. Zkratka w/o označuje referenční model na signálech bez přidaného šumu. Symbol * značí statisticky významný rozdíl oproti w/o.

Aditivní bílý šum prokázal významně negativní dopad na segmentaci izolovaných síňových stahů, které se nachází v záznamech s přechodem fibrilace síní do jiného, více organizovaného rytmu. Již od hladiny 10 dB docházelo k poklesu celkového počtu identifikovaných EGM komplexů, a to o 5.8 % až o výsledných 27.3 % v případě 0 dB, viz Obrázek 2.21. Takový výsledek naznačuje buď úplné selhání identifikace izolovaných komplexů nebo naopak jejich zahrnutí do segmentu, který náleží AF.



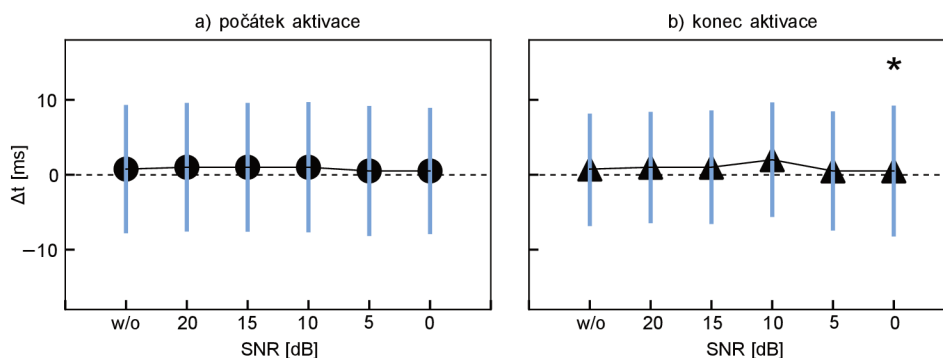
Obr. 2.21: Vliv hladiny bílého šumu na přesnost určení počtu izolovaných EGM.

U zarušení síťovým brumem byl zaznamenán pokles kvality segmentace obdobně od hranice SNR 5 dB. Ovlivněna byla primárně hodnota P^+ , která byla při 5 dB snížena oproti výchozímu stavu o 0.011 bodů. Při 0 dB došlo následně k poklesu až o 0.024 bodů. Zhoršení metriky F_1 bylo kompenzováno nárůstem senzitivity, avšak jen v případě záznamů s fibrilací síní, a to z hodnoty 0.877 na 0.930. U segmentace zbylých typů síňových stahů docházelo při srovnání s účinností na nezarušených datech ke snížení F_1 o 0.031, Se o 0.013 a P^+ o 0.026 bodů. Zmíněná závislost je demonstrována v grafu na Obrázku 2.22.



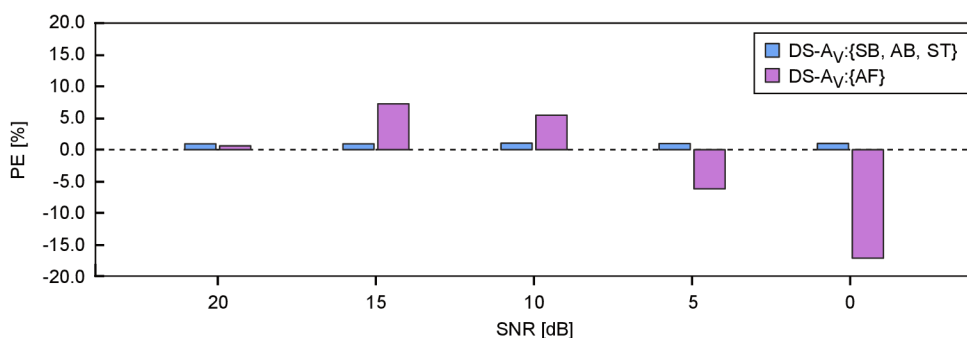
Obr. 2.22: Srovnání F_1 , Se a P^+ v závislosti na hladině síťového brumu u DS-A_V.

Z Obrázku 2.23 je zřejmé, že v případě správné identifikace síňového komplexu byla zachována přesnost nalezení počáteční hranice segmentu, a to i při hladině šumu 0 dB. Hodnoty RMSE se pohybují u počátku v rozsahu 8.4–8.6 ms, bez statisticky významných rozdílů v rozptylu reziduí. U přesnosti stanovení koncové hranice se hodnoty RMSE pohybují v rozmezí 7.5–8.7 ms, kde nejvýraznějšího zhoršení na 8.7 ms bylo dosaženo u hladiny SNR 0 dB ($p < 0.001$).



Obr. 2.23: Vliv hladiny síťového brumu na přesnost nalezení počátku a konce síňové aktivity v celé DS- A_V . Zkratka w/o označuje referenční model na signálech bez přidaného šumu.

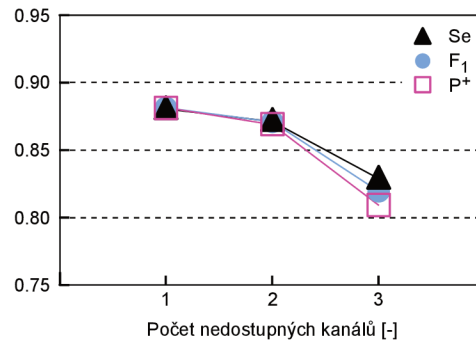
Síťový brum měl dále dopad na stabilitu detekce individuálních komplexů v případě fibrilačních EGM. Při hladině SNR 15 a 10 dB docházelo k navýšení počtu detekovaných segmentů oproti počtu referencí o 7.4, resp. 5.8 %. Od hladiny 5 dB pak byla zřejmá změna trendu a počet detekovaných segmentů byl nižší než počet anotací s celkovým poklesem při 0 dB o 17.4 %. Tato závislost je vyobrazena na Obrázku 2.24.



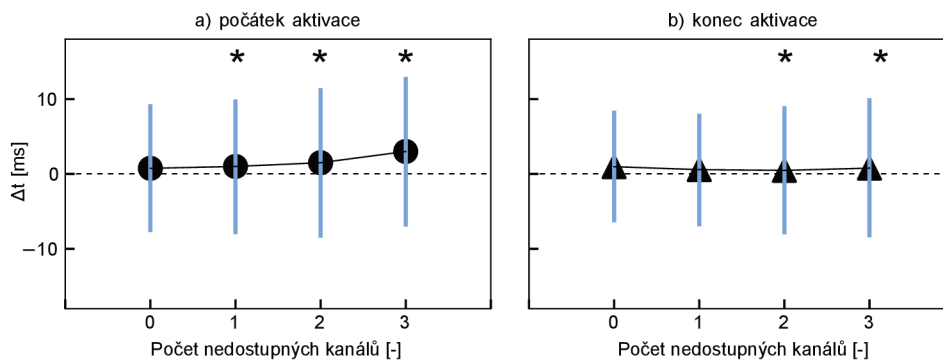
Obr. 2.24: Vliv hladiny síťového brumu na přesnost stanovení počtu izolovaných EGM.

Odstranění montáže v případě zarušení svodů patří mezi běžná akutní řešení problémů technického charakteru v průběhu EFV. Z tohoto důvodu bylo provedeno hodnocení kvality segmentace v závislosti na počtu chybějících kanálů. Vzhledem k tomu, že vstup modelu vyžaduje konstantní počet kanálů, byly vybrané svody nahrazeny sekvencí nul. Aby byl snížen vliv pořadí, byla selekce nulovaného kanálu pro každý záznam provedena náhodně.

Z Obrázku 2.25 je patrné, že k poklesu všech metrik kvality docházelo systematicky již po odstranění 1 z kanálů. Hodnota F_1 skóre byla snížena při odstranění dvou a tří kanálů na 0.872, respektive 0.823 bodů. Ovlivněna byla i přesnost stanovení hranic jednotlivých komplexů. Hodnota RMSE postupně narůstala z výchozích 8.4 ms až 10.0 ms v případě počátku segmentu a z 7.5 ms až k 9.3 ms u koncové hranice, viz Obrázek 2.26.



Obr. 2.25: Srovnání F_1 , Se a P^+ pro model M_{lopt} v závislosti na počtu dostupných kanálů na celé DS- A_V .



Obr. 2.26: Vliv počtu dostupných kanálů na přesnost nalezení (a) počátku a (b) konce síňové aktivity pro model M_{lopt} v celé DS- A_V . Symbol * označuje statisticky významný rozdíl oproti výchozímu stavu.

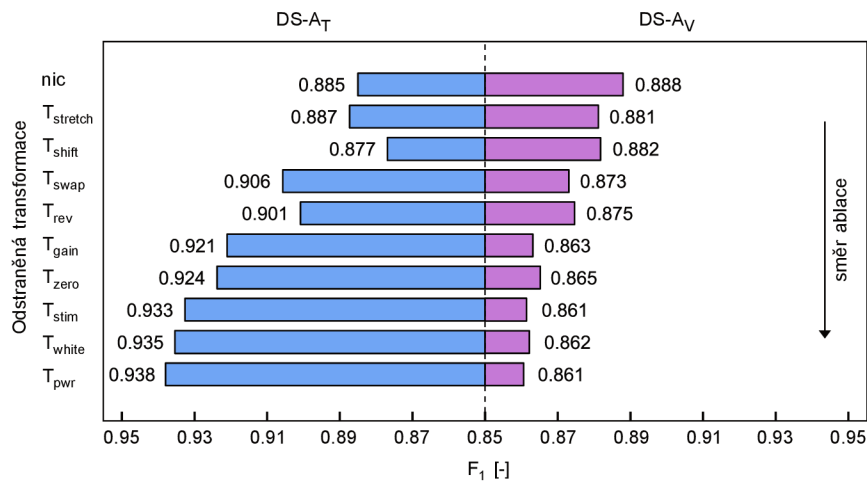
Rozdíly v rozptylu reziduí oproti segmentaci záznamů bez odstraněných kanálů jsou přitom statisticky významné již po odstranění prvního kanálu v případě počátku ($p=0.011$, následně $p<0.0001$), a od druhého kanálu u stanovení konce (v obou případech $p<0.0001$). Graf (a) na Obrázku 2.26 dále ukazuje rostoucí nadhodnocení detekce počátku v závislosti na počtu nulovaných kanálů, a to v průměru až o 3.6 ms v nejhorším případě. Hladina statistické významnosti $\alpha = 0.0166$ byla korigována dle Bonferonniho.

2.11.6 Vliv augmentace dat

Vliv jednotlivých augmentačních technik na výsledky segmentace EGM je znázorněn na Obrázku 2.27. Výsledky byly získány pomocí ablační studie. Princip ablační studie byl založen na opakovaném učení modelu. V každém učebním cyklu byla provedena náhodná inicializace jeho parametrů a současně byl deaktivován jeden z augmentačních bloků. Postupně tedy byly odstraněny všechny aplikované transformace. Pro snížení vlivu náhodné inicializace byly v každém kroku realizovány 3 experimenty. Z nich na Obrázku 2.27 je uveden jejich medián.

Z výsledků je patrné, že k největšímu poklesu F_1 metriky na DS- A_V došlo po odstranění náhodného zesílení signálu (T_{gain}) a bloku pro prohození pořadí dvou kanálů (T_{swap}), a to o 0.012 respektive 0.009 bodů. Třetí nejvýznamnější dopad měl blok pro roztažení signálu v časové ose (T_{stretch}) s poklesem o 0.007 bodů. Skupina augmentací založená na vkládání uměle generovaného šumu neměla z pohledu F_1 skóre na výsledky významný vliv. Odlišná situace by mohla nastat při cílené selekci zarušených signálů, či po vložení šumu do DS- A_V . Tento scénář nebyl pro zachování přehlednosti výsledků uvažován.

Celkově došlo ke snížení F_1 na DS- A_V o 0.027 bodů. Jedná se o mírné zhoršení oproti výchozímu stavu. Naopak na trénovací sadě DS- A_T je zřejmé významné navýšení F_1 po odstranění zhruba první poloviny augmentačních technik. Reprezentace síňových signálů hypoteticky ztrácí na dostatečné tvarové variabilitě, která by zabraňovala přeučení modelu.



Obr. 2.27: Vliv augmentačních technik na F_1 skóre v průběhu ablační studie u modelu M_{lopt} . Vlevo výsledky na trénovací sadě DS- A_T ; vpravo na validační sadě DS- A_V .

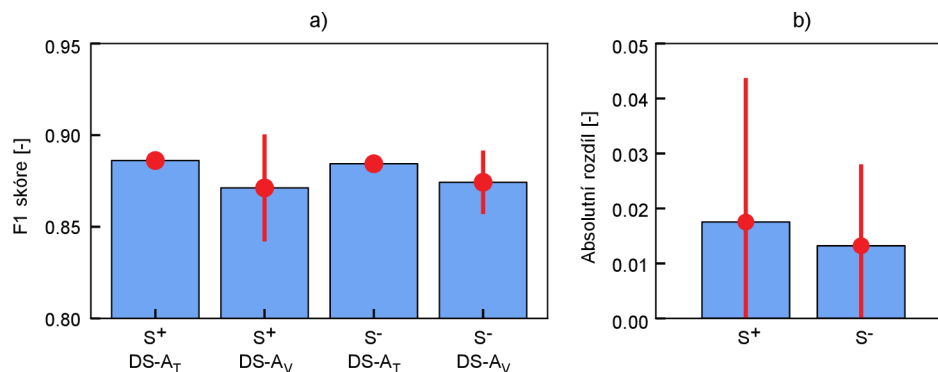
Kromě uvedených technik byla provedena dvojice experimentů s augmentací dat založená na vytvoření aditivní směsi dvou nezávislých EGM signálů. V prvním případě dochází k součtu dvou různých záznamů, které jsou váhovány na základě poměru generovaného z rozložení pravděpodobnosti β (tzv. metoda *MixUp*). V druhém případě je do EGM signálu zařazen jeden či více krátkých EGM segmentů, které byly náhodně vybrány z jiného záznamu (tzv. metoda *CutMix*). V obou případech se nepodařilo dosáhnout nastavení

parametrů, které by vedly ke stabilnímu procesu učení, a proto tyto metody nebyly do řetězce předzpracování zařazeny.

2.11.7 Vliv stratifikace datové sady

Při náhodném generování trénovací (DS- A_T) a validační (DS- A_V) datové sady došlo pouze v 24.4 % případů z celkem 1000 pokusů k vygenerování dvou subsetů se statisticky shodným rozdělením (S^-) diagnostických tříd (Pearsonův χ^2 , $p=0.994$). Relativní entropie mezi rozděleními pro množinu semínek S^- dosáhla $(18.19 \pm 5.78) \times 10^{-3}$, zatímco u množiny semínek s významnými rozdíly (S^+) v distribuci tříd dosáhla $(53.20 \pm 28.48) \times 10^{-3}$. Vliv rozdělení na věrohodnost měřené metriky kvality pro skupiny semínek S^+ a S^- je zobrazen na Obrázku 2.28 na následující straně.

V průměrném případě dosáhne na (DS- A_T) i (DS- A_V) model takřka stejných výsledků, bez ohledu na způsob rozdělení datasetu. Při opakovaném učení totožného modelu je však v případě skupiny S^+ metrika kvality významně více rozptýlena (F-test, $p<0.001$) než u S^- . Při stratifikaci dat podle diagnostických tříd došlo dále k mírnému snížení rozdílů ve sledované metrice mezi DS- A_T a validační DS- A_V , byť na hranici statistické významnosti (Wilcoxonův test, $p=0.022$).

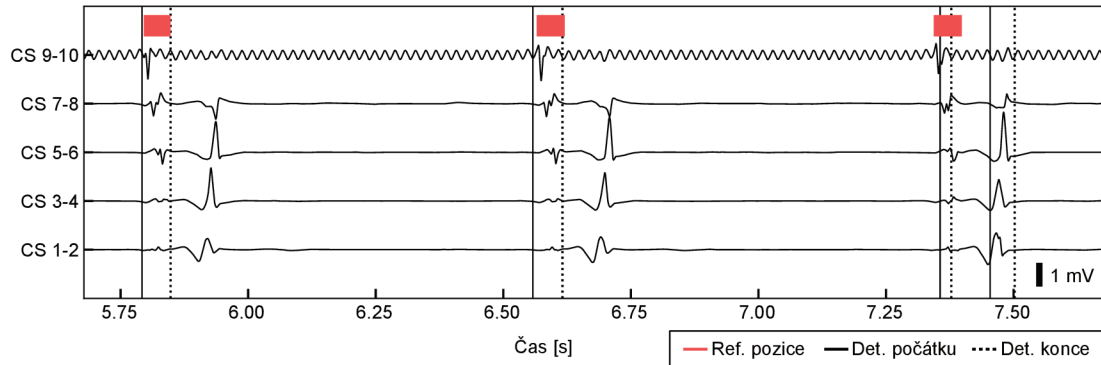


Obr. 2.28: Krabicové grafy F_1 hodnot M_{aopt} pro stratifikované a nestratifikované rozdělení dat do trénovací a validační množiny.

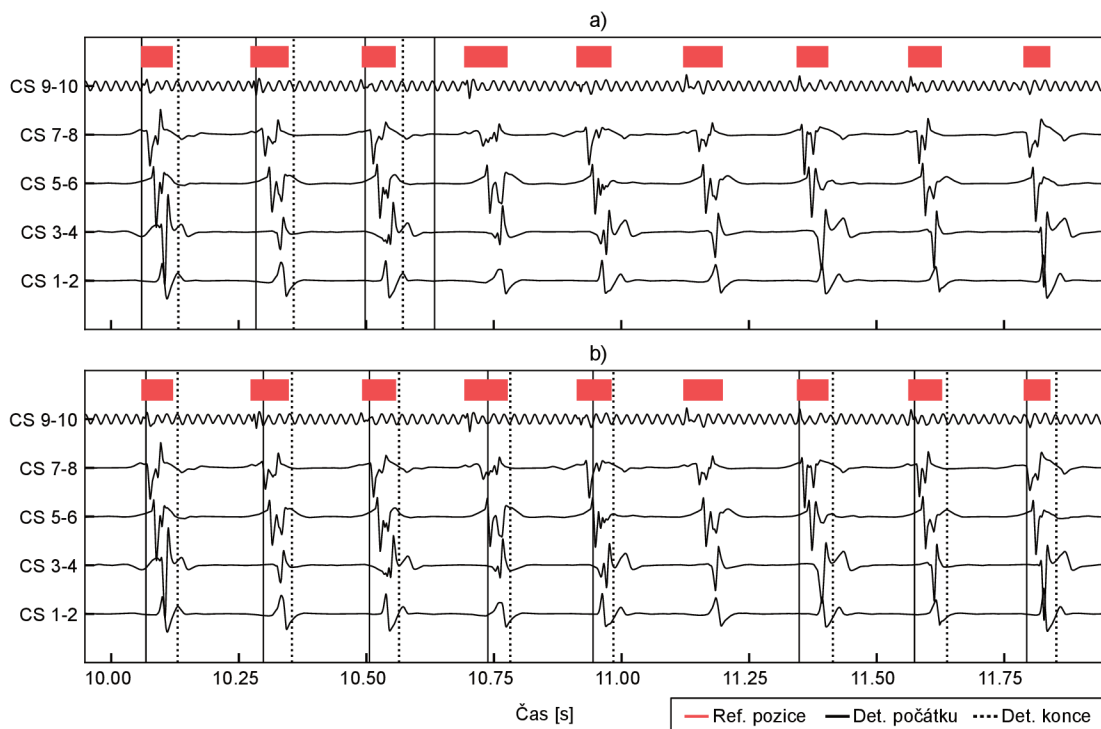
2.12 Ukázky segmentace

V této sekci jsou prezentovány ukázky segmentace na signálech z DS- A_V . Selekcce byla provedena na základě analýzy metrik kvality pro individuální záznamy. Zařazeny byly primárně ty signály, kde docházelo k systematickému selhávání.

Na Obrázku 2.29 je zobrazen EGM signál se sinusovým rytmem a přítomností hroznatého komorového komplexu s vysokou amplitudou. Zhruba v každém třetím stahu dochází k chybné interpretaci komorového komplexu jako síňový stah. Jedná se o signál s nejnižším dosaženým $F_1 = 0.61$ z celé validační sady.



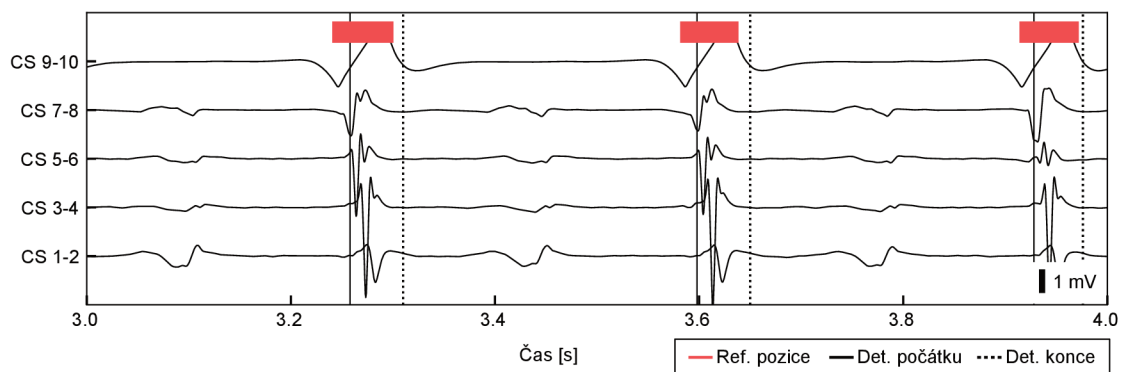
Obr. 2.29: Ukázka automatické segmentace modelem M_{lopt} s intermitentní záměnou komorového signálu za síňový.



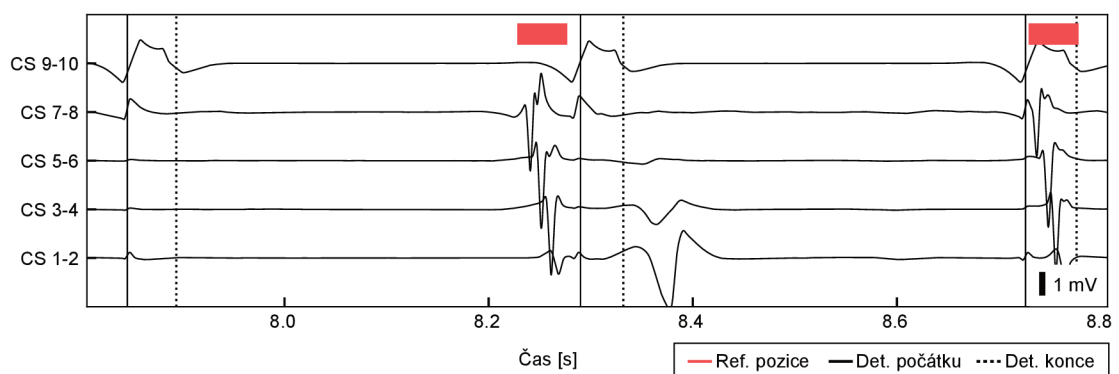
Obr. 2.30: Ukázka automatické segmentace modelem M_{lopt} v případě typického flutteru síní; a) M_{lopt} trénovaný s anotací AF; b) M_{lopt} trénovaný bez anotované AF.

Na Obrázku 2.30 je uvedeno srovnání výsledků segmentace modelem M_{lopt} . Ten byl v případě varianty (a) trénován se standardním nastavením a varianty (b) s vynecháním všech referenčních anotací fibrilace síní (AF). V obou případech se jedná o totožný úsek záznamu s typickým flutterem síní. Na průběhu EGM je vidět kolísání morfologie i časových intervalů síňových EGM, které model interpretuje v druhé části záznamu jako AF. Ukončení intervalu přerušovanou čarou je mimo rozsah ukázky. Jedná se o druhou nejnižší dosaženou hodnotu $F_1 = 0.62$. V situaci (b) model dokáže správně reagovat na individuální síňové stahy, avšak s nepřesným stanovením počátku a konce ($F_1 = 0.72$).

Na Obrázcích 2.31 a 2.32 je zachycena dvojice průběhů s programovanou stimulací aplikovanou přímo do jednoho kanálu CS 9-10. Stimulace způsobuje saturaci zesilovače a vysoká amplituda signálu i přes aplikovaný ořez způsobuje selhání modelu. V 2.31 je stimulace efektivně převáděna na myokard, čímž je časově synchronizována se síňovou aktivitou, a zdánlivě je postižena jen přesnost stanovení hranic. Vliv artefaktu je však demaskován v 2.32, kde došlo k neefektivní stimulaci a je patrné, že model síňovou aktivitu zcela minul.

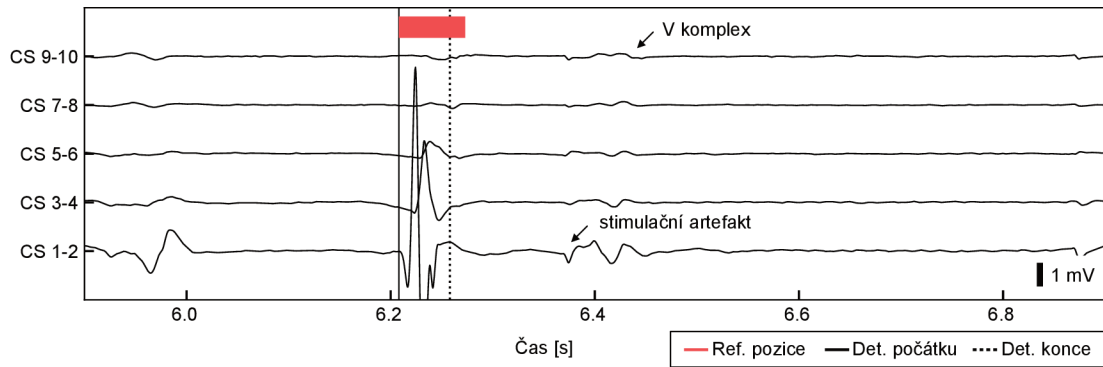


Obr. 2.31: Ukázka automatické segmentace modelem M_{lopt} v případě programované stimulace. Kanál CS 9-10 je pro lepší zobrazení oproti zbytku 5x zmenšen.

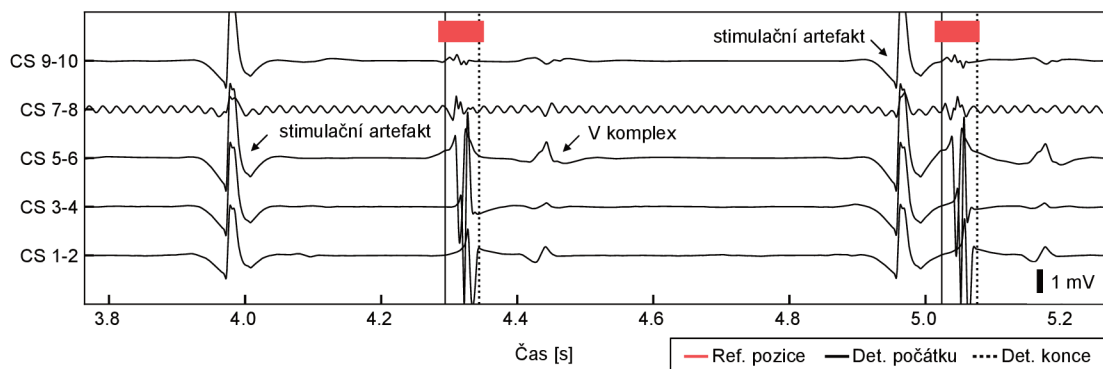


Obr. 2.32: Ukázka automatické segmentace modelem M_{lopt} při programované stimulaci.

V Obrázku 2.33 a 2.34 jsou zobrazeny další scénáře s přítomností stimulačních artefaktů. V 2.33 model správně interpretuje ($F_1 = 0.84$) vysokofrekvenční komplex jako síňovou aktivitu. Posun zadní hranice je zřejmě způsoben typickým far-field komorového EGM v CS 9-10 a 7-8, který je v této lokalizaci typický. V 2.34 je záznam se sinusovým rytmem, který byl augmentován uměle generovanými stimulačními artefakty s amplitudou převyšující síňové EGM. V tomto byly artefakty správně odlišeny od síňových komplexů s $F_1 = 0.89$.

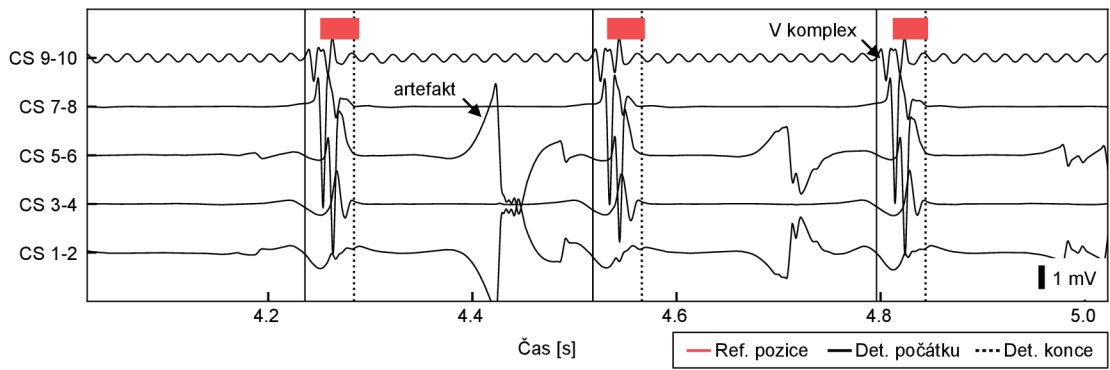


Obr. 2.33: Ukázka automatické segmentace modelem M_{lopt} v případě programované stimulace z oblasti Hisova svazku, která není převedena na síň.

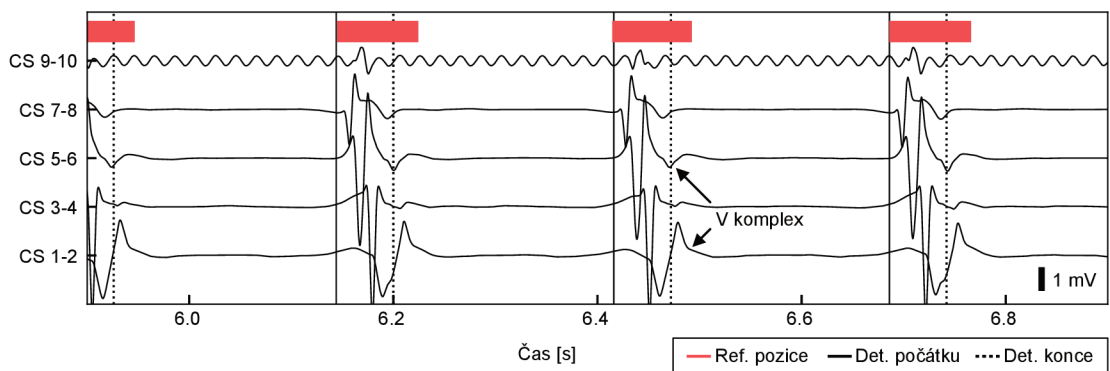


Obr. 2.34: Ukázka automatické segmentace modelem M_{lopt} v případě augmentace signálu uměle generovaným stimulačním artefaktem.

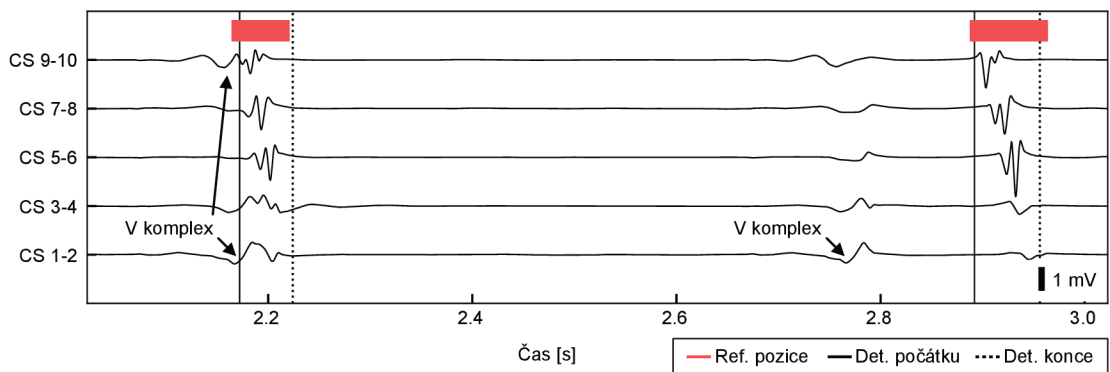
Dále jsou uvedeny ukázky chování modelu v přítomnosti fúzovaných síňových a komorových stahů. Nejslabší výsledek $F_1 = 0.75$ byl dosažen v ukázce 2.35, kde došlo vlivem velmi podobné morfologie síňového i komorového EGM ke klasifikaci obou subjektů do stejné třídy (projevono jako předčasné stanovení počátku). Na Obrázcích 2.36 a 2.37 je dosaženo vynikajícího rozlišení síňové složky od nízkofrekvenčního komorového EGM s $F_1 = 0.85$, resp. 0.87. V případě ukázky 2.36 je odchylka stanovení konce komplexu dána chybnou referenční anotací, což bylo vizuálně verifikováno na dalších záznamech stejného pacienta.



Obr. 2.35: Ukázka segmentace modelem M_{lopt} v případě fúzovaných komplexů při AVNRT.

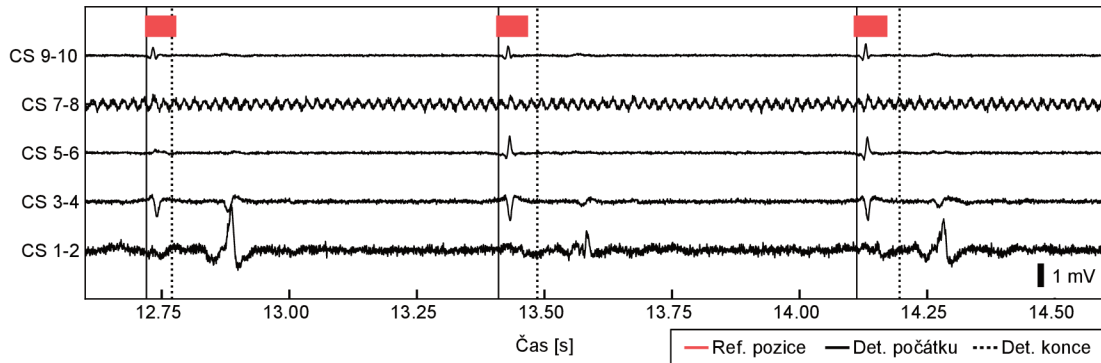


Obr. 2.36: Ukázka segmentace modelem M_{lopt} v případě fúzovaných komplexů při AVNRT. Jiný subjekt než na Obrázku 2.35.

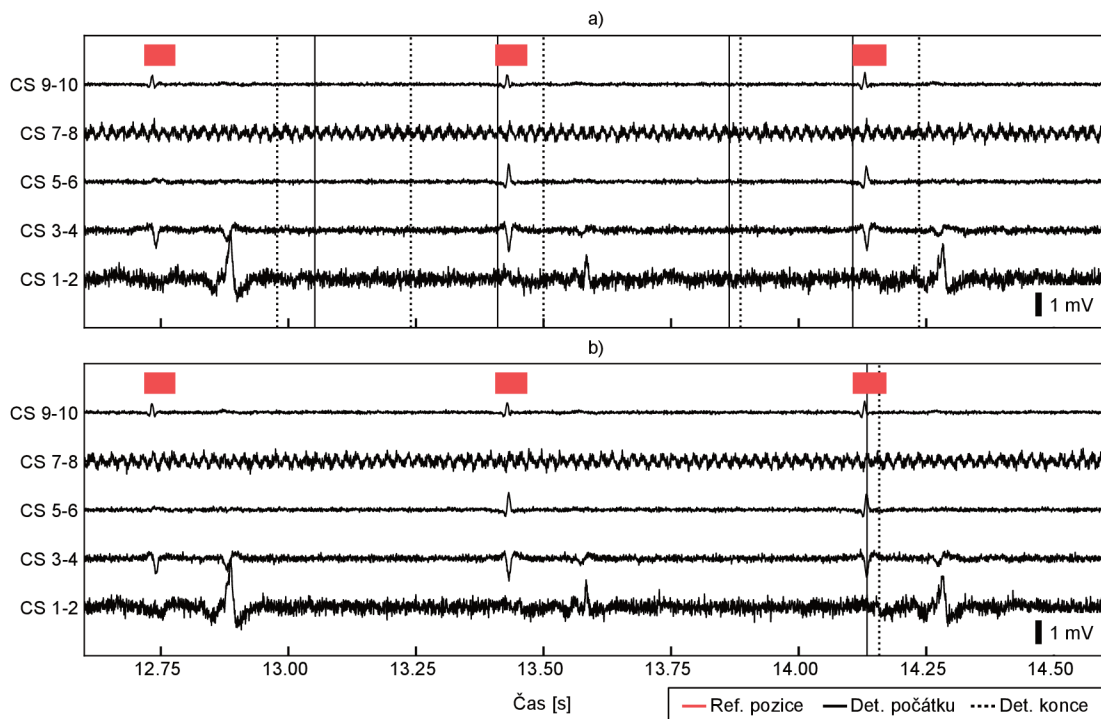


Obr. 2.37: Ukázka segmentace modelem M_{lopt} v případě komoro-síňové disociace.

Na Obrázcích 2.38 a 2.39 je znázorněn totožný segment s EGM záznamy s celkově sníženou energií signálu. Záznam obsahuje normální sinusový rytmus. Při umělém zarušení bílým šumem s výslednou úrovní $SNR = 5$ dB model správně rozlišil síňové a komorové segmenty, avšak za cenu snížené přesnosti ve stanovení zadní hranice. F_1 skóre je v tomto případě sníženo na hodnotu 0.78.

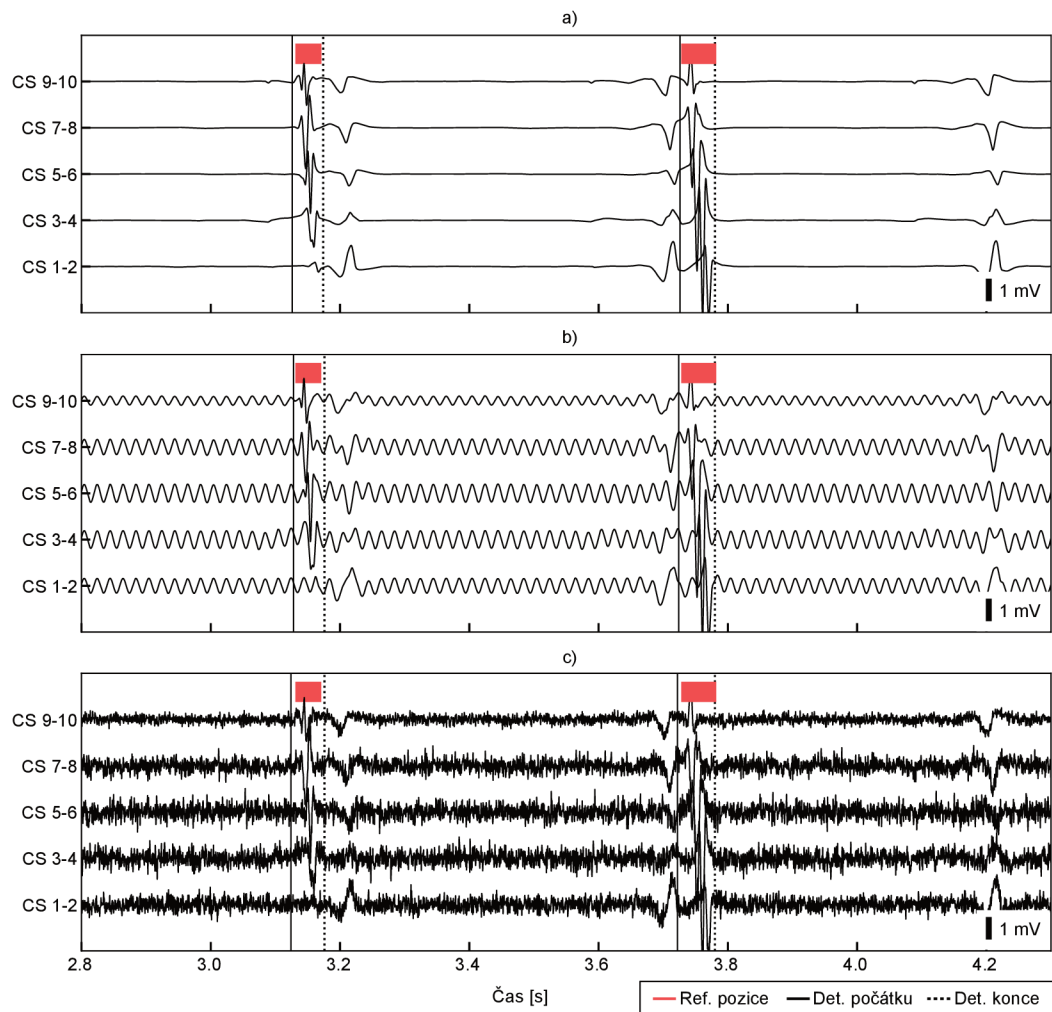


Obr. 2.38: Ukázka segmentace modelem M_{lopt} s přidaným bílým šumem, $SNR=5$ dB.



Obr. 2.39: Ukázka segmentace modelem M_{lopt} : a) reakce na přidaný bílý šum u modelu trénovaného s AF anotacemi; b) model trénovaný bez AF anotací; $SNR=0$ dB.

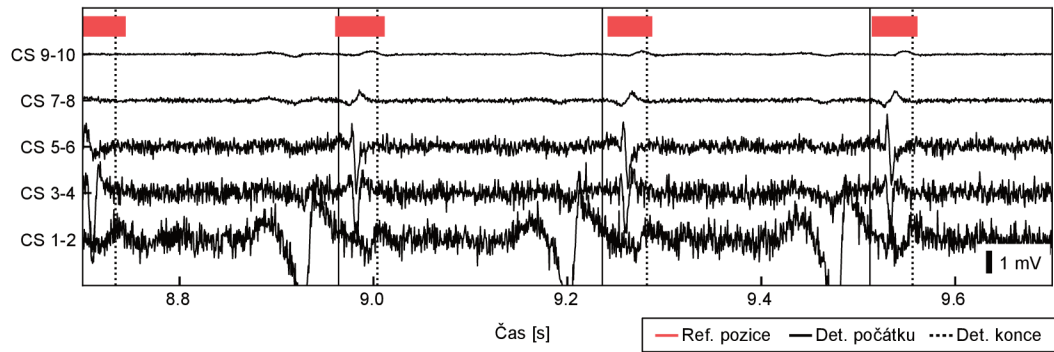
V 2.39 je hladina SNR nastavena na 0 dB. V případě (a) i (b) se jedná o výstup modelu M_{lopt} . Ve scénáři (b) však byl model trénován bez zařazení referenčních segmentů označující fibrilaci síní. V horní části (a) byl výstup generovaný modelem v podstatě náhodný. Šum je v tomto případě zřejmě zaměněn za fibrilaci síní. Celkové F_1 skóre je 0.30. V situaci, kdy model neměl při učení k dispozici anotace značící AF úseky sice také dochází k selhání identifikace síňového komplexu, šum je však správně zařazen do negativní klasifikační třídy. F_1 skóre nabývá v tomto případě hodnoty 0.51.



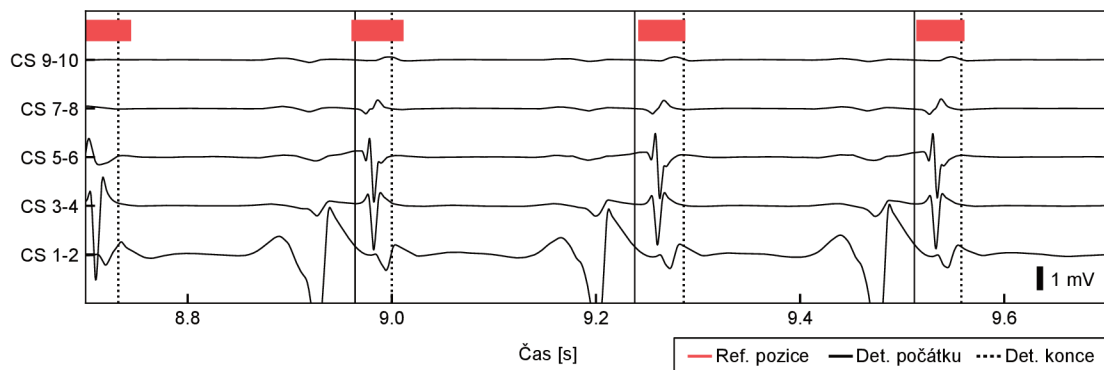
Obr. 2.40: Ukázka segmentace modelem M_{lopt} v případě a) bez přidaného šumu; b) generovaný střídavý signál o frekvenci 50 Hz, SNR = 0 dB; c) generovaný bílý šum, SNR=0 dB.

Na Obrázcích 2.40 a 2.41 demonstrujeme chování modelu v případě výskytu aditivního rušení typu síťový brum – 2.40 (b) a bílý šum – 2.40 (c) a 2.41. V obou případech se jedná o uměle generované zarušení s energií odpovídající SNR = 0 dB. Model je schopen správně detekovat síňové komplexy v každém ze zvolených scénářů. Významný je však dopad na přesnost stanovení hranic. V případě bílého šumu na Obrázku 2.40 došlo k poklesu F_1 z původní 0.90 na hodnotu 0.88. V případě ukázky 2.41 došlo po přidání bílého šumu

k propadu F_1 o 0.025 bodů oproti originálnímu záznamu 2.42, což je patrné na posunu hranic. Jedná se nicméně o záznam, který obsahuje významně menší počet kanálů s dobře rozlišitelným síňovým komplexem. Vliv může mít dále blízký komorový komplex v kanálu CS 1-2, který je důsledkem arytmie typu AVRT.

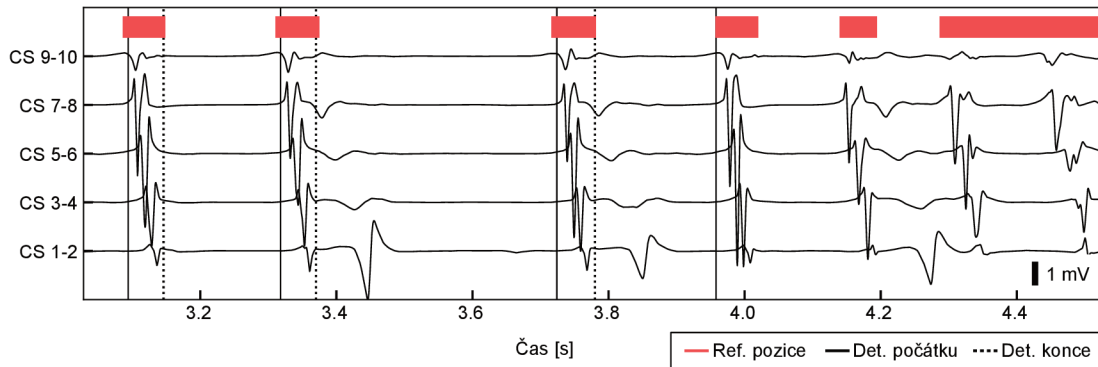


Obr. 2.41: Segmentace modelem $M_{l_{opt}}$ při výskytu generovaného bílého šumu, SNR=0 dB.



Obr. 2.42: Segmentace modelem $M_{l_{opt}}$ bez aditivního bílého šumu. Stejný úsek jako v 2.41.

Poslední ukázka na Obrázku 2.43 znázorňuje kvalitu segmentace v přítomnosti více typů síňových arytmií. Zhruba v druhé třetině záznamu dochází k přechodu do fibrilace síní. Tu by měl model zachytit jako jeden kontinuální segment s hranicí konce mimo znázorněný rozsah. To se modelu daří, avšak do tohoto intervalu je zařazen také první, relativně morfologicky homogenní stah. Je však nutné dodat, že z CS záznamu není zřejmá homogenita šíření ve zbylých částech síní, a přesné určení referenčního počátku AF může být z těchto EGM sporné. Celkově dosažené F_1 je v tomto záznamu 0.97.



Obr. 2.43: Ukázka automatické segmentace modelem M_{lopt} v přítomnosti dvou různých typů síňových arytmií.

2.13 Diskuze

V případě elektrofyziologických akvizičních systémů stále dominují anotační nástroje založené na jednoduchých filtračních technikách a manuálním měření. Důvody mohou spočívat v obecné nedůvěře s ohledem na replikovatelnost měření u pokročilých metod zpracování, přes výpočetní nároky, až po zdlouhavý proces certifikace nových technologií. Anotace lokálních aktivačních časů může přitom být prováděna v případě komplexních procedur pro řádově až tisíce individuálních elektrogramů.

2.13.1 Kvalita segmentace pomocí DNN modelů

V této části disertační práce byl realizován návrh a implementace modelů hlubokého učení pro segmentaci síňových EGM v intrakardiálních záznamech. Za výchozí model byla zvolena vlastní implementace modifikované architektury ResNet s dekódující vrstvou založenou na pyramidovém skládání příznaků. Architektura modelu i proces učení byly optimalizovány pomocí Bayesovské optimalizace. Pro výsledný model M_{lopt} byla provedena explorativní studie zaměřená na vliv vybraných parametrů na kvalitu segmentace. Srovnání modelu M_{lopt} s konvenčními technikami je vzhledem k použití interní datové sady obtížné. Autoři El Haddad et al. [69] nicméně pro měření lokálního aktivačního času definují tři úrovně přesnosti: a) vysoce přesné s chybou měření do ± 5 ms; b) přesné s chybou měření do ± 10 ms a c) nepřesné s chybou měření nad ± 10 ms. Ze zkušenosti autora odpovídá hladina ± 5 ms běžné interindividuální chybě několika hodnotitelů, která může v závislosti na způsobu zobrazení a filtraci dosáhnout i významně vyšších hodnot. V tomto kontextu dokáže model M_{lopt} ve vícesvodovém EGM určit přesnost počátku či konce aktivace s chybou < 5 ms v 54 %. V 81 % případů dokáže stanovit hranici segmentu odchylkou < 10 ms. Tyto hranice byly specifikovány pro záznamy s fokální síňovou tachykardií. V použité datové sadě DS-A se oproti tomu vyskytuje celá řada poruch srdečního rytmu, u kterých dochází k fúzi síňového a komorového komplexu, a hranice segmentů jsou bez znalosti kontextu obtížně stanovitelné i pro zkušeného operátora. Při aplikaci jedné z konvenčních

technik založené na prahování výkonové obálky na podmnožinu záznamů v DS-A bylo studentkou Martinů v práci [111] dosaženo celkové hodnoty F_1 skóre 0.763. To je o celých 0.126 bodů méně ve srovnání s hlubokou neuronovou sítí.

Ani DNN model si nedokázal poradit se všemi problematickými typy morfologií. Zatímco v případě nízkofrekvenčních komorových komplexů lze nalézt s vysokou mírou shody oba konce segmentu, v případě blízkosti spektrálních charakteristik síňových a komorových EGM může dojít při klasifikaci k jejich záměně či sloučení do jednoho segmentu. U stimulovaných síňových rytmů se opakovaně ukázal stimulační artefakt jako zdroj častého selhání. To se projevilo buď záměnou artefaktu za síňový EGM (ve dvou případech), častěji však jeho zahrnutím do výsledného segmentu. Tento efekt se nepodařilo potlačit ani augmentací s využitím uměle generovaných stimulačních artefaktů, a to zřejmě pro nedostatečnou tvarovou variabilitu. Zařazením více záznamů s programovanou stimulací následně docházelo k přeučení modelu a artefakty byly systematicky klasifikovány jako síňová aktivita.

2.13.2 Vliv fibrilačních záznamů

Specifickou roli ve vývoji modelu hrály záznamy s fibrilací síní a způsob jejich anotace. Fibrilační vlny se mohou lokálně projevovat jako náhodně organizovaná elektrická aktivita, ale také jako časově organizované komplexy, byť s variabilní amplitudou a délkou trvání. V případě DS-A byly úseky fibrilace síní bez ohledu na míru organizace signálu označeny jako kontinuální segment, což významně komplikuje optimalizační cíle. Jednou ze vstupních hypotéz pro zvýšení úspěšnosti segmentace AF signálů bylo rozšíření receptivní oblasti modelu pomocí dilatovaných konvolučních masek. Po jejich deaktivaci skutečně došlo k významnému poklesu hlavní metriky u fibrilačních záznamů, a to především díky nárůstu falešně pozitivních klasifikací. Z toho lze usoudit, že rozšíření receptivní oblasti zásadním způsobem nepřispívá k identifikaci fibrilačních vln. Zřejmě však umožňuje přesnější odlišení mezi organizovanou a neorganizovanou síňovou aktivitou v oblasti přechodu do jiného rytmu. Je třeba zmínit, že efektivní recepce širších časových souvislostí je možné v hlubokém modelu realizovat celou řadou operací, od rekurzivních neuronových jednotek [112] až po pozornostní vrstvy [113]. V obou případech se jedná o komplexní a výpočetně náročné architektonické struktury vyžadující odpovídající přístup k optimalizaci. Chování modelu za opačného extrému bylo ověřeno pomocí experimentu s trénováním a validací bez přítomnosti anotací AF rytmů. Optimalizačním cílem bylo zařadit AF úseky do negativní klasifikační třídy. V tomto případě došlo ke snížení všech metrik kvality u AF záznamů pod akceptovatelnou míru, což nutně nesouviselo s úplným selháváním modelu. Post-hoc vizuální analýza odhalila, že model i v takovém případě identifikuje izolované, částečně organizované EGM, správně jako síňovou aktivitu. Zajímavě v kontextu experimentu působilo snížení kvality segmentace u síňových stahů při sinusovém rytmu a supra-ventrikulárních arytmií. Variabilní morfologie fibrilačních EGM by mohla přispívat k lepší regularizaci modelu. Přestože binární klasifikace takových úseků představuje přílišné zjednodušení optimalizačního cíle, měla být v případě možnosti do učení modelu zahrnuta.

2.13.3 Vlivu šumu

S ohledem na chování algoritmu v situacích, kdy je do EGM signálu zaneseno uměle generované rušení, se jako kritická ukázala hranice SNR 5 dB, a to jak v případě bílého šumu, tak síťového brumu. Od této hranice docházelo k systematickému nárůstu chyby ve stanovení hranic segmentů. V případě izolovaných síťových stahů však nebyla významně snížena schopnost jejich správné identifikace. Hodnota F_1 skóre i relativní četnost identifikovaných segmentů byla v tomto případě zachována. Problematické se opět ukázaly záznamy s fibrilací síní, kde s rostoucí hladinou obou typů rušení došlo k nárůstu falešně pozitivních detekcí. Z post-hoc analýzy je zřejmé, že model v takovém případě zaměňuje rušení za fibrilační vlny. Typickou ukázkou je možné vidět na Obrázku 2.43. Hypotézu potvrzuje i segmentace provedená modelem, který byl učen bez přítomnosti označených AF úseků. V takovém případě je citlivost na rušení významně snížena ve prospěch korektní identifikace síťových komplexů. Chybné rozlišení šumu a fibrilačních vln je zajímavé i s ohledem na fakt, že signály byly v průběhu učení augmentovány totožnými technikami aditivního rušení. Hladina SNR při augmentaci byla zvolena na úrovni 5 dB, tedy shodně s kritickou hladinou odhalenou při testování. Úroveň nebyla dále snižována. Proto nelze dovodit příčinnou souvislost mezi zvoleným nastavením augmentace a výsledným vlivem na metriky kvality. Dále je třeba poznamenat, že do testování nebyly zahrnuty vyšší harmonické složky síťového brumu ani komplexnější artefakty spojené s pohybem katétrů či saturací zesilovače. Všechny zmíněné jevy se přitom v průběhu EFV mohou běžně vyskytnout.

Z dalších typů rušení byl testován vliv nulování napěťových hodnot jednotlivých kanálů, které má za úkol simulovat situaci s nižším počtem dostupných svodů. V tomto kontextu došlo již při odstranění jednoho svodu k významnému zvýšení odchylek i průměrné odchylky časové lokalizace segmentů od referenční pozice. Systematicky přitom dochází k nadhodnocení pozic, což neodpovídá situaci, kdy vlivem chybějícího kanálu dojde k celkovému zúžení nalezeného segmentu. Jedním z možných vysvětlení je schopnost modelu aproximovat průměrnou časovou distribuci vícekanálových EGM. Při chybějící informaci může dojít k detekci, která není podpořena relevantními příznaky, ale vybranou statistickou charakteristikou použitého datasetu. Konstantní počet vstupních kanálů a globální způsob anotací zřejmě budou významnou limitací navržené metody. Eliminovat tento efekt je samozřejmě možné manuálním rozměření všech lokálních EGM. Taková činnost je však časově náročná a nesouvisí s cíli disertační práce. V rámci optimalizačního procesu by mohly být přínosem transformační techniky založené na vzájemném posunu jednotlivých svodů či regrese pozic některého z extrémů lokálních EKG. V obou případech může být problematické věrohodně stanovit vztah ke globální referenční anotaci. Proměnlivý počet vstupních kanálů by bylo možné řešit adekvátní úpravou architektury ve vstupní části modelu pomocí agregačních operátorů.

2.13.4 Vliv nastavení penalizační funkce

Dílejší část optimalizačních experimentů byla zaměřena na hledání optimální kritériální funkce, a to zejména s ohledem na relativně řídký výskyt a časové omezení doby trvání síňových komplexů. Pro podobný účel byla autory Nabila et al. [93] odvozena kritériální funkce vycházející z exponenciálně váhovaného Tverskyho indexu (FTL). Oproti výsledkům v [93] nebylo v této práci dosaženo významného zlepšení v F_1 skóre ve srovnání s konvenční funkcí odvozenou z Sørensen–Dice koeficientu (SDC) nebo neváhované verze Tverskyho indexu. Přínos deklarovaný autory bude nejspíše odvozen od vlastností datové sady a komplexnosti řešeného problému. V našem případě směřoval trend nastavení exponentu podobně jako ve zmíněné práci směrem k potlačení vlivu snadno rozpoznatelných vzorků signálu. Z pohledu metrik kvality vedla tato optimalizace k vyrovnání senzitivity a pozitivní prediktivní hodnoty. Praktický důsledek spočíval primárně v přesnosti stanovení oblasti přechodu u fibrilačních záznamů a intervalů u stimulovaných síňových stahů. Z klinického hlediska může jít o užitečný nástroj pro kalibraci hlubokých modelů s ohledem na specifické klinické cíle. Skupina testovaných kritériálních funkcí totiž minimálně v našich podmínkách prokázala vysokou separační schopnost pravděpodobnostní distribuce příznaků negativní a pozitivní klasifikační třídy. Konvenční kalibrace prahu pro normalizovanou skóre vrácená modelem neměla na poměr mezi senzitivitou a pozitivní prediktivní hodnotou v podstatě žádný vliv. Kromě FTL byla navíc testována dvojice kritériálních funkcí, a to váhovaná verze křížové entropie a spojitá aproximace Matthewsova korelačního koeficientu (MCC). V případě MCC nedošlo k významnému zlepšení sledované metriky (validační $F_1 = 0.873$). U křížové entropie nebylo možné dosáhnout efektivního učebního procesu v případě standardního váhování pozitivní a negativní třídy. Až při kombinaci váhování a odstranění vlivu negativních vzorků ve více než 50 % délky signálu pomocí nulování gradientů bylo možné docílit stabilního učení (validační $F_1 = 0.805$). Kritériální funkce odvozené od SDC lze tudíž v případě segmentace řídce se vyskytujících objektů pokládat za preferovanou variantu.

2.13.5 Vliv stratifikace datové sady

Empiricky byl zjištěn významný rozptyl v metrikách kvality při konvenčním náhodném rozdělení na trénovací a validační sadu. Podíl na této asymetrii měly primárně atypické a obtížně segmentovatelné rytmy, z nichž některé jsou v DS-A relativně málo zastoupeny. Po zavedení stratifikované randomizace EGM došlo k redukci směrodatné odchylky rozdílů v F_1 skóre pozorovaných mezi validační a trénovací sadou téměř o 40 %. Z toho je patrné, že klinická stratifikace u malých souborů hraje významnou roli pro věrohodnost použitých metrik. Podobnou implikaci lze vyvodit i ze závěru empirické studie od autorů Althnian et al.[114]. Autoři argumentují, že robustnost a generalizační schopnost klasifikátoru je více než celkovým rozsahem trénovacích dat určena tím, do jaké míry reprezentuje skutečné rozdělení příznaků v populaci. Při vzorkování populace by tudíž měla být dodržena podobná pravděpodobnostní charakteristika obou datových sad. V opačném případě nemusí rozdíl v metrice kvality znamenat sníženou obecnou generalizační schopnost modelu, jako

spíš selhání u vzorků, které jsou v jedné z částí nadreprezentovány. Navíc je třeba doplnit, že kromě samotné věrohodnosti výsledků, může mít zvýšený rozptyl metriky negativní vliv na kontrolní mechanismy ovlivňující proces učení. Tím může být nepřímo ovlivněna kvalita modelu i reprodukovatelnost použité metodologie při nutnosti manuálních zásahů. Zde implementovaná metoda je založena na jednoduchém náhodném prohledávání. Metoda nezaručuje nalezení nejlepšího možného řešení. V relativně malém počtu iterací ale bylo možné vygenerovat dvojici datových sad s přibližně stejnou distribucí diagnostických tříd. V případě výskytu více neexkluzivních tříd (multi-label) se přitom nejedná o triviální problém. Jedna z mála dostupných metod je založena na hladovém algoritmu [115]. Díky chybějící randomizaci ovšem její použití zvyšuje riziko zanesení výběrové chyby.

2.13.6 Limitace

Práce se týkají některé limitace, které je nutné zmínit. Limitace na straně datové sady již byly uvedeny v diskuzi Kapitoly 1. Mezi další omezení patří především konstantní počet svodů, které je model schopen analyzovat. Z výsledků je patrné, že pro správnou funkci je v ideálním případě zapotřebí všech 5 svodů z katétru v koronárním sinu. Přestože je toto zapojení v praxi běžně užívané, nelze vyloučit použití katétru s menším či větším počtem svodů. Možnosti řešení již byly nastíněny v předchozích odstavcích. S použitými kanály souvisí také specifická lokalizace, ze které byly EGM snímány. I když byla demonstrována vysoká morfologická variabilita síňových i komorových komplexů v této oblasti, nebyl model ověřen na signálech z jiných anatomických struktur. To se týká zejména rozmezí síní a komor, kde jsou zaznamenávány EGM obou dutin. Model dále nebyl testován na signálech snímaných ze strukturálně postiženého myokardu, unipolárních signálech, signálech snímaných odlišným hardwarem. Kvalita segmentace nebyla kvantitativně ověřena na zcela nezávislé množině testovacích dat. Vizualní analýza DS-B, která vznikla právě pomocí automatického hodnocení zde implementovaným algoritmem, však neodhalila jiná závažná selhání než ta, která jsou reportována v rámci výsledků.

3 Slabě dozorovaná segmentace P vln

3.1 Současný stav

Elektrická depolarizace srdečních síní je v povrchovém EKG signálu reprezentována vlnou P, která má za fyziologických podmínek v rámci P-QRS-T sekvence pevnou časovou pozici. Parametry P vlny jako jsou amplituda, šířka, morfologie a uspořádání v čase jsou významnými klinickými markery pro diagnostiku a stratifikaci rizika u celé řady srdečních onemocnění [116]. Abnormální rozšíření P vlny, indikující snížení rychlosti šíření depolarizační vlny, je považováno za nezávislý prediktor vzniku paroxysmální fibrilace síní [117, 118, 119] podobně jako četný výskyt síňových extrasystol a paroxysmů síňových tachykardií [120, 118]. Časový vztah mezi P vlnou a QRS komplexem hraje významnou roli v diferenciální diagnostice atrioventrikulárních (AV) bloků, supraventrikulárních tachykardií (SVT) a tachykardií s širokým QRS komplexem. Obraz poslední zmíněné arytmie může přitom vznikat v důsledku přítomnosti SVT s aberovaným vnitřní komorovým vedením i komorové tachykardie. Správné stanovení diagnózy je přitom stěžejní pro určení vhodné terapie a celkové prognózy pacienta. Zatímco u komorové tachykardie je typickým znakem atrioventrikulární disociace, tj. frekvence síňové aktivity je nezávislá či nekonzistentní s frekvencí aktivity komorové [121, 122], u supraventrikulární tachykardie s širokým QRS komplexem je obvyklý různý stupeň spřažení P vlny a QRS komplexu. Variabilita ve frekvenci P vlny u síňové tachykardie je proto vodítkem pro odlišení fokální síňové tachykardie a isthmus-dependetního flutteru síní [123]. Naopak úplná absence P vlny v EKG záznamu je typická pro běžící fibrilaci síní. Odlišení paroxysmů fibrilace síní od salv síňových extrasystol a dalších poruch rytmu je tedy jedním z předpokladů pro kvantitativní hodnocení časové zátěže fibrilace síní. Ta představuje dosud málo probádaný klinický parametr spojený se zvýšenou incidencí cévní mozkové příhody, srdečního selhání a celkovou mortalitou [124].

Automatická detekce síňové aktivity umožňuje zvýšit diagnostickou výtěžnost EKG záznamů a snížit ekonomickou zátěž spojenou s klinickou interpretací, především pak u dlouhodobých EKG měření. U modelů hlubokého učení se může s ohledem na jejich komplexnost a výpočetní nároky jevit detekce síňové aktivity jako příliš úzké vymezení problému. Zabudování dílčích architektonických prvků, které by umožnily modelům extrakci příznaků korelujících s výskytem či typem síňové aktivity, může přispět k lepšímu pochopení vazeb a interpretaci vnitřních stavů hlubokých neuronových sítí.

Současné metody detekce síňové aktivity ze standardního EKG signálu nedosahují takové přesnosti jako algoritmy zaměřené na elektrickou aktivitu komor. Přesnost detekce a segmentace se navíc významně snižuje s výskytem patofyziologických rytmů v EKG signálu. Obtížnost úlohy a jejího řešení je pak dána primárně následujícími důvody [125, 126]: a) nízká úroveň energie síňového signálu a s tím spojená nízká úroveň poměru signál-šum (SNR); b) hlavní složky frekvenční charakteristiky se nacházejí v oblasti 5–30 Hz [127] a významně se překrývají se spektrem QRS komplexu a T vlny; c) v případě výskytu poruch srdečního rytmu obvykle dochází k abnormálnímu časovému uspořádání P vlny s následným překrytím tvarově výraznějším QRS komplexem nebo T vlnou; d) vysoká interindividuální variabilita morfologie P vlny či nepravidelná polymorfní elektrická aktivita u komplexních arytmí a rozsáhlých strukturálních změn síňového myokardu a e) nedostupnost dat s věrohodně stanovenými referenčními pozicemi síňové aktivity pro potřeby strojového učení.

Metody pro identifikaci síňové aktivity v EKG záznamu lze rozdělit do tří základních kategorií. První skupina je založena na filtračních technikách a prohledávání lokální oblasti před QRS komplexem. Druhá skupina využívá pro nalezení P vlny subtraktivní a statistické metody pro potlačení komorové aktivity. Do poslední kategorie spadají metody využívající hluboké učení. Jednotlivé oblasti jsou v současné literatuře pokryty množstvím prací, jejichž kompletní výčet spadá mimo rozsah disertační práce. Následující text je proto zaměřen primárně na popis nejvíce významných publikací v této oblasti.

3.1.1 Metody založené na prohledávání lokální oblasti

První skupina metod je nejrozšířenější a zahrnuje přístupy, které aproximují síňovou aktivitu pomocí jedné či více bázových funkcí, obvykle odvozených z Gausiánu. Síňová aktivita je zvládněna na základě korelace P vlny a zvolené bázové funkce. Prohledávání je následně lokalizováno do oblasti mezi koncem T vlny a QRS komplexem.

Mezi typické reprezentanty patří metody založené na vlnkové transformaci (WT¹) [19, 128, 129], která pomocí banky filtrů rozkládá signál do časově-frekvenční oblasti s flexibilním časově-frekvenčním rozlišením. Charakteristické rysy P vlny jako jsou extrémy a inflexní body jsou tak transformovány na průchody signálu nulovou hladinou, což umožňuje jejich detekci. Obvyklou volbou tvaru bázové funkce je kvadratický spline [19, 128] nebo derivace Gausiánu různých řádů [19]. Popularita metody vedla k řadě prací zaměřených na optimalizaci tvaru a parametrů modulujících časově-frekvenční vlastnosti bázových funkcí [130, 52], pravidla pro prahování koeficientů WT [52], či výpočetní náročnosti pro využití v embedded systémech [131].

McSharry et al. [48] navrhl nelineární parametrický model založený na numerickém řešení diferenciálních rovnic, které umožňují genezi syntetických EKG záznamů. Autoři v práci [23] a [132] následně rozšířili model o další parametry a optimalizační algoritmus tak, aby byla umožněna regrese modelu a reálného EKG signálu. Každá P vlna je aproximována pomocí dvojice Gaussových funkcí s modulovanou amplitudou, šířkou a časovou

¹Z anglického Wavelet Transform

lokalizací. Vzhledem k nelinearitě modelu je odhad parametrů modelu proveden pomocí rozšířeného Kalmanova filtru. Identifikované parametry následně korelují s pozicí, šířkou a amplitudou mono- či bifázických P vln.

Z dalších prací lze zmínit metody využívající neparametrické a nelineární transformace signálu do prostoru latentních proměnných, které potenciálně korelují s tvarem P vlny jako jsou např. metoda empirického rozkladu [25], fázorová transformace [133, 134, 135] či optimalizace pomocí evolučních algoritmů [24].

3.1.2 Metody založené na potlačení komorové aktivity

Naivní přístup pro potlačení komorové aktivity zahrnuje přímou subtrakci QRS-T komplexu z původního signálu [136, 137, 138]. Požadavkem je podobně jako u prohledávání lokální oblasti předchozí znalost pozice R vln, potažmo QRS-T úseků. Referenční pozice slouží pro přesnou lokalizaci subtrahované oblasti a také jako zdroj dat pro vytvoření reprezentativní QRS-T sekvence, tzv. templátu. Pro snížení vlivu variability QRS-T komplexu vlivem dýchání a změn v tepové frekvenci bývá templát tvarově adaptován na lokální signál pomocí přízpůsobeného filtru nebo dynamického borcení osy.

Odlišný přístup, navržený autory v [139] a [140], využívá slepé separace zdrojů pro extrakci síňové aktivity z 12-svodových EKG záznamů. Princip metody je založen na předpokladu, že měřený signál je lineární směsí šumu a dvou statisticky nezávislých zdrojů s jiným než Gaussovským rozložením pravděpodobnosti vzorků. Dvojice nezávislých zdrojů je v tomto případě tvořena síňovou a komorovou aktivitou a jejich separace je provedena pomocí lineárního rozkladu na nezávislé komponenty.

3.1.3 Metody využívající hluboké učení

Oproti konvenčním technikám umožňují hluboké neuronální modely automatickou extrakci latentních proměnných, které potenciálně reprezentují síňovou aktivitu, s minimální nutností předchozího zpracování. Diskriminativní schopnost hlubokých modelů u úloh spojených s interpretací EKG záznamů se blíží či převyšuje klinické experty [141], což odpovídá oblibě a trendu publikovaných prací v této oblasti. Ten shrnují autoři Hong et al. v rozsáhlé rešerši [142]. Jednotlivé přístupy se zpravidla liší pouze v architektuře použitého modelu, spíše výjimečně se vývoj hlouběji věnuje samotnému optimalizačnímu procesu či problematice trénovacích dat. Proto se v této části omezíme pouze na stručný výčet architektur hlubokých modelů, které byly publikovány v souvislosti se segmentací síňové aktivity. Typickými představiteli jsou práce autorů Costandy et al. [21] a Đuraj et al. [143], kteří pro segmentaci využili široce rozšířenou architekturu U-Net [86] s 2D, respektive 1D konvolučními jádery. Reziduální neuronovou síť s 1D konvolučními jádery navrhuje autoři v práci [18]. Konvoluční neuronové sítě v tomto případě spoléhají na podobnost vzorků síňového signálu v rámci úzkého receptivního pole daného šířkou konvolučního jádra. Vyšší úroveň abstrakce je dosaženo kompozicí konvolučních vrstev a vrstev pro agregaci dat v časové ose (pooling). Naproti tomu autoři v [22, 20, 144] navrhují využití rekurentní neuronové sítě typu long-short term memory (LSTM), která díky své zřetězené struktuře

a rekurzivním vazbám umožňuje zachytit vývoj signálu v delším časovém horizontu bez nutnosti časového poolingu.

3.1.4 Shrnutí

Zmíněné metody vykazují různě závažné nedostatky, které znesnadňují přesné zachycení síňové aktivity. Vlnková transformace a příbuzné metody využívající matematické modely jsou citlivé na inter- a intraindividuální variabilitu a selhávají v situaci, kdy se morfologie křivky významně odchyluje od tvaru modelované funkce. Matematické modely vyžadují předchozí inicializaci parametrů. Ty se mohou pro různé kohorty pacientů a komplexnost signálů lišit. Nevhodný způsob optimalizace tudíž může vést ke ztrátě klinicky důležité informace a vést k podučení systému.

Významnou nevýhodou takřka všech metod využívajících lokální prohledávání je předpoklad, že se síňová aktivita bude nacházet právě v oblasti před QRS komplexem. Přesto, že jde z pohledu dostupnosti klinických anotací o logický krok, diskvalifikuje tento přístup využití algoritmu pro analýzu značné části patologických signálů. To se týká především populace s poruchami srdečního rytmu, která tvoří značnou část pacientů indikovaných k měření EKG.

U metod založených na potlačení komorové aktivity je nespornou výhodou schopnost alespoň částečně extrahovat síňovou aktivitu skrytou uvnitř QRS-T komplexu. Tyto metody nalézají uplatnění primárně v analýze fibrilačních vln [136, 139, 140] a síňových ektopických stahů [137, 138]. Subtrakce QRS-T komplexu je z principu citlivá na chybnou detekci QRS, nestacionární charakter signálu a výskyt polymorfních komorových signálů. Adekvátní separace síňové aktivity pomocí analýzy nezávislých komponent je možné dosáhnout pouze při splnění statistických předpokladů metody. K významnému zkreslení může dojít vlivem prostorové variace zdroje elektrického signálu při pohybu síňového myokardu i výskytem multifokálních ložisek, jejichž počet převyšuje počet měřených kanálů [139]. Latentní proměnné navíc vyžadují výpočetně náročnou post-hoc analýzu pro určení komponent korelujících se síňovou aktivitou.

Kvalita hlubokých modelů je obecně odvozena od kvality použité datové sady. Modely jsou známy sníženou generalizační schopností na odlišné kohortě pacientů, než byla data na kterých byl model trénován [145, 146]. Přes jednoznačný pokrok v této oblasti nesou hluboké modely memento černé skříňky, jejíž rozhodovací proces je obtížné interpretovat.

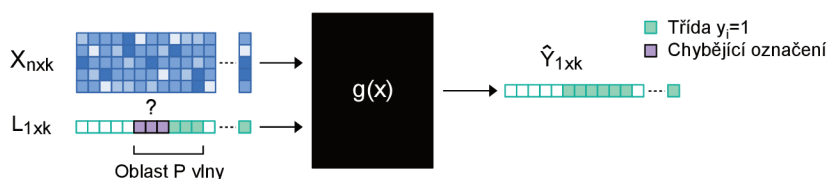
3.2 Dílčí cíle

Detekcí síňové aktivity v kontextu patofyziologických srdečních rytmů se dosud zabývaly dvě práce [135, 126]. V obou případech nebyly z již uvedených důvodů do hodnocení zařazeny skryté P vlny. Cílem této sekce je ověřit možnost zabudování vektorové reprezentace P vln do hlubokého modelu s využitím intrakardiálních EGM referencí z DS-B. Anotace nesou pouze limitovanou znalost o délce trvání elektrické aktivity síní, což převádí konvenční dozorované učení na problém učení s neúplnou informací. Na této bázi byly proto stanoveny dílčí cíle následovně:

1. Návrh a implementace metod hlubokého učení založených na víceinstančním učení pro zabudování vektorové reprezentace P vln do modelu.
2. Kvantitativní porovnání konvenční plně dozorované segmentace a segmentace s využitím víceinstančního učení.
3. Kvantitativní porovnání metod pro záznamy se sinusovým rytmem a pro celý dataset arytmiických záznamů DS-B.
4. Kvalitativní hodnocení pomocí vizuální analýzy agregovaných příznakových map.

3.3 Definice problému segmentace s neúplnou informací

Plně dozorované učení vychází z předpokladu, že každému vzorku segmentovaného objektu je v trénovací množině přiřazena referenční anotace. V reálných aplikacích je tento předpoklad málokdy naplněn a referenční hodnoty jsou buď zkruseny šumem, nejsou úplné nebo zcela chybí. V případě učení s neúplnou informací jsou referenční hodnoty poskytnuty pouze pro: a) malou podmnožinu objektů; b) malou podmnožinu vzorků reprezentujících objekt. V kontextu této práce, kde byly anotace síňové aktivity získány z intrakardiálních EGM signálů se jedná o druhý zmíněný případ. Obecně je cílem nalézt transformační funkci $g(x)$ podobně jako v případě Kapitoly 2 za předpokladu, že trénovací data sestávají z množiny $D = \{(x_{1,1}, l_{1,1}), \dots (x_{1,j}, l_{1,j}) \dots (x_{i,j}, l_{i,j}), x_{1,j+1}, \dots x_{i,m}\}$, kde i je počet trénovacích vzorků, j je počet vzorků signálu s přiřazenou anotací a $m - j$ je počet vzorků náležících síňové aktivitě, avšak bez přiřazené klasifikační třídy. Úkolem transformační funkce $g(x)$ je tedy přiřazení pozitivní klasifikační třídy všem vzorkům EKG signálu, které reprezentují síňovou aktivitu, bez ohledu na neúplnost trénovací sady. Vstupní matice dat je v našem případě tvořena obecně vícekanálovým EKG signálem z povrchového snímání. Výstupní matice je opět tvořena binární sekvencí, kde $\hat{y}_i = 1$ představuje vzorek odpovídající síňové aktivitě.



Obr. 3.1: Obecný princip učení s neúplnou informací v případě segmentace P vlny.

V oblasti metod strojového učení existují dvě základní techniky pro řešení daného problému. V prvním případě se jedná o tzv. aktivní učení. To je založeno na iterativním procesu, v němž je model trénován jen s využitím anotovaných dat. Následně je pomocí modelu provedeno značení zbylé části datové sady pomocí pseudo-referencí. Metody staví na předpokladu, že chybějící části anotací je možné doplnit na základě sémantické korelace mezi extrahovanými příznaky a konkrétní klasifikační třídou. Typicky jsou aplikovány statistické metody založené na odhadu maximální věrohodnosti nebo Lasso algoritmu pro výběr kandidátní třídy s nejvyšší posteriorní pravděpodobností [147, 148, 149, 150].

Mezi druhou skupinu technik patří tzv. víceinstanční učení (MIL²). To našlo efektivní využití v oblasti značení audio signálů [151], lokalizace objektů [152] a celé řady dalších úloh [153, 154]. Oproti konvenčnímu dozorovanému učení není v tomto případě vyžadována klasifikace individuálních instancí. Místo toho jsou instance sdruženy do množin a referenční třída je následně přiřazena celé množině. Standardním předpokladem u většiny MIL

²Z anglického Multiple-instance learning

je označení všech instancí v množině pozitivní klasifikační třídou, pokud do ní spadá alespoň jedna zařazená instance [155]. Zařazení množiny instancí do $L_i = 0$ platí proto pouze v případě že všechna $l_{i,j} = 0$. Za instanci přitom může být považován jakýkoliv příznakový vektor. Od vstupního EKG signálu až po hluboké příznakové mapy extrahované modelem. Takový způsob organizace příznaků zanáší do trénovací sady nejednoznačnost v přiřazení klasifikační třídy a modelu je tím poskytnuta flexibilita v nastavení rozhodovacích pravidel s ohledem na vzájemnou podobnost mezi instancemi [156].

3.4 Architektura modelu

Segmentaci P vlny lze zařadit do stejného okruhu problémů, které byly řešeny v Kapitole 2. Také charakteristika signálů se do značné míry překrývá. Rozdíly jsou dány primárně prostorovým uspořádáním měřicího systému a jeho fyzikálními parametry. Z tohoto důvodu se i použitá architektura a metodologie do značné míry překrývá. V následujícím textu jsou proto specifikovány jen části, ve kterých se nastavení algoritmů významně lišilo.

3.4.1 Hyperparametry

Vstupní část modelu byla tvořena 50-vrstvým reziduálním modelem (ResNet-50³) s architekturou dle návrhu v Sekci 2.4.1. Hyperparametry sítě v tomto případě nebyly z důvodu časové náročnosti učení optimalizovány. Hyperparametry proto vycházejí ze specifikace ResNet-50 a snahy o maximální kapacitu modelu dosažitelnou v rámci použitého hardwaru. Hodnoty hyperparametrů byly konstantní pro všechny provedené experimenty a jsou uvedeny v Tabulce 3.1.

Tab. 3.1: Specifikace prostoru hyperparametrů modelu pro parametry architektury

Název hyperparametru	Symbol	Hodnota
Šířka konvolučních masek	f_{RB}	7
Šířka konvoluční masky v L_{end}	f_{L0}	17
Počet vrstev v enkodéru	K_e	4
Počet RB ve vrstvách	N_{RB}	{9, 12, 18, 9}
Počet filtrů v RB	p_{RB}	{128, 256, 512, 1024}
Počet filtrů dekodéru	CH_C	256

3.4.2 Regionální příznakové mapy

Úpravy dekodérové části modelu vychází z původních prací [157] a [158]. Zde je paradigma MIL aplikováno nikoliv konvenčně na instance vstupních dat, ale na úzce vymezené regiony příznakových map generovaných dekodérem. Příznakové mapy $z(x)_{L_k}$ jsou vyvedeny z prvních tří hierarchických vrstev dekodéru. Jedná se o časově agregované příznaky, které

³Z anglického Residual Network

reprezentují vstupní signál v měřítku $\frac{1}{8}$, $\frac{1}{16}$ a $\frac{1}{32}$ ve vztahu k původní vzorkovací frekvenci vstupního EKG signálu. Jeden vzorek v příznakové mapě tudíž časově odpovídá zhruba 32, 64, resp. 128 ms. Mapy $z(x)_{L_2}$ a $z(x)_{L_3}$ jsou pro snadnější vymezení regionů se shodným počtem příznaků převzorkovány do měřítka $\frac{1}{8}$ s využitím lineární interpolace.

S ohledem na zachování informace o původní poloze vzorků je následně provedeno vnoření pevného pozičního kódování dle autorů Vaswani et al. [113]. Kódování přiřazuje každému vzorku příznakové mapy pozičně unikátní číslo. Dle autorů technika umožňuje modelu brát ohled nejen na samotnou hodnotu příznaku, ale také pozici výskytu v mapě. Metoda pozičního kódování je definována jako:

$$z^*(x)_{L_k} = z(x)_{L_k} + P, \quad (3.1)$$

$$P = \begin{cases} \sin\left(\frac{m}{5000\frac{2i}{d}}\right) & \text{pro } i = \{n \in \mathbb{Z} \mid 2n\}, \\ \cos\left(\frac{m}{5000\frac{2i}{d}}\right) & \text{pro } i = \{n \in \mathbb{Z} \mid 2n + 1\}, \end{cases} \quad (3.2)$$

kde P je d -dimenzionální matice pozičních hodnot, m je pořadí vzorku v časové ose a $i = \{1, 2, \dots, d\}$ je pořadí příznaku v matici. Každý řádek tedy odpovídá sinusoidě s variabilní frekvencí a fázovým posunem $\{0, \frac{\pi}{2}\}$.

Mapy $z(x)_{L_k}$ jsou následně rozděleny na ekvidistantní regiony o konstantní šířce. Tyto regiony se mohou vzájemně překrývat. Každý z nich reprezentuje jednu množinu instancí (vybraných vzorků příznakové mapy) a je mu přidělena anotace na základě předpokladů MIL zmíněných v předchozím textu. Šířka a míra překryvu jednotlivých segmentů zavádí do modelu nové hyperparametry, jejichž analýza by vyžadovala samostatnou studii. Z tohoto důvodu byla empiricky zvolena šířka regionu 3 vzorky s překryvem $\frac{1}{3}$. To pro sadu map $z(x)_{L_k}$ představuje dostatečně širokou receptivní oblast (přibližně 36, 192 a 384 ms) pro hodnocení příznaků vymezujiících P vlnu i její nejbližší okolí.

3.5 Normalizovaná křížová entropie pro kontrastní učení

Kontrastní učení je na rozdíl od konvenční klasifikace (kategorizace do tříd) založeno na transformaci objektu do vnořené vektorové reprezentace. Primárním cílem je vytvořit vektorovou reprezentaci s podobnou charakteristikou pro dva objekty ze stejné kategorie a naopak, maximalizovat vzdálenost mezi vektory pro dvojici nesouhlasných objektů. Základní princip metody vychází z původní práce autorů Chopra et al. [159]. Ti definovali nedozorované kontrastní učení na základě optimalizace Euklidovské vzdálenosti mezi vektorovou reprezentací dvojice objektů. Koncept byl dále rozšířen pro triplety objektů [160] a vícečetný výskyt pozitivních a negativních párů [161] v trénovacích dávkách. Autoři Chen et al. [162] dále optimalizovali proces kontrastního učení zavedením kosinové podobnosti a křížové entropie do kriteriální funkce. Khosla et al. [163] poté rozšířili kriteriální funkci z [162] pro použití v úlohách dozorovaného učení.

V rámci této části byla adoptována penalizační funkce *SimCLR* z práce autorů Khosla et al. [163]. Účelem bylo vnoření vektorové reprezentace úseků se síňovou aktivitou do jednotlivých vrstev dekódující části modelu. Princip optimalizace pomocí *SimCLR* spočívá v nalezení dvou oddělených shluků na povrchu jednotkové hyperkoule. Objekty náležící do stejné klasifikační třídy jsou v normalizovaném prostoru vzájemně přitahovány a současně dochází k maximalizaci mezishlukové vzdálenosti pro dvě odlišné kategorie. Pro shlukování jednotlivých úseků EKG byla využita kosinová podobnost, kterou můžeme obecně definovat jako:

$$\text{sim}(z_i, z_k) = \frac{z_i^T \cdot z_k}{\|z_i\| \cdot \|z_k\|}, \quad (3.3)$$

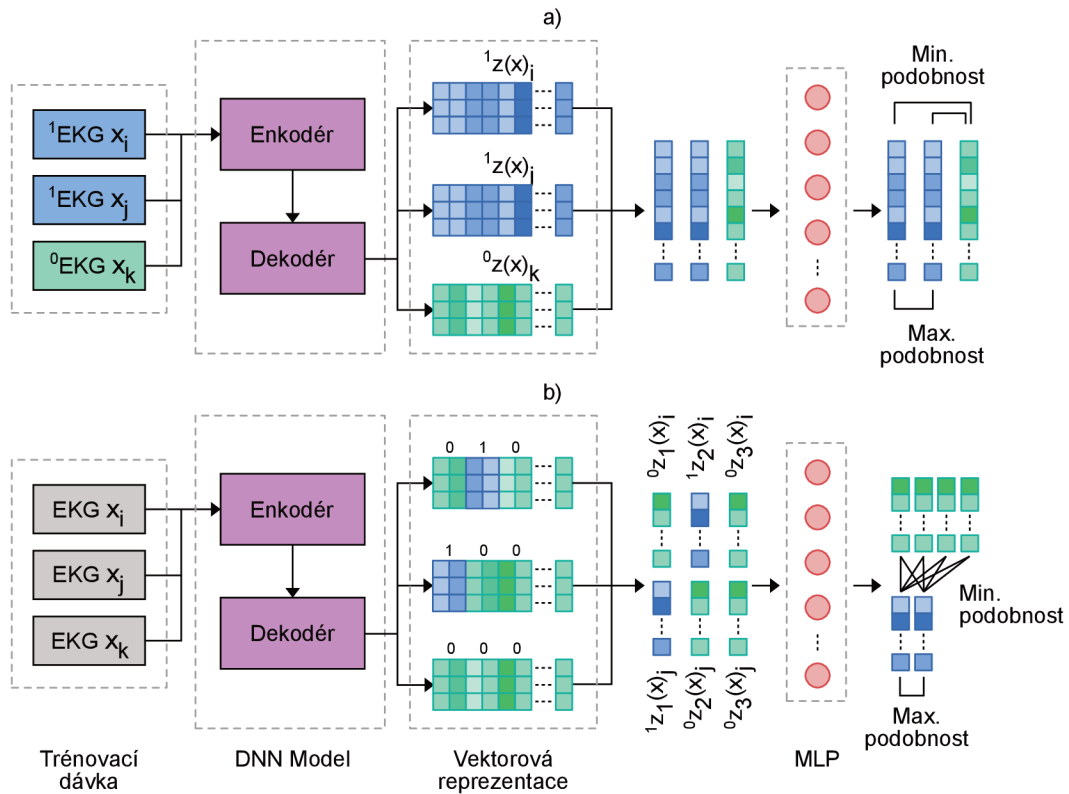
kde proměnné z_i a z_k značí dvojici příznakových vektorů a symbol $\|\cdot\|$ reprezentuje jejich L^2 vektorovou normu. Penalizace pro dvojici vektorů s přítomností P vlny je následně určena pomocí standardní křížové entropie:

$$\text{SimCLR} = -\log \frac{e^{\text{sim}(z_i, z_j)/\tau_S}}{\sum_{k=1}^{2N} \mathbb{P}[i \neq k] e^{\text{sim}(z_i, z_k)/\tau_S}}. \quad (3.4)$$

Parametr $p \in \{0, 1\}$ nabývá hodnoty 1 pokud $i \neq k$, čímž je v děliteli zanedbána soběpodobnost vektoru z_i . Koefficient τ_S slouží k váhování vlivu obtížně klasifikovatelných segmentů z negativní třídy na penalizaci. Pro vyšší hodnoty τ_S dochází k nárůstu strmosti gradientu kriteriální funkce v blízkosti pozitivních vzorků. Tím je navýšena i penalizace v případě malé vzdálenosti mezi shluky negativních a pozitivních vzorků [164].

3.5.1 Kontrastní učení v lokalizované oblasti signálu

Originální práce autorů Khosla et al. se zabývá využitím *SimCLR* pro optimalizaci vektorové reprezentace na úrovni celých trénovacích vzorů vstupujících do modelu. Trénovací vzory jsou v našem případě tvořeny individuálními EKG záznamy. Tato situace je v rámci principu *SimCLR* ilustrována na Obrázku 3.2 (a). Anotované úseky síňové aktivity jsou ovšem lokalizovány do mnoha relativně úzce vymezených oblastí signálu, a to u většiny EKG v DS-B. Z tohoto důvodu není celková meziobjektová podobnost použitelná pro segmentaci P vln. Za tímto účelem byla navržena modifikace procesu učení se *SimCLR* tak, aby ke stanovení podobnosti/nepodobnosti docházelo mezi lokalizovanými oblastmi EKG signálu. Základní princip je znázorněn na Obrázku 3.2 (b).



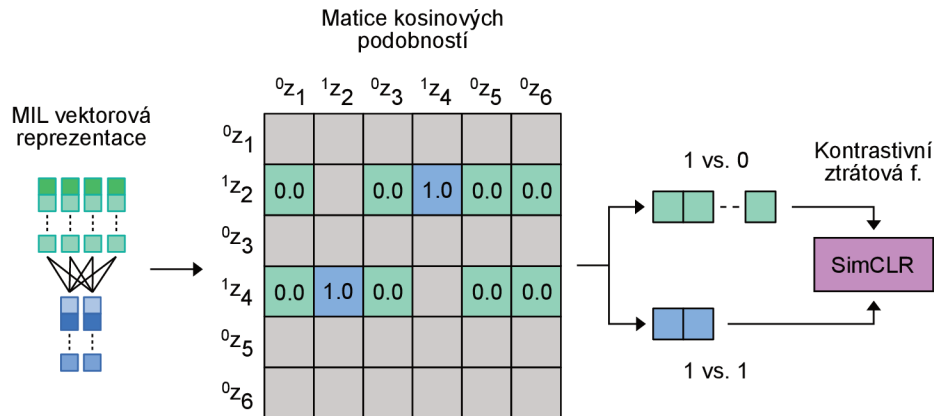
Obr. 3.2: Schématické znázornění navrženého principu penalizace podobnosti lokálních úseků příznakových map pomocí *SimCLR*: a) standardní regularizace na základě porovnání tripletu vstupních trénovacích vzorů; b) porovnání dvojic lokálních segmentů z pozitivní a negativní klasifikační třídy. MLP – vícevrstvý perceptron.

Prvním krokem je extrakce příznakových map z jednotlivých vrstev pyramidového dekodéru pro všechny signály v trénovací dávce. Tím jsou obdrženy víceúrovňové příznaky s různou mírou abstrakce a časovým rozlišením. Příznakové mapy jsou následně rozděleny na úseky o konstantní délce. Úseku je přiřazena pozitivní klasifikační třída, jestliže je na základě anotace alespoň jeden jeho vzorek označen jako síňová aktivita. Tento krok je postaven na paradigmatu MIL, viz předchozí text. Instance jsou na rozdíl od konvenčního MIL tvořeny vzorky hluboké příznakové mapy podobně jako v práci [157], a v úlohách rozpoznávání objektů s návrhem kandidátních segmentů [158]. Na rozdíl od [158] není v této práci prováděn výběr kandidátů. Vybrané úseky proto rovnoměrně pokrývají celou délku signálu. Každý úsek tvoří samostatnou vektorovou reprezentaci lokální oblasti EKG signálu, pro kterou byla provedena optimalizace, viz Obrázek 3.3

Před výpočtem penalizace v *SimCLR* byla provedena nelineární transformace lokální vektorové reprezentace pomocí série plně propojených vrstev (MLP⁴) a ReLU aktivační funkce. Následně je stanovena míra shody úseků s využitím kosinové podobnosti. Penalizace je vypočtena pouze pro podobnost mezi dvojicí pozitivních tříd (1 vs. 1) a mezi

⁴Z anglického Multi-layer Perceptron

pozitivní a negativní třídou (1 vs. 0). Tím je rozvolněn požadavek na vztah mezi vnořenou reprezentací zbylých EKG komplexů, izoelektrické linie a šumu.



Obr. 3.3: Princip výpočtu penalizace pomocí *SimCLR* a kosinové vzdálenosti určené pro dílčí úseky příznakových map EKG signálu. Modrou barvou jsou znázorněny vektory obsahující P vlnu a optimální míra jejich podobnosti. V zelené barvě jsou vektory označené jako “pozadí” signálu.

Na Obrázku 3.3 je lokální oblast s přítomností P vlny označena modrou barvou. “Pozadí” EKG signálu, tedy vektorová reprezentace ostatních úseků, je zobrazena zeleně. Podobnost úseků je určena napříč příznakovou mapou i přes jednotlivé trénovací vzory celé trénovací dávky. Výsledkem je matice podobností, kde hodnota 1.0 reprezentuje ideální shodu dvou úseků z pozitivní třídy. Hodnota 0.0 znázorňuje optimum pro odlišení úseků s P vlnou od pozadí signálu.

Minimalizaci kosinové podobnosti mezi pozitivní a negativní třídou lze zjednodušeně chápat jako dekorelaci příznaků síňové aktivity a zbývajících úseků signálu. Díky MIL paradigmatu je modelu umožněno nalézt příznaky korelované s přítomností síňové aktivity bez ohledu na přítomnost neúplné anotace. Kontrastní učení s pomocí uvedených modifikací *SimCLR* navíc umožňuje zachytit inter- i intraindividuální korelace mezi P vlnami. Při dostatečném počtu trénovacích epoch může být dosaženo párového srovnání všech dostupných P vln v datové sadě. Pomocí penalizace *SimCLR* jsou učeny individuální vrstvy dekodéru. Vnořené vektorové reprezentace jsou následně využity ve finální klasifikační vrstvě.

3.6 MIL vrstva

Vrstva MIL transformuje pravděpodobnostní rozložení příznakových map z úrovně instancí signálu na pravděpodobnost zařazení celé množiny. Klíčovou roli u MIL hrají agregační operátory. Agregaci množiny lze realizovat pomocí maxima [165], normalizované exponenciály [166], self-attention mechanismu [167] a dalších operací. V této práci byla zvolena agregace příznakových map do prostoru klasifikačních tříd pomocí operátoru maxima, který dosahuje kvalitních výsledků i ve srovnání s více komplexními protějšky [168]. Agregace je provedena na normalizované mapě a lze ji formálně vyjádřit jako:

$$z_{ij} = \max_{i=\{k,\dots,s\},j=\{0,\dots,m\}} \left(\frac{1}{1 + e^{-z_{i,j}}} \right) \quad (3.5)$$

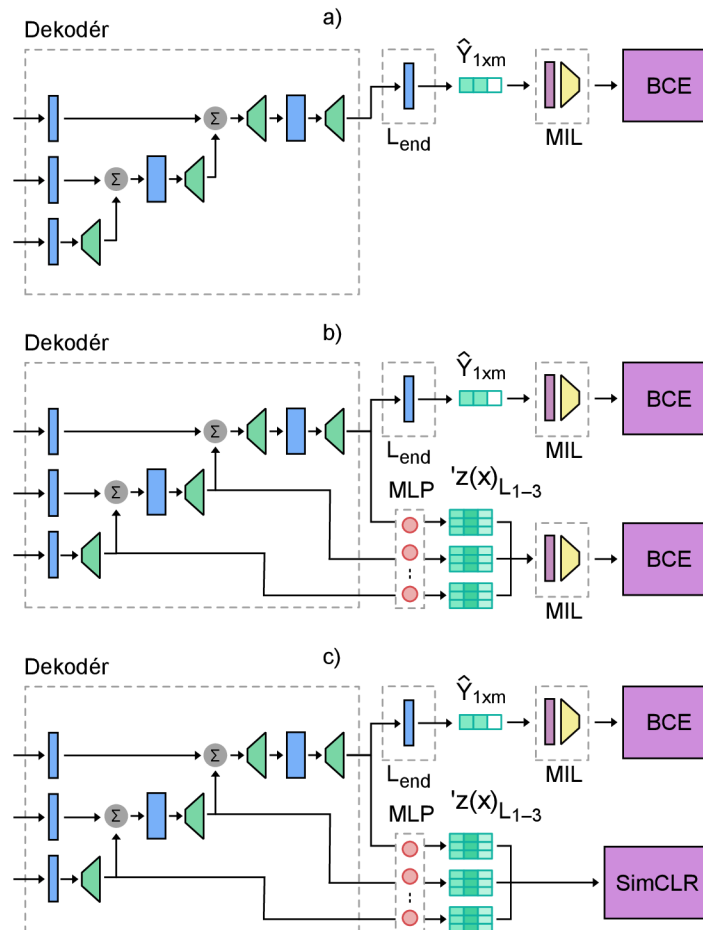
kde výraz v závorce reprezentuje normalizaci pomocí logistické sigmoidální funkce, $z_{i,j}$ je segment příznakové mapy v rozsahu vzorků $i = \{k, \dots, s\}$ a příznaků $j = \{0, \dots, m\}$, $z_{i,j}$ odpovídá agregované hodnotě pro lokálně vymezenou oblast. Výraz lze tedy interpretovat jako klouzavý nelineární filtr typu maximum s šířkou okna τ_i .

Pro výstupní vrstvu L_{end} byl rozsah τ_i stanoven na konstantní hodnotu 160 ms, aby bylo možné potenciálně pokrýt celou šířku P vlny sinusového původu. Ta je odhadována na 120 ms u zdravé populace a až 145 ms u pacientů s fibrilací síní [169]. Poruchy vedení v oblasti fibrotické tkáně a jejich vliv na prodloužení síňové aktivace v případě makroreentry arytmií byl v tomto případě zanedbán.

3.7 Popis experimentů

Ověření navržených metod bylo provedeno pro 3 různé způsoby učení, jejichž princip je schématicky znázorněn na Obrázku 3.4 na následující straně. Jemná segmentace byla u všech variant provedena za výstupní vrstvou dekodéru L_{end} , kde dochází k agregaci příznaků pomocí MIL vrstvy. Penalizace MIL příznakových map byla realizována pomocí váhované verze binární křížové entropie (BCE⁵). Nejjednodušší scénář v (a) odpovídal standardnímu scénáři s učením přes finální klasifikační vrstvu. Výsledkem byl model M_{MIL} . Ve scénáři (b) byly MIL agregace a BCE penalizace rozšířeny na regionální příznakové mapy ze samostatných vrstev dekodéru. Výstupní model byl označen jako $M_{3\text{MIL}}$. V případě (c) byly dekodérové vrstvy modelu M_{SIM} učeny s využitím kontrastní křížové entropie (*SimCLR*) aplikované na lokální oblasti příznaků.

⁵Z anglického Binary Cross Entropy



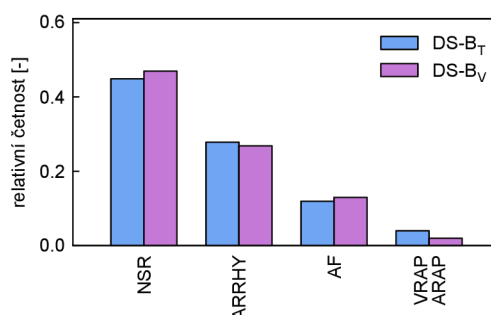
Obr. 3.4: Schémata testovaných scénářů s různým typem penalizace: a) MIL+BCE pouze pro vrstvu L_{end} (M_{MIL}); b) MIL+BCE ve vrstvách L_{end} a L_K (M_{3MIL}); c) MIL+BCE ve vrstvě L_{end} a *SimCLR* ve vrstvách L_K (M_{SIM})

3.8 Trénovací data

Pro účely této části práce je využita datová sada DS-B, která obsahuje povrchové 12-svodové EKG i intrakardiální EGM. Z dostupných záznamů bylo vybráno 8 nezávislých kanálů: a) dva bipolární povrchové svody I, II a b) unipolární hrudní svody V1–V6. Povrchové svody III, aVR, aVL a aVF nejsou měřeny přímo, ale signál je odvozen na základě lineární kombinace vybraných dvojic nezávislých svodů. Vzhledem k tomu, že redundance EKG kanálů může negativně působit na generalizační schopnost modelu [170], nebyly tyto kanály zařazeny.

Dataset byl rozdělen na trénovací $DS-B_T$ a validační $DS-B_V$ množinu v poměru 70:30 metodou popsanou v Sekci 2.7. S ohledem na dostupnost limitovaného počtu klasifikačních skupin byly pro stratifikaci využity třídy normální sinusový rytmus (NSR), fibrilace síní

(AF⁶), supraventrikulární tachykardie/arytmie (ARRHY⁷) a stimulovaný rytmus (ARAP, VRAP⁸). Procentuální zastoupení stimulovaných rytmů bylo cíleně sníženo tak, aby v datasetu nebyla obsažena příliš silná statistická vazba mezi výskytem síňové aktivity a stimulačního artefaktu. Ze zkušenosti autora a experimentů provedených v rámci disertační práce způsobuje četný výskyt stimulovaných rytmů tendence k přeučení modelu. Rozdělení pravděpodobnosti jednotlivých diagnostických tříd v DS-B_T a validační DS-B_V po optimalizaci je uvedeno na Obrázku 3.5



Obr. 3.5: Histogram relativního rozložení diagnostických tříd v DS-B_T a DS-B_V po stratifikovaném rozdělení.

3.9 Předzpracování dat

Pro běžné diagnostické využití je za optimální frekvenční pásmo EKG signálu považován rozsah 0.05–100 Hz s možným snížením až do pásma 0.5–40 Hz [55]. V prvním kroku je proto provedeno podvzorkování signálu z původních 2000 Hz na 250 Hz pomocí decimace a aplikace anti-aliasingového FIR filtru. Následně jsou aplikovány FIR filtry typu horní (T_{HP}) a dolní propust (T_{LP}) s mezní frekvencí 0.5 Hz, respektive 40 Hz.

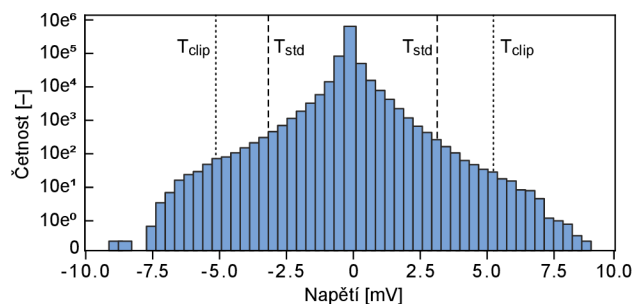
Malá část povrchových EKG záznamů obsahuje strmé pohybové artefakty s následnou saturací zesilovače. Tyto extrémy mohou mít negativní vliv na stabilní odhad gradientů v průběhu učení. Z histogramu napěťových vzorků všech dostupných EKG na Obrázku 3.6 byla proto určena prahová hodnota pro ořezání T_{clip} na úrovni 5.26 mV (99.99 percentil). Jedná se o konzervativní hladinu s dostatečnou rezervou v případě výskytu vysokonapěťových QRS komplexů, jejichž vrchol může dosahovat až několik jednotek mV. Standardizace napěťové škály (T_{std}) byla provedena do jednotkového mezipercentilového rozpětí, a to v rozsahu percentilů [0.1, 99.9]. Zvolený rozsah odpovídá napětí v oblasti [−3.15, 3.15] mV.

V průběhu trénování byly zavedeny vybrané augmentační transformace popsané v Kapi- tole 2. Jejich detailní nastavení je uvedeno v Příloze A.1.2. Záměna dvou po sobě jdoucích

⁶Z anglického Atrial Fibrillation

⁷Z anglického Arrhythmia

⁸Z anglického Atrial/Ventricular Rhythm from Artificial Pacing



Obr. 3.6: Distribuce napěťových hodnot EKG v DS-B s úrovní prahů pro standardizaci signálu; T_{clip} – hladina ořezu; T_{std} – rozsah škálování. Četnost je pro přehlednost prezentována v logaritmickém měřítku.

kanálů T_{swap} byla aplikována oproti EGM pouze na hrudní svody V1-V6. Tyto svody jsou unipolární a fyzická záměna jejich zapojení vede ke změně pořadí snímaných signálů. U bipolárních svodů dochází záměnou zapojení ke změně polarity signálu, což bylo simulováno blokem T_{rev} .

Do řetězce byl navíc přidán blok T_{time} , jehož účelem je náhodné převrácení časové posloupnosti EKG signálu. Tato transformace nemění jeho původní morfologii ani spektrální vlastnosti. Výsledkem je reverze standardního řazení komplexů P-QRS-T. Převrácení časové osy by mohlo být z klinického pohledu hodnoceno jako nelogický krok. V případě této práce však předběžné výsledky ukázaly tendence DNN modelu aproximovat časové uspořádání P vlny a QRS komplexu, a to i v případech, kdy se před QRS komplexem zjevně žádná P vlna nevyskytovala.

3.10 Nastavení procesu učení

Optimalizace modelu byla provedena pomocí algoritmu AdamW. Hyperparametry procesu učení nebyly pro vysokou časovou náročnost v této sekci stanoveny pomocí Bayesovské optimalizace, ale na základě hrubého prohledání stavového prostoru metodou *Grid-search*. Nastavení bylo konstantní pro všechny provedené experimenty. Přehled hyperparametrů, u kterých byla oproti Kapitole 2 provedena modifikace, je uveden v Tabulce 3.2

Tab. 3.2: Nastavení hyperparametrů procesu učení a kriteriálních funkcí. BS_T – počet trénovacích vzorů v dávce; η_0 – výchozí krok učení; N_{warm} – počet epoch v zahřívací fázi; w_d – faktor poklesu kroku učení.

	Proces optimalizace				SimCLR
Hyperparam.	BS_T	η_0	N_{warm}	w_d	τ_S
Hodnota	64	0.001	20	0.1	0.15

3.11 Hodnocení kvality segmentace

Metriky kvality byly zvoleny s ohledem na fakt, že v DS-B jsou k dispozici pouze slabé, respektive neúplné anotace. U prahované příznakové mapy z výstupní klasifikační vrstvy L_{end} proto byla hodnocena Se a P^+ z rovnic 2.9 a 2.10. Hodnocení bylo provedeno pro správnost stanovení počátku P_{on} , hypotetického vrcholu P_{peak} a konce P_{end} detekovaného segmentu s P vlnou. Toleranční pásmo správné detekce τ_{tol} bylo nastaveno více benevolentně, aby byl zohledněn vztah mezi reálnou šířkou P vlny a referencí. V případě expertních anotací trvá jeden síňový stah průměrně 59 ± 17 ms (viz Sekce 1.6.3), což odpovídá přibližně polovině doby trvání P vlny. Toleranční pásmo proto bylo stanoveno shodně pro všechny tři významné body na úroveň $\tau_{\text{tol}} = 60$ ms.

Metoda MIL byla shodně u všech tří variant aplikována pro učení výstupní klasifikační vrstvy L_{end} . Proto byla v tomto případě hodnocena úspěšnost zařazení celého lokálního agregovaného segmentu do klasifikační třídy, a to na základě F_1 skóre uvedeného v Rovnici 2.8. Volba F_1 opět vychází z nerovnoměrného zastoupení úseků s výskytem P vlny, které by mohlo významně vychýlit jiné, běžně užívané metriky.

3.12 Výsledky

V Tabulce 3.3 jsou uvedeny výsledky segmentace pro výstupní klasifikační vrstvu modelu. F_1 skóre reprezentuje kvalitu klasifikace po agregaci příznaků pomocí MIL. Jedná se proto o hrubý odhad schopnosti správně identifikovat P vlnu v okně o délce přibližně 160 ms. Okna měla vlivem augmentací plovoucí pozici ve vztahu k DS- A_T . V DS- A_V byly pozice oken vůči signálu konstantní.

F_1 přesáhlo pro všechny tři varianty hodnotu 0.8. Nejlepšího výsledku na validačních datech $F_1 = 0.819$ bylo dosaženo u kontrastního modelu M_{SIM} . Rozdíl mezi DS- A_T a DS- A_V se pohyboval v rozmezí zhruba 0.06–0.08 bodů. To by mohlo naznačovat mírné tendence k přeučení, bez ohledu na použitou techniku učení.

Tab. 3.3: Výsledky segmentace po MIL agregaci příznaků z klasifikační vrstvy L_{end} . DS- A_T – trénovací sada; DS- A_V – validační sada.

	DS- A_T			DS- A_V		
	M_{MIL}	$M_{3\text{MIL}}$	M_{SIM}	M_{MIL}	$M_{3\text{MIL}}$	M_{SIM}
F_1	0.864	0.892	0.893	0.801	0.811	0.819

V Tabulce 3.4 jsou prezentovány výsledky rozměření významných bodů P vlny po převedení příznakové mapy na jednotlivé instance reprezentující P vlnu. Konverze byla provedena binarizací normalizované mapy s prahem stanoveným na běžnou úroveň 0.5. P_{on} značí počátek P vlny, P_{end} její konec a P_{peak} reprezentuje hypotetický vrchol. P_{peak} byl

určen jako geometrický střed segmentu, který reprezentoval P vlnu v prahované reprezentaci. Kvůli nejasnému umístění anotace byl za úspěšnou detekci nalezení významného bodu v okolí $\tau_{\text{tol}} = 60$ ms.

Tab. 3.4: Hodnoty Se a P^+ významných bodů po převodu příznakové mapy na instance P vln. Výsledky pro celou DS-B. Tučně vyznačena nejvyšší skóre.

		DS-B _T			DS-B _V		
Model		P_{on}	P_{peak}	P_{end}	P_{on}	P_{peak}	P_{end}
M_{MIL}	Se	0.66	0.67	0.68	0.59	0.60	0.59
	P^+	0.71	0.73	0.71	0.62	0.62	0.60
$M_{3\text{MIL}}$	Se	0.69	0.71	0.71	0.63	0.63	0.64
	P^+	0.79	0.81	0.78	0.69	0.69	0.66
M_{SIM}	Se	0.78	0.79	0.78	0.71	0.72	0.69
	P^+	0.83	0.87	0.85	0.80	0.82	0.80

V Tabulce 3.4 lze modely M_{MIL} a $M_{3\text{MIL}}$ považovat za referenční metody pro učení s neúplnou informací. V případě M_{MIL} byl model optimalizován jen přes výstupní klasifikační vrstvu. U modelu $M_{3\text{MIL}}$ byly kromě klasifikační vrstvy učeny také jednotlivé větve dekodéru. Jedná se o jeden z konvenčních přístupů pro zlepšení komprimované reprezentace v pyramidových vrstvách.

Předmětem zájmu jsou primárně hodnoty Se a P^+ pro P_{peak} , které jsou z uvedených významných bodů nejméně zkresleny metodikou hodnocení. Metriky kvality u P_{on} a P_{off} by s ohledem na toleranční pásmo mohly ukázat na extrémní nadhodnocení či podhodnocení segmentů. V případě této práce jsou metriky konzistentní napříč významnými body a takový závěr tedy není možné vyvodit.

Modely M_{MIL} i $M_{3\text{MIL}}$ dosáhly v případě P_{peak} na celé DS-B_V podprůměrných výsledků se sníženými hodnotami Se i P^+ do rozsahu 0.60–0.63, respektive 0.62–0.69. Podobný pokles v kvalitě detekce významných bodů u konvenčních metod bylo, s ohledem na zastoupení rytmů, možné očekávat. Při srovnání vedla navržená metoda kontrastního učení k významnému zlepšení výsledků. Model M_{SIM} dosáhl ve srovnání s M_{MIL} a $M_{3\text{MIL}}$ zlepšení senzitivity o 12 %, respektive 9 %. K významnému nárůstu došlo u pozitivní prediktivní hodnoty, a to o celých 20 %, respektive 13 %. Kontrastní učení jednoznačně přispělo ke kvalitnější extrakci vektorové reprezentace síňové aktivity z EKG signálu a předčilo konvenčně užívané metody.

V Tabulce 3.5 uvádíme výsledky detekce významných bodů jen pro skupinu sinusových rytmů (NSR). Do podmnožiny NSR byly zařazeny méně závažné až zcela běžné poruchy srdečního rytmu jako jsou předčasné síňové stahy, komorové extrasystoly, sinusové tachykardie nebo blokády AV uzlu 1. stupně. Tento případ byl vybrán pro demonstraci chování modelů na datové sadě, která se složením nejvíce blíží standardním databázím.

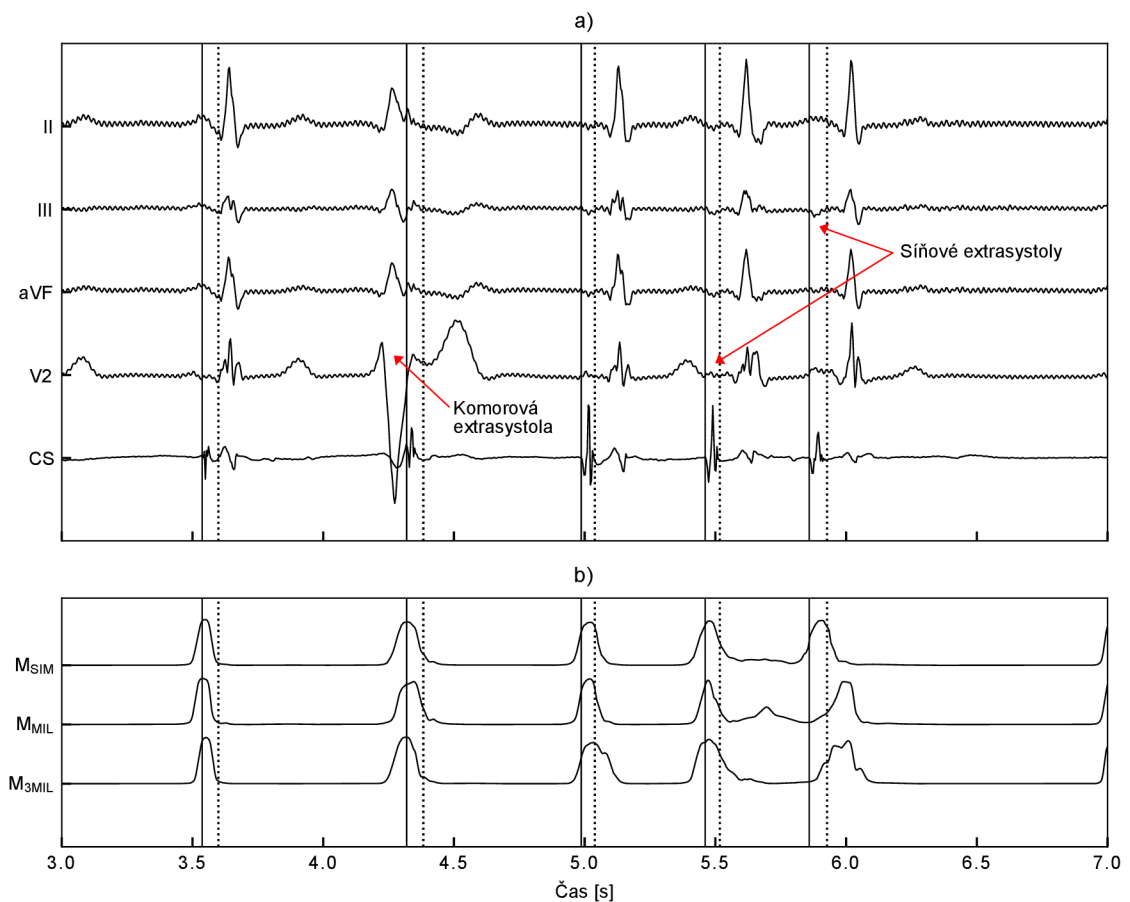
Tab. 3.5: Hodnoty Se a P^+ významných bodů po převodu příznakové mapy na instance P vln. $DS-B_V:\{NSR\}$ – podmnožina validační sady se záznamy se sinusovým rytmem. Tučně vyznačena jsou nejvyšší skóre.

		$DS-B_V:\{NSR\}$		
Model		P_{on}	P_{peak}	P_{end}
M_{MIL}	Se	0.83	0.83	0.83
	P^+	0.84	0.84	0.84
M_{3MIL}	Se	0.85	0.84	0.85
	P^+	0.84	0.85	0.84
M_{SIM}	Se	0.88	0.88	0.88
	P^+	0.82	0.86	0.86

V této skupině signálů dosáhly všechny tři modely hodnot Se a P^+ nad 80 %. To lze považovat za dobrý výsledek i v kontextu dalších prací validovaných na standardních databázích. Modely M_{MIL} i M_{3MIL} dosáhly velmi podobných hodnot Se a P^+ , a to v případě P_{peak} na úrovni 0.83 a 0.84, respektive 0.84 a 0.85. I v tomto případě ve většině parametrů dominuje kontrastně učený model M_{SIM} . Pouze v případě stanovení počátku P vlny došlo ke snížení P^+ na hodnotu 0.82, což je ve srovnání se zbylými dvěma modely nejhorší výsledek.

3.13 Ukázky segmentace

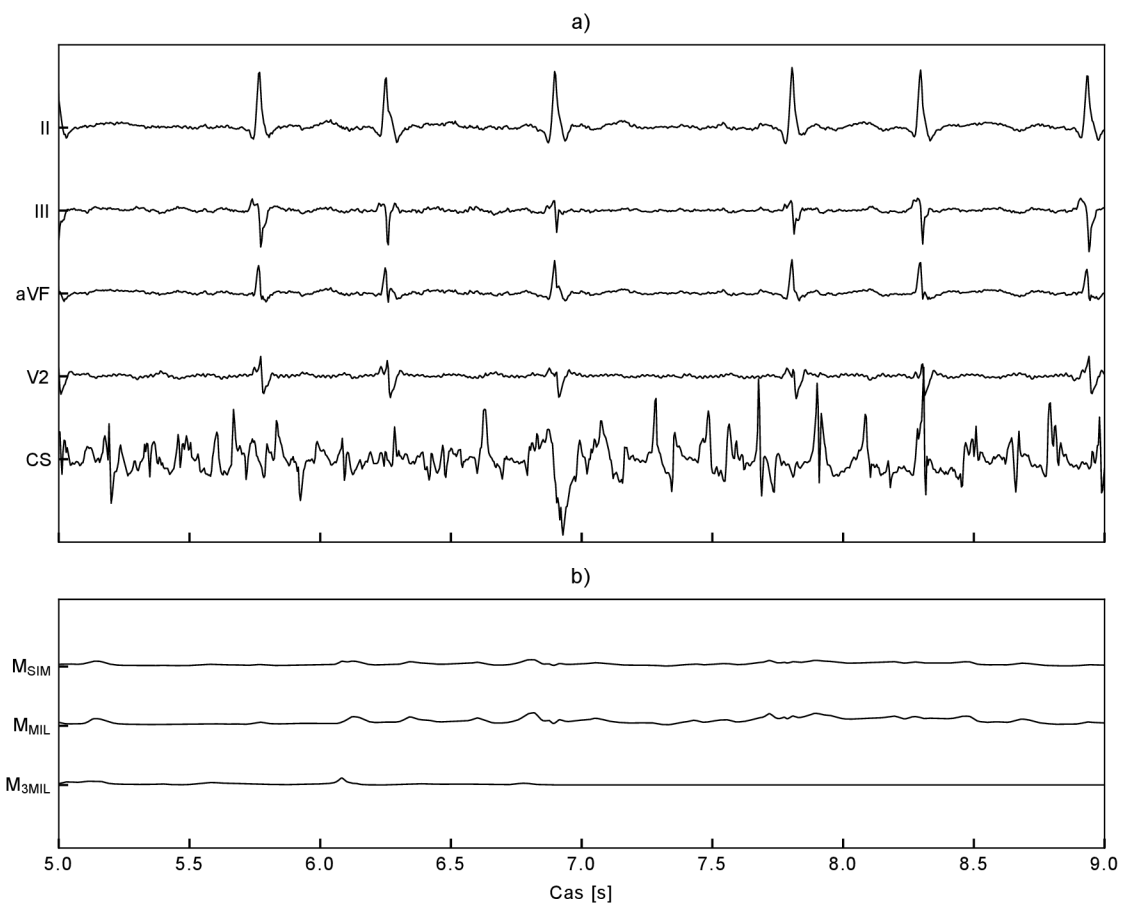
V první ukázce na Obrázku 3.7 je zobrazena segmentace P vlny v přítomnosti dvou síňových a jedné komorové extrasystoly. Za zmínku stojí především posun vrcholů v příznakových mapách v místě komorové extrasystoly směrem ke QRS komplexu. To naznačuje, že všechny tři modely pravděpodobně označily aberovaný QRS komplex chybně za P vlnu. Jedná se o jeden z obvyklých scénářů selhání. Podobnou situaci lze sledovat u M_{3MIL} a M_{MIL} v posledním, předčasném, síňovém stahu. Model M_{SIM} v tomto případě naopak správně interpretoval abnormální síňovou aktivaci před QRS komplexem.



Obr. 3.7: Ukázka segmentace EKG pomocí modelů M_{MIL} , M_{3MIL} a M_{SIM} za přítomnosti komorové a síňových extrasystol: a) výběr povrchových EKG svodů + EGM z proximálního CS; b) příznakové mapy z vrstvy L_{end} normalizované do intervalu $[0, 1]$. Svislé čáry znázorňují automatické anotace z intrakardiálních EGM.

Na Obrázku 3.8 je ukázáno chování modelů na EKG signálech s fibrilací síní. Zcela náhodná aktivita je patrná především v intrakardiálním svodu CS. Příznakové mapy vykazují jen malou míru fluktuace, která by mohla vést k falešné pozitivitě v případě nízko zvoleného prahu. Všechny tři modely přesto správně vyhodnotily, že se v signálu žádná P vlna nenachází. Nedošlo ani k záměně za vlnu T, která může vykazovat podobnou morfologickou a frekvenční charakteristiku.

Detekce fibrilace síní patří s ohledem na riziko komplikací mezi široce probíranou problematikou. Konvenčně bývá při vývoji modelů záznam s fibrilací síní zařazen do pozitivní klasifikační třídy. Z toho vyplývá i problematická interpretace pozitivně aktivovaných příznakových map u hlubokých modelů. Chybějící P vlny naproti tomu patří mezi jednoznačný a srozumitelný klinický parametr.



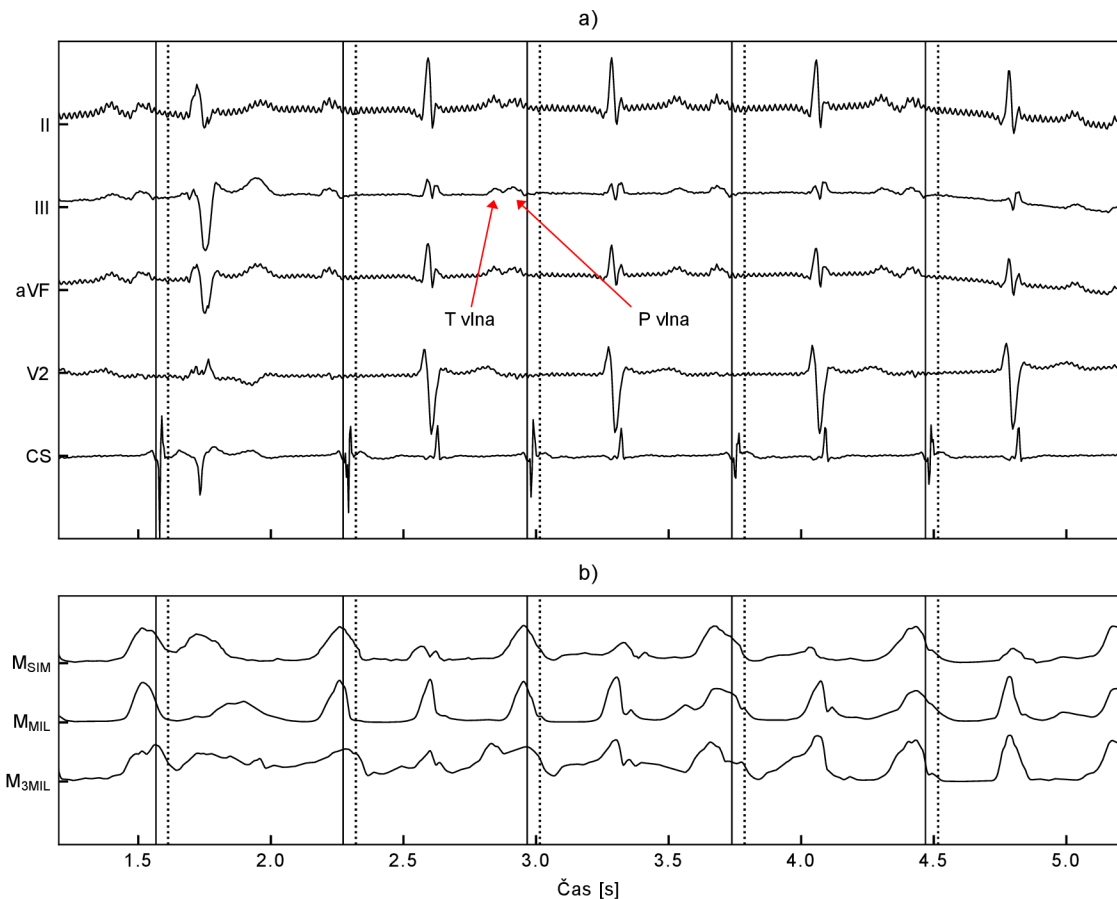
Obr. 3.8: Ukázka segmentace EKG pomocí modelů M_{MIL} , M_{3MIL} a M_{SIM} v přítomnosti fibrilace síní: a) výběr povrchových EKG svodů + EGM z proximálního CS; b) příznakové mapy z vrstvy L_{end} normalizované do intervalu $[0, 1]$. Svislé čáry znázorňují automatické anotace z intrakardiálních EGM.

Na další ukázce na Obrázku 3.9 je zobrazen EKG záznam s běžící atrio-ventrikulární reentry tachykardií (AVRT) se štíhlým QRS. Vzruch se v daném případě šíří přes AV uzal na komory a síň je poté retrográdně aktivována přes přídatné síňo-komorové vedení. Projevem na EKG jsou P vlny umístěné v náběžné hraně T vln. Ty mohou být identifikovány zkušeným kardiologem. Pro většinu současných automatických metod segmentace ovšem bude podobný scénář pravděpodobným důvodem selhání. U tohoto signálu si nejlépe vedl model M_{SIM} , který našel jednoznačně oddělené segmenty s P vlnou. Zejména v případě základního modelu M_{MIL} lze spatřit nehomogenní průběh agregované mapy s občasnou aktivací v místě QRS komplexu, který je potenciálním zdrojem falešně pozitivních detekcí.



Obr. 3.9: Ukázka segmentace EKG pomocí modelů M_{MIL} , M_{3MIL} a M_{SIM} v přítomnosti supraventrikulární tachykardie typu AVRT: a) výběr povrchových EKG svodů + EGM z proximálního CS; b) příznakové mapy z vrstvy L_{end} normalizované do intervalu $[0, 1]$. Svislé čáry znázorňují automatické anotace z intrakardiálních EGM.

Poslední ukázka na Obrázku 3.10 je věnována problematickému záznamu s bloádou AV uzlu 1. stupně a aberovaným QRS komplexem (vlevo). Každý ze tří modelů správně odhalil vlnu P těsně navazující na vlnu T, byť s nadhodnocením její šířky. V případě modelů M_{3MIL} a M_{MIL} došlo k opakované aktivaci v oblasti QRS komplexu, který byl tímto zaměněn za P vlnu. U modelu M_{SIM} by pravděpodobně tuto situaci bylo možné částečně kompenzovat volbou vhodné metody prahování. I model M_{SIM} nicméně selhává v případě prvního, aberovaného, QRS komplexu.



Obr. 3.10: Ukázka segmentace EKG pomocí modelů M_{MIL} , M_{3MIL} a M_{SIM} v přítomnosti bloády AV uzlu 1. stupně: a) výběr povrchových EKG svodů + EGM z proximálního CS; b) příznakové mapy z vrstvy L_{end} normalizované do intervalu $[0, 1]$. Svislé čáry znázorňují automatické anotace z intrakardiálních EGM.

3.14 Diskuze

Tato část dizertační práce byla věnována návrhu algoritmu pro segmentaci P vln v záznamech s četným výskytem poruch srdečního rytmu. Algoritmus byl založen na vlastní implementaci konvoluční neuronové sítě s pyramidovým skládáním příznakových map. Za účelem učení byly využity reference síňové aktivity z intrakardiálních záznamů. Pro optimalizaci modelu byla navržena modifikace kontrastní penalizační funkce založená na vzájemné korelaci lokalizovaných úseků EKG. Metodologie byla srovnána s konvenčním multi-instančním učením pro neúplné anotace na úrovni jedné a více vrstev dekodéru.

Z dosažených výsledků je patrné, že detekci P vlny v přítomnosti četných arytmii nelze považovat za triviální problém. Nejlepšího výsledku bylo dosaženo pomocí navržené modifikace kontrastního učení s hodnotou senzitivity a pozitivní prediktivní hodnoty v případě P_{peak} 72 %, respektive 82 %. Post-hoc vizuální analýza ukázala, že největší obtíže činilo nalezení P vln zcela překrytých QRS komplexem. Ty se mohou vyskytovat u několika typů supraventrikulárních tachykardií i komorových extrasystol. Tento typ záznamů měl vliv nejen na snížení Se, ale pro zvýšenou falešnou pozitivitu také na P^+ . V některých případech aberovaných a předčasných QRS komplexů měly modely tendenci umístit P vlnu doprostřed QRS. Tento efekt byl zřejmý zejména v situacích, kdy se v blízkém okolí žádná P vlna nevyskytovala. Modely ve specifických případech aproximovaly pravděpodobnostní charakteristiku P-QRS-T sekvence v trénovacím datasetu. S těmito scénáři si optimálně nedokázal poradit ani jeden z testovaných modelů. Výsledek je s ohledem na složení databáze a formulaci úlohy logický. Binární klasifikace nutí modely v případě překryté P vlny a QRS komplexu extrahovat příznaky, které jsou do značné míry korelované. Bez zásadních zásahů do metodologie učení nelze u $M_{3\text{MIL}}$ a M_{MIL} očekávat v tomto ohledu významné zlepšení. Kontrastní učení u M_{SIM} by nicméně bylo možné rozšířit na problém klasifikace do více tříd, z nichž jedna by reprezentovala komorovou depolarizaci. Penalizaci pomocí *SimCLR* aplikovanou napříč klasifikačními třídami by šlo poté připodobnit k hluboké separaci zdrojů, což by potenciálně mohlo přinést další zlepšení modelu. Tato varianta bohužel nemohla být realizována, neboť v době dokončení dizertační práce nebyly k dispozici anotace QRS komplexů.

3.14.1 Jak věrohodná je segmentace P vlny hlubokým modelem?

Závěr vizuální analýzy poněkud znevěrohodňuje výsledky hlubokých modelů pro segmentaci P vlny, které byly učeny na některé ze standardních databází [32, 33, 35, 34]. Nejčastěji [18, 19, 20, 21, 22, 23, 25] za tímto účelem bývá využívána databáze QT-DB. Publikované algoritmy hlubokého učení [20, 21] reportují na QT-DB senzitivitu a specifitu shodně nad 98 % v případě obou metrik. Práce autorů Saclova et al. [171, 134] zaměřená na srovnání několika různých metod ovšem ukázala, že jsou tyto výsledky značně nadhodnoceny ve prospěch snadno identifikovatelných P vln. Hodnota senzitivity a pozitivní prediktivní hodnoty v jejich případě kolísala v rozmezí 78.1—93.1 %, respektive 72.0—88.6 % na BUT-PDB, zatímco na databázi QT-DB v rozmezí od 99.2—99.9 % a 99.1—99.8 %. V tomto

případě se jednalo o konvenční techniky detekce s manuální extrakcí příznaků. Podobného efektu lze ovšem očekávat i u hlubokých modelů.

Výstupy této práce ukázaly, že si model může P vlnu "domyslet" tam, kde se v rámci trénovacích dat obvykle vyskytuje. V případě DS-B jsou to četné supraventrikulární tachykardie s P vlnou uvnitř QRS komplexu. Při validaci na EKG záznamech s primárně sinusovým rytmem lze stěží odhalit, jestli je detekce podložena smysluplnou extrakcí příznaků nebo pravděpodobnostním očekáváním. Tato vlastnost může být typická pro reziduální architektury, kde dochází k exponenciální agregaci dat a ztrátě časově-prostorové informace v hlubších vrstvách. Rekurentní neuronové sítě by z principu mohly být vůči tomuto jevu imunní. Podrobné srovnání se sice neprotíná s cíli této práce, nicméně by stálo za pozornost. V rámci dizertační práce se jako dostatečné řešení ukázala vhodná augmentace dat založená na inverzi časové osy signálu. Tím byla do určité míry narušena přirozená posloupnost P-QRS-T sekvence v trénovacích datech a model byl donucen hledat odlišný typ souvislostí. Principiálně by šlo tuto metodu využít i pro testování robustnosti segmentačních algoritmů založených na hlubokém učení.

3.14.2 Využití intrakardiálních referencí

Intrakardiální reference byly pro vývoj detekčního algoritmu dosud použity jen v práci Lenis et al. [52]. Autoři provedli snímání intrakardiálních EGM pomocí multi-elektrodového košíkového katétru. Z nich byly následně stanoveny referenční pozice počátku a konce elektrické aktivace síní. Signály z košíkového katétru zachycují širší část časově-prostorové distribuce napěťové vlny oproti záznamům z koronárního splavu. Jsou tedy s ohledem na segmentaci P vlny více reprezentativní než signály využívané v této práci. Práce autorů Lenis et al. ovšem využila jen 32 síňových stahů snímaných u jednoho pacienta, což je pro robustní testování algoritmů nedostatečné.

3.14.3 Limitace

S ohledem na neúplné anotace jsou hodnoty metrik kvality pouze orientační. Výsledky mohou být ovlivněny zvolenou tolerancí a mírou agregace u MIL map. Ve výsledcích je nicméně možné pozorovat jednoznačný trend, ze kterého lze usoudit kvalitu testovaných metod. V rámci kvantitativního hodnocení nebylo provedeno srovnání na některé ze standardních databází. Zde je nutno zmínit, že žádná z existujících databází s anotovanou P vlnou nedosahuje co do počtu subjektů a zastoupení různých typů rytmů takové variability jako DS-A a DS-B. Anotace částečně skrytých P vln jsou k dispozici jen v [171].

Jako jednu z hlavních limitací je potřeba uvést chybějící optimalizaci hyperparametrů modelu a procesu učení. Hyperparametry byly nastaveny konstantně pro všechny experimenty. Dílčí úpravy na straně optimalizačního algoritmu byly provedeny jen v případě, kdy nebylo možné pro testovanou variantu dosáhnout stabilního procesu učení. Některé metody by mohly, při optimální volbě hyperparametrů, teoreticky dosáhnout stejných výsledků jako nejlepší z modelů, byť např. na úkor výpočetní efektivity.

Závěr

Dizertační práce se zabývá vývojem algoritmů pro segmentaci elektrokardiografických záznamů z povrchového a intrakardiálního snímání se zaměřením na síňovou aktivitu. Hlavní cíle dizertační práce byly stanoveny s ohledem na některé nedostatky současných metod hlubokého učení v oblasti interpretace EKG. Mezi ně patří zejména snížená kvalita, případně rozmanitost, trénovacíchází a omezené možnosti při posouzení vztahu mezi hlubokými příznaky a klinickou predikcí modelu.

Návrh řešení byl založen na zabudování informace ze záznamů z intervenčního elektrofyziologického vyšetření srdce do hlubokého modelu. Data z těchto procedur poskytují zlatý standard pro diagnostiku poruchy rytmu u indikovaných pacientů. Intrakardiální signály lze označit za věrohodnou referenci pro specifické EKG komplexy, jakým je například P vlna.

Využití elektrofyziologických dat je komplikováno uzavřeným systémem pro akvizici klinických signálů. V práci proto byly navrženy softwarové nástroje pro export dat do otevřeného formátu. Díky tomu bylo možné sestavit trénovací datové sady s širokým spektrem poruch srdečního rytmu z kohort mladých dospělých a pacientů středního věku. U datasetů byla z části provedena expertní anotace autorem práce a ve spolupráci s klinickým pracovištěm FN u Sv. Anny v Brně. Z části byly záznamy anotovány navrženým automatickým algoritmem.

V Kapitole 2 byl navržen a implementován algoritmus pro automatickou segmentaci síňových elektrogramů (EGM) pro různé arytmiické scénáře. Algoritmus využívá vlastní implementaci 1D konvoluční neuronové sítě s reziduálním spojením. Pro rozšíření receptivní oblasti modelu řádově na jednotky sekund byl do modelu implementován vlastní návrh vrstvy s kompresními a dilatovanými konvolučními jádry. Tato vrstva se ukázala jako stěžejní pro korektní segmentaci v oblasti přechodu mezi fibrilací síní a jiným, více pravidelným rytmem. Pro algoritmus byla provedena série experimentů sledující vliv nastavení a předzpracování datové sady na kvalitu segmentace. Jako zásadní činitel se ukázala přítomnost fibrilace síní v datasetu a způsob její anotace. Výsledky naznačily, že přítomnost fibrilačních záznamů má pozitivní vliv na regularizaci modelu a zlepšuje přesnost segmentace i v případě izolovaných EGM. Navržená metoda prokázala schopnost odlišit síňové EGM od záznamu vzdáleného elektromagnetického pole komorového myokardu i v případě sdružených síňo-komorových komplexů. Jedná se o problematické úseky, se kterými si konvenčně užívané metody dokáží poradit jen obtížně. Na obdobných záznamech bylo v rámci studentské práce dosaženo o celých 12.6 % horší skóre než u modelu prezentovaného v této práci. Navržená metoda je schopna v komplexních datech přesně rozměřit až 81 % síňových komplexů.

V práci byl dále navržen nový přístup v segmentaci P vln s využitím neúplných anotací. Algoritmus byl založen na metodě hlubokého kontrastního učení. Původní přínos spočívá v modifikaci kontrastní penalizační funkce pro vzájemné odlišení lokálních segmentů s P vlnou od zbývajících částí signálu. Navržená metoda umožňuje nalézt množství P vln, které jsou zcela skryté v QRST komplexu. V současnosti existuje pouze limitovaný počet

prací, které řeší detekci P vlny pouze v případě částečného překrytí. I tyto metody navíc v celé řadě arytmiických scénářů selhávají.

U hluboké segmentace P vln se jako stěžejní ukázala vhodná augmentace dat. Hluboké modely mají tendenci upřednostňovat pravděpodobnost časového uspořádání P-QRS-T sekvence před relevantními morfologickými příznaky. Tento vliv se nepodařilo eliminovat převzorkováním datové sady pomocí bootstrap metody do různých poměrů zastoupení diagnostických tříd. Provedená analýza znevěhodňuje výsledky hlubokých modelů v problematice detekce P vlny. Částečně lze efekt časového uspořádání zmírnit prostou reverzí časové osy v průběhu učení. I přes principiální jednoduchost by se mohlo jednat o doplňkový způsob testování algoritmů pro segmentaci EKG založených na hlubokých modelech.

Kontrastní učení bylo modifikováno pro optimalizaci vzájemné korelace napříč jedním signálem, mezi všemi signály v trénovací dávce, a dále na různých úrovních abstrakce extrahovaných příznakových map. Tento přístup umožnil vnoření vektorové reprezentace síňové aktivity z lokálních elektrogramů do hlubokého modelu pro segmentaci povrchového EKG. Naučený model je teoreticky možné přenést do dalších oblastí analýzy EKG se zachováním vnořené informace. Na základě hodnocení extrahovaných příznakových map lze hypoteticky dosáhnout lepší interpretability hlubokého modelu.

Literatura

- [1] Paul Kligfield. The centennial of the einthoven electrocardiogram. *Journal of electrocardiology*, 35:123, 2002.
- [2] W Einthoven, AS Jaffe, P Venge, and B Lindahl. Galvanometrische registratie van het menselijk electrocardiogram. *Herinneringsbundel Professor SS Rosenstein*, pages 101–107, 1902.
- [3] Paul Kligfield, Leonard S. Gettes, James J. Bailey, Rory Childers, Barbara J. Deal, E. William Hancock, Gerard van Herpen, Jan A. Kors, Peter Macfarlane, David M. Mirvis, Olle Pahlm, Pentti Rautaharju, and Galen S. Wagner. Recommendations for the standardization and interpretation of the electrocardiogram. *Circulation*, 115(10):1306–1324, March 2007. doi:10.1161/circulationaha.106.180200.
- [4] J J Bailey, A S Berson, A Garson, L G Horan, P W Macfarlane, D W Mortara, and C Zywiets. Recommendations for standardization and specifications in automated electrocardiography: bandwidth and digital signal processing. a report for health professionals by an ad hoc writing group of the committee on electrocardiography and cardiac electrophysiology of the council on clinical cardiology, american heart association. *Circulation*, 81(2):730–739, February 1990. doi:10.1161/01.cir.81.2.730.
- [5] JA Kors and JH Van Bommel. Classification methods for computerized interpretation of the electrocardiogram. *Methods of Information in Medicine*, 29(04):330–336, 1990.
- [6] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015. doi:10.1038/nature14539.
- [7] Ding-Xuan Zhou. Universality of deep convolutional neural networks. *Applied and Computational Harmonic Analysis*, 48(2):787–794, March 2020. doi:10.1016/j.acha.2019.06.004.
- [8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In F Pereira, C J C Burges, L Bottou, and K Q Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012. URL: <https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.
- [9] Clement Farabet, Camille Couprie, Laurent Najman, and Yann LeCun. Learning hierarchical features for scene labeling. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1915–1929, 2012.
- [10] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath,

- et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine*, 29(6):82–97, 2012.
- [11] Michael K. K. Leung, Hui Yuan Xiong, Leo J. Lee, and Brendan J. Frey. Deep learning of the tissue-regulated splicing code. *Bioinformatics*, 30(12):i121–i129, June 2014. doi:10.1093/bioinformatics/btu277.
- [12] Junshui Ma, Robert P. Sheridan, Andy Liaw, George E. Dahl, and Vladimir Svetnik. Deep neural nets as a method for quantitative structure–activity relationships. *Journal of Chemical Information and Modeling*, 55(2):263–274, February 2015. doi:10.1021/ci500747n.
- [13] Sulaiman Somani, Adam J Russak, Felix Richter, Shan Zhao, Akhil Vaid, Fayzan Chaudhry, Jessica K De Freitas, Nidhi Naik, Riccardo Miotto, Girish N Nadkarni, Jagat Narula, Edgar Argulian, and Benjamin S Glicksberg. Deep learning and the electrocardiogram: review of the current state-of-the-art. *EP Europace*, 23(8):1179–1191, February 2021. doi:10.1093/europace/euaa377.
- [14] Grégoire Montavon, Wojciech Samek, and Klaus-Robert Müller. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73:1–15, February 2018. doi:10.1016/j.dsp.2017.10.011.
- [15] Alon Halevy, Peter Norvig, and Fernando Pereira. The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2):8–12, 2009. doi:10.1109/MIS.2009.36.
- [16] Xiangxin Zhu, Carl Vondrick, Charless C. Fowlkes, and Deva Ramanan. Do we need more training data? *International Journal of Computer Vision*, 119(1):76–92, March 2015. doi:10.1007/s11263-015-0812-2.
- [17] George Koulaouzidis, Tomasz Jadczyk, Dimitris K. Iakovidis, Anastasios Koulaouzidis, Marc Bisnaire, and Dafni Charisopoulou. Artificial intelligence in cardiology—a narrative review of current status. *Journal of Clinical Medicine*, 11(13), 2022. URL: <https://www.mdpi.com/2077-0383/11/13/3910>, doi: 10.3390/jcm11133910.
- [18] Jinlei Liu, Yunqing Liu, Yanrui Jin, Xiaojun Chen, Liqun Zhao, and Chengliang Liu. P-wave detection using a parallel convolutional neural network in electrocardiogram. In *2021 4th International Conference on Information Communication and Signal Processing (ICICSP)*, pages 157–161, 2021. doi:10.1109/ICICSP54369.2021.9611986.
- [19] J.P. Martinez, R. Almeida, S. Olmos, A.P. Rocha, and P. Laguna. A wavelet-based ecg delineator: evaluation on standard databases. *IEEE Transactions on Biomedical Engineering*, 51(4):570–581, 2004. doi:10.1109/TBME.2003.821031.

- [20] Abdolrahman Peimankar and Sadasivan Puthusserypady. An ensemble of deep recurrent neural networks for p-wave detection in electrocardiogram. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1284–1288, 2019. doi:10.1109/ICASSP.2019.8682307.
- [21] Rana N. Costandy, Safa M. Gasser, Mohamed S. El-Mahallawy, Mohamed W. Fakhr, and Samir Y. Marzouk. P-wave detection using a fully convolutional neural network in electrocardiogram images. *Applied Sciences*, 10(3):976, February 2020. doi:10.3390/app10030976.
- [22] Jin Qi, Peng Shi, Lin Hu, Tao Zhang, and Shenghua Xie. ECG characteristic wave detection based on deep recursive long short-term memory. *Journal of Medical Imaging and Health Informatics*, 9(9):1920–1924, December 2019. doi:10.1166/jmih.2019.2815.
- [23] Mahsa Akhbari, Mohammad B Shamsollahi, Christian Jutten, Antonis A Armoundas, and Omid Sayadi. ECG denoising and fiducial point extraction using an extended kalman filtering framework with linear and nonlinear phase observations. *Physiological Measurement*, 37(2):203–226, January 2016. doi:10.1088/0967-3334/37/2/203.
- [24] D Panigrahy and P K Sahu. P and T wave detection and delineation of ECG signal using differential evolution (DE) optimization strategy. *Australas. Phys. Eng. Sci. Med.*, 41(1):225–241, March 2018.
- [25] Md Billal Hossain, Syed Khairul Bashar, Allan J. Walkey, David D. McManus, and Ki H. Chon. An accurate qrs complex and p wave detection in ecg signals using complete ensemble empirical mode decomposition with adaptive noise approach. *IEEE Access*, 7:128869–128880, 2019. doi:10.1109/ACCESS.2019.2939943.
- [26] G B Moody and R G Mark. The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol. Mag.*, 20(3):45–50, May 2001.
- [27] R Bousseljot, D Kreiseler, and A Schnabel. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das internet. *Biomed. Tech. (Berl.)*, pages 317–318, July 2009.
- [28] Patrick Wagner, Nils Strodthoff, Ralf-Dieter Bousseljot, Dieter Kreiseler, Fatima I Lunze, Wojciech Samek, and Tobias Schaeffter. PTB-XL, a large publicly available electrocardiography dataset. *Sci. Data*, 7(1):154, May 2020.
- [29] S D Greenwald, R S Patil, and R G Mark. Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information. In *[1990] Proceedings Computers in Cardiology*. IEEE Comput. Soc. Press, 2002.

- [30] Simona Petrutiu, Alan V Sahakian, and Steven Swiryn. Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans. *Europace*, 9(7):466–470, July 2007.
- [31] Hui Liu, Dan Chen, Da Chen, Xiyu Zhang, Huijie Li, Lipan Bian, Minglei Shu, and Yinglong Wang. A large-scale multi-label 12-lead electrocardiogram database with standardized diagnostic statements. *Sci. Data*, 9(1):272, June 2022.
- [32] Pablo Laguna, Roger G Mark, Ary L Goldberger, and George B Moody. The qt database, 1997. URL: <https://physionet.org/content/qtadb/>, doi:10.13026/C24K53.
- [33] JL Willems, P Arnaud, JH van Bommel, R Degani, PW Macfarlane, and C Zywertz. Common standards for quantitative electrocardiography: goals and main results. cse working party. *Methods of information in medicine*, 29(4):263—271, September 1990. URL: <http://europepmc.org/abstract/MED/2233372>.
- [34] Lucie Maršánová, Andrea Němcová, Radovan Smíšek, Tomáš Goldmann, Martin Vítek, and Lukáš Smital. Automatic detection of p wave in ECG during ventricular extrasystoles. In *IFMBE Proceedings*, pages 381–385. Springer Singapore, May 2018. doi:10.1007/978-981-10-9038-7_72.
- [35] Lucie Maršánová, Andrea Nemcova, Radovan Smisek, Lukas Smital, and Martin Vitek. Brno university of technology ecg signal database with annotations of p wave (but pdb), 2021. URL: <https://physionet.org/content/but-pdb/1.0.0/>, doi:10.13026/HWVJ-5B53.
- [36] Alena Kalyakulina, Igor Yusipov, Viktor Moskalenko, Alexander Nikolskiy, Konstantin Kosonogov, Nikolai Zolotykh, and Mikhail Ivanchenko. Lobachevsky university electrocardiography database, 2021.
- [37] Alena I Kalyakulina, Igor I Yusipov, Viktor A Moskalenko, Alexander V Nikolskiy, Konstantin A Kosonogov, Grigory V Osipov, Nikolai Yu Zolotykh, and Mikhail V Ivanchenko. LUDB: A new open-access validation tool for electrocardiogram delineation algorithms. *IEEE Access*, 8:186181–186190, 2020.
- [38] Tawfiq Hasanin, Taghi M. Khoshgoftaar, Joffrey L. Leevy, and Richard A. Bauder. Severely imbalanced big data challenges: investigating data sampling approaches. *Journal of Big Data*, 6(1), November 2019. doi:10.1186/s40537-019-0274-4.
- [39] Matthew A Reyna, Nadi Sadr, Erick A Perez Alday, Annie Gu, Amit J Shah, Chad Robichaux, Ali Bahrami Rad, Andoni Elola, Salman Seyedi, Sardar Ansari, Hamid Ghanbari, Qiao Li, Ashish Sharma, and Gari D Clifford. Issues in the automated classification of multilead ecgs using heterogeneous labels and populations. *Physiol. Meas.*, 43(8):084001, August 2022.

- [40] N P Hughes, S J Roberts, and L Tarassenko. Semi-supervised learning of probabilistic models for ECG segmentation. In *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2005.
- [41] Cristina Gallego Vázquez, Alexander Breuss, Oriella Gnarra, Julian Portmann, Antonio Madaffari, and Giulia Da Poian. Label noise and self-learning label correction in cardiac abnormalities classification. *Physiol. Meas.*, 43(9):094001, September 2022.
- [42] Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. Understanding deep learning requires rethinking generalization. In *International Conference on Learning Representations*, 2017. URL: <https://openreview.net/forum?id=Sy8gdB9xx>.
- [43] Rion Snow, Brendan O’Connor, Daniel Jurafsky, and Andrew Ng. Cheap and fast – but is it good? evaluating non-expert annotations for natural language tasks. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 254–263, Honolulu, Hawaii, October 2008. Association for Computational Linguistics. URL: <https://aclanthology.org/D08-1027>.
- [44] Benoît Frénay and Michel Verleysen. Classification in the presence of label noise: a survey. *IEEE Trans. Neural Netw. Learn. Syst.*, 25(5):845–869, May 2014.
- [45] José A Sáez, Mikel Galar, Julián Luengo, and Francisco Herrera. Analyzing the presence of noise in multi-class problems: alleviating its influence with the One-vs-One decomposition. *Knowl. Inf. Syst.*, 38(1):179–206, January 2014.
- [46] Bin Huang, Yaohai Lin, and Chaoyang Xu. Contrastive label correction for noisy label learning. *Information Sciences*, 611:173–184, September 2022. doi:10.1016/j.ins.2022.08.060.
- [47] Karishma Sharma, Pinar Donmez, Enming Luo, Yan Liu, and I. Zeki Yalniz. NoiseRank: Unsupervised label noise reduction with dependence models. In *Computer Vision – ECCV 2020*, pages 737–753. Springer International Publishing, 2020. doi:10.1007/978-3-030-58583-9_44.
- [48] P.E. McSharry, G.D. Clifford, L. Tarassenko, and L.A. Smith. A dynamical model for generating synthetic electrocardiogram signals. *IEEE Transactions on Biomedical Engineering*, 50(3):289–294, 2003. doi:10.1109/TBME.2003.808805.
- [49] Vajira Thambawita, Jonas L. Isaksen, Steven A. Hicks, Jonas Ghouse, Gustav Ahlberg, Allan Linneberg, Niels Grarup, Christina Ellervik, Morten Salling Olesen, Torben Hansen, Claus Graff, Niels-Henrik Holstein-Rathlou, Inga Strümke, Hugo L. Hammer, Mary M. Malekar, Pål Halvorsen, Michael A. Riegler, and Jørgen K. Kanters. DeepFake electrocardiograms using generative adversarial networks are

- the beginning of the end for privacy issues in medicine. *Scientific Reports*, 11(1), November 2021. doi:10.1038/s41598-021-01295-2.
- [50] Elisa Schenone, Annabelle Collin, and Jean-Frédéric Gerbeau. Numerical simulation of electrocardiograms for full cardiac cycles in healthy and pathological conditions. *International Journal for Numerical Methods in Biomedical Engineering*, 32(5):e02744, September 2015. doi:10.1002/cnm.2744.
- [51] M. A. Quiroz-Juárez, O. Jiménez-Ramírez, R. Vázquez-Medina, V. Breña-Medina, J. L. Aragón, and R. A. Barrio. Generation of ECG signals from a reaction-diffusion model spatially discretized. *Scientific Reports*, 9(1), December 2019. doi:10.1038/s41598-019-55448-5.
- [52] Gustavo Lenis, Nicolas Pilia, Tobias Oesterlein, Armin Luik, Claus Schmitt, and Olaf Dössel. P wave detection and delineation in the ECG based on the phase free stationary wavelet transform and using intracardiac atrial electrograms as reference. *Biomedical Engineering / Biomedizinische Technik*, 61(1):37–56, February 2016. doi:10.1515/bmt-2014-0161.
- [53] The HDF Group. Hierarchical data format version 5, 2000-2010. URL: <http://www.hdfgroup.org/HDF5>.
- [54] F Plesinger, J Jurco, J Halamek, and P Jurak. SignalPlant: an open signal processing software platform. *Physiological Measurement*, 37(7):N38–N48, May 2016. doi:10.1088/0967-3334/37/7/n38.
- [55] K.L. Venkatachalam, Joel E. Herbrandson, and Samuel J. Asirvatham. Signals and signal processing for the electrophysiologist. *Circulation: Arrhythmia and Electrophysiology*, 4(6):965–973, December 2011. doi:10.1161/circep.111.964304.
- [56] D. F Dickinson. The normal ECG in childhood and adolescence. *Heart*, 91(12):1626–1630, December 2005. doi:10.1136/hrt.2004.057307.
- [57] Kirti Khunti. Accurate interpretation of the 12-lead ECG electrode placement: A systematic review. *Health Education Journal*, 73(5):610–623, April 2013. doi:10.1177/0017896912472328.
- [58] Demosthenes G. Katritsis, Giuseppe Boriani, Francisco G. Cosio, Gerhard Hindricks, Pierre Jaïs, Mark E. Josephson, Roberto Keegan, Young-Hoon Kim, Bradley P. Knight, Karl-Heinz Kuck, Deirdre A. Lane, Gregory Y. H. Lip, Helena Malmborg, Hakan Oral, Carlo Pappone, Sakis Themistoclakis, Kathryn A. Wood, Carina Blomström-Lundqvist, Bulent Gorenek, Nikolaos Dargès, Gheorge-Andrei Dan, Marc A Vos, Gulmira Kudaiberdieva, Harry Crijns, Kurt Roberts-Thomson, Yenn-Jiang Lin, Diego Vanegas, Walter Reyes Caorsi, Edmond Cronin, and Jack Rickard. European heart rhythm association (EHRA) consensus document on the management of supraventricular arrhythmias, endorsed by heart rhythm society (HRS), asia-pacific

- heart rhythm society (APHRS), and sociedad latinoamericana de estimulación cardiaca y electrofisiología (SOLAECE). *EP Europace*, 19(3):465–511, November 2016. doi:10.1093/europace/euw301.
- [59] John M. Miller, Rahul Jain, Gopi Dandamudi, Seeongwook Han, and Thomas R. Kambur. Entrainment mapping in supraventricular tachycardias. In *Cardiac Mapping*, pages 665–677. John Wiley & Sons, Ltd, April 2019. doi:10.1002/9781119152637.ch55.
- [60] Ziad F. Issa, John M. Miller, and Douglas P. Zipes. Electrophysiological mechanisms of cardiac arrhythmias. In *Clinical Arrhythmology and Electrophysiology*, pages 51–80. Elsevier, 2019. doi:10.1016/b978-0-323-52356-1.00003-7.
- [61] Daniel W. Kaiser, Henry H. Hsia, Anne M. Dubin, L. Bing Liem, Mohan N. Viswanathan, Paul C. Zei, Paul J. Wang, Sanjiv M. Narayan, and Mintu P. Turakhia. The precise timing of tachycardia entrainment is determined by the postpacing interval, the tachycardia cycle length, and the pacing rate: Theoretical insights and practical applications. *Heart Rhythm*, 13(3):695–703, March 2016. doi:10.1016/j.hrthm.2015.11.032.
- [62] Roger S. Damle, Willie Choe, Nabil M. Kanaan, Frederick A. Ehlert, Jeffrey J. Goldberger, and Alan H. Kadish. Atrial and accessory pathway activation direction in patients with orthodromic supraventricular tachycardia: Insights from vector mapping. *Journal of the American College of Cardiology*, 23(3):684–692, March 1994. doi:10.1016/0735-1097(94)90755-2.
- [63] Daniel Steven, Jens Seiler, Kurt C. Roberts-Thomson, Keiichi Inada, and William G. Stevenson. Mapping of atrial tachycardias after catheter ablation for atrial fibrillation: Use of bi-atrial activation patterns to facilitate recognition of origin. *Heart Rhythm*, 7(5):664–672, May 2010. doi:10.1016/j.hrthm.2010.01.009.
- [64] Patrizio Pascale, Ashok J. Shah, Laurent Roten, Daniel Scherr, Yuki Komatsu, Amir S. Jadidi, Khaled Ramoul, Matthew Daly, Arnaud Denis, Stephen B. Wilton, Nicolas Derval, Frédéric Sacher, Mélèze Hocini, Michel Haïssaguerre, and Pierre Jaïs. Pattern and timing of the coronary sinus activation to guide rapid diagnosis of atrial tachycardia after atrial fibrillation ablation. *Circulation: Arrhythmia and Electrophysiology*, 6(3):481–490, June 2013. doi:10.1161/circep.113.000182.
- [65] Hidekazu Miyazaki, William G. Stevenson, Kent Stephenson, Kyoko Soejima, and Laurence M. Epstein. Entrainment mapping for rapid distinction of left and right atrial tachycardias. *Heart Rhythm*, 3(5):516–523, May 2006. doi:10.1016/j.hrthm.2006.01.014.
- [66] Madison S Spach and Paul C Dolber. Relating extracellular potentials and their derivatives to anisotropic propagation at a microscopic level in human cardiac muscle.

- evidence for electrical uncoupling of side-to-side fiber connections with increasing age. *Circulation research*, 58(3):356–371, 1986.
- [67] M Biermann, M Shenasa, M Borggreffe, G Hindricks, W Haverkamp, and G Breithardt. The interpretation of cardiac electrograms. *Cardiac mapping*, pages 11–34, 1993.
- [68] Gjin Ndrepepa, Edward B. Caref, Hong Yin, Nabil El-Sheris, and Mark Retivo. Activation time determination by high-resolution unipolar and bipolar extracellular electrograms in the canine heart. *Journal of Cardiovascular Electrophysiology*, 6(3):174–188, March 1995. doi:10.1111/j.1540-8167.1995.tb00769.x.
- [69] Milad El Haddad, Richard Houben, Roland Stroobandt, Frederic Van Heuverswyn, Rene Tavernier, and Mattias Duytschaever. Algorithmic detection of the beginning and end of bipolar electrograms: Implications for novel methods to assess local activation time during atrial tachycardia. *Biomedical Signal Processing and Control*, 8(6):981–991, November 2013. doi:10.1016/j.bspc.2012.11.005.
- [70] Carl F Pieper, Russell Blue, and Aantonio Pacifico. Simultaneously collected monopolar and discrete bipolar electrograms: comparison of activation time detection algorithms. *Pacing and Clinical Electrophysiology*, 16(3):426–433, 1993.
- [71] Jan De Pooter, Milad El Haddad, Michael Wolf, Thomas Philips, Frederic Van Heuverswyn, Liesbeth Timmers, René Tavernier, Sebastien Knecht, Yves Vandekerkhove, and Mattias Duytschaever. Clinical assessment and comparison of annotation algorithms in high-density mapping of regular atrial tachycardias. *Journal of Cardiovascular Electrophysiology*, 29(1):177–185, November 2017. doi:10.1111/jce.13371.
- [72] Mark Nothstein, Armin Luik, Amir Jadidi, Jorge Sánchez, Laura A. Unger, Eike M. Wülfers, Olaf Dössel, Gunnar Seemann, Claus Schmitt, and Axel Loewe. CVAR-seg: An automated signal segmentation pipeline for conduction velocity and amplitude restitution. *Frontiers in Physiology*, 12, May 2021. doi:10.3389/fphys.2021.673047.
- [73] Ernesto F. Treo, Daniel O. Cervantes, and Edward J. Ciaccio. Automated detection and mapping of electrical activation when electrogram morphology is complex. *Biomedical Signal Processing and Control*, 8(1):41–49, January 2013. doi:10.1016/j.bspc.2012.04.006.
- [74] J.F. Kaiser. On a simple algorithm to calculate the 'energy' of a signal. In *International Conference on Acoustics, Speech, and Signal Processing*. IEEE, 1990. doi:10.1109/icassp.1990.115702.

- [75] C. Schilling, M. P. Nguyen, A. Luik, C. Schmitt, and O. Dössel. Non-linear energy operator for the analysis of intracardial electrograms. In *IFMBE Proceedings*, pages 872–875. Springer Berlin Heidelberg, 2009. doi:10.1007/978-3-642-03882-2_233.
- [76] Alejandro Alcaine, David Soto-Iglesias, David Andreu, Juan Fernandez-Armenta, Antonio Berruezo, Pablo Laguna, Oscar Camara, and Juan Pablo Martínez. Wavelet-based electrogram onset identification for ventricular electroanatomical mapping. In *Computing in Cardiology 2013*, pages 615–618. IEEE, 2013.
- [77] Jorge Sánchez, Giorgio Luongo, Mark Nothstein, Laura A. Unger, Javier Saiz, Beatriz Trenor, Armin Luik, Olaf Dössel, and Axel Loewe. Using machine learning to characterize atrial fibrotic substrate from intracardiac signals with a hybrid in silico and in vivo dataset. *Frontiers in Physiology*, 12, July 2021. doi:10.3389/fphys.2021.699291.
- [78] B Nowak. Diagnostic value of onset-recordings and marker annotations in dual chamber pacemaker stored electrograms. *Europace*, 5(1):103–109, January 2003. doi:10.1053/eupc.2002.0276.
- [79] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015. URL: <http://arxiv.org/abs/1512.03385>, arXiv:1512.03385.
- [80] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision – ECCV 2016*, pages 630–645, Cham, 2016. Springer International Publishing.
- [81] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 1026–1034, 2015. doi:10.1109/ICCV.2015.123.
- [82] Andrew L Maas, Awni Y Hannun, Andrew Y Ng, et al. Rectifier nonlinearities improve neural network acoustic models. Proceedings of the 30th International Conference on Machine Learning, Atlanta, USA, 17–19 Jun 2013.
- [83] Krishna C. Ravi, David E. Krummen, Ailinh J. Tran, John R. Bullinga, and Sanjiv M. Narayan. Electrocardiographic measurements of regional atrial fibrillation cycle length. *Pacing and Clinical Electrophysiology*, 32:S66–S71, March 2009. doi:10.1111/j.1540-8159.2008.02229.x.
- [84] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(04):834–848, apr 2018. doi:10.1109/TPAMI.2017.2699184.

- [85] Shuyu Gong, Zhewei Wang, Tao Sun, Yuanhang Zhang, Charles D. Smith, Li Xu, and Jundong Liu. Dilated FCN: listening longer to hear better. *CoRR*, abs/1907.11956, 2019. URL: <http://arxiv.org/abs/1907.11956>, arXiv:1907.11956.
- [86] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science*, pages 234–241. Springer International Publishing, 2015. doi:10.1007/978-3-319-24574-4_28.
- [87] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12):2481–2495, December 2017. doi:10.1109/tpami.2016.2644615.
- [88] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection, 2016. URL: <https://arxiv.org/abs/1612.03144>, doi:10.48550/ARXIV.1612.03144.
- [89] Pedro O. Pinheiro, Tsung-Yi Lin, Ronan Collobert, and Piotr Dollár. Learning to refine object segments, 2016. URL: <https://arxiv.org/abs/1603.08695>, doi:10.48550/ARXIV.1603.08695.
- [90] Hyeonwoo Noh, Seunghoon Hong, and Bohyung Han. Learning deconvolution network for semantic segmentation. *CoRR*, abs/1505.04366, 2015. URL: <http://arxiv.org/abs/1505.04366>, arXiv:1505.04366.
- [91] Aaron Carass, Snehashis Roy, Adrian Gherman, Jacob C. Reinhold, Andrew Jesson, Tal Arbel, Oskar Maier, Heinz Handels, Mohsen Ghafoorian, Bram Platel, Ariel Birenbaum, Hayit Greenspan, Dzung L. Pham, Ciprian M. Crainiceanu, Peter A. Calabresi, Jerry L. Prince, William R. Gray Roncal, Russell T. Shinohara, and Ipek Oguz. Evaluating white matter lesion segmentations with refined sørensen-dice analysis. *Scientific Reports*, 10(1), May 2020. doi:10.1038/s41598-020-64803-w.
- [92] Rongjian Zhao, Buyue Qian, Xianli Zhang, Yang Li, Rong Wei, Yang Liu, and Yinggang Pan. Rethinking dice loss for medical image segmentation. In *2020 IEEE International Conference on Data Mining (ICDM)*, pages 851–860, 2020. doi:10.1109/ICDM50108.2020.00094.
- [93] Nabila Abraham and Naimul Mefraz Khan. A novel focal tversky loss function with improved attention u-net for lesion segmentation. *CoRR*, abs/1810.07842, 2018. URL: <http://arxiv.org/abs/1810.07842>, arXiv:1810.07842.
- [94] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015. URL: <http://arxiv.org/abs/1412.6980>.

- [95] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL: <https://openreview.net/forum?id=Bkg6RiCqY7>.
- [96] Ashia C. Wilson, Rebecca Roelofs, Mitchell Stern, Nathan Srebro, and Benjamin Recht. The marginal value of adaptive gradient methods in machine learning. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS’17, page 4151–4161, Red Hook, NY, USA, 2017. Curran Associates Inc.
- [97] Dami Choi, Christopher J. Shallue, Zachary Nado, Jaehoon Lee, Chris J. Maddison, and George E. Dahl. On empirical comparisons of optimizers for deep learning, 2020. URL: <https://openreview.net/forum?id=HygrAR4tPS>.
- [98] Dami Choi, Christopher J. Shallue, Zachary Nado, Jaehoon Lee, Chris J. Maddison, and George E. Dahl. On empirical comparisons of optimizers for deep learning. *CoRR*, abs/1910.05446, 2019. URL: <http://arxiv.org/abs/1910.05446>, arXiv:1910.05446.
- [99] Frank Schneider, Lukas Balles, and Philipp Hennig. Deepobs: A deep learning optimizer benchmark suite. *CoRR*, abs/1903.05499, 2019. URL: <http://arxiv.org/abs/1903.05499>, arXiv:1903.05499.
- [100] Priya Goyal, Piotr Dollár, Ross B. Girshick, Pieter Noordhuis, Lukasz Wesolowski, Aapo Kyrola, Andrew Tulloch, Yangqing Jia, and Kaiming He. Accurate, large minibatch SGD: training imagenet in 1 hour. *CoRR*, abs/1706.02677, 2017. URL: <http://arxiv.org/abs/1706.02677>, arXiv:1706.02677.
- [101] Jerry Ma and Denis Yarats. On the adequacy of untuned warmup for adaptive optimization. *CoRR*, abs/1910.04209, 2019. URL: <http://arxiv.org/abs/1910.04209>, arXiv:1910.04209.
- [102] Alon Halevy, Peter Norvig, and Fernando Pereira. The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2):8–12, 2009. doi:10.1109/MIS.2009.36.
- [103] Chen Sun, Abhinav Shrivastava, Saurabh Singh, and Abhinav Kumar Gupta. Revisiting unreasonable effectiveness of data in deep learning era. *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 843–852, 2017.
- [104] Connor Shorten and Taghi M. Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), July 2019. doi:10.1186/s40537-019-0197-0.
- [105] Jonas Mockus. *Bayesian Approach to Global Optimization*. Springer Netherlands, 1989. doi:10.1007/978-94-009-0909-0.

- [106] James Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011. URL: <https://proceedings.neurips.cc/paper/2011/file/86e8f7ab32cfd12577bc2619bc635690-Paper.pdf>.
- [107] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2005. doi:10.7551/mitpress/3206.001.0001.
- [108] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.
- [109] Monika Butkuvienė, Andrius Petrėnas, Andrius Sološenko, Alba Martín-Yebra, Vaidotas Marozas, and Leif Sörnmo. Considerations on performance evaluation of atrial fibrillation detectors. *IEEE Transactions on Biomedical Engineering*, 68(11):3250–3260, 2021. doi:10.1109/TBME.2021.3067698.
- [110] Frank Hutter, Holger Hoos, and Kevin Leyton-Brown. An efficient approach for assessing hyperparameter importance. In Eric P. Xing and Tony Jebara, editors, *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 754–762, Beijing, China, 22–24 Jun 2014. PMLR. URL: <https://proceedings.mlr.press/v32/hutter14.html>.
- [111] Martinů Žaneta. Extrakce a klasifikace síňové aktivity z intrakardiálních elektrogramů. Bachelor’s thesis.
- [112] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, November 1997. doi:10.1162/neco.1997.9.8.1735.
- [113] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2017. URL: <https://arxiv.org/abs/1706.03762>, doi:10.48550/ARXIV.1706.03762.
- [114] Alhanoof Althnian, Duaa AlSaeed, Heyam Al-Baity, Amani Samha, Alanoud Bin Dris, Najla Alzakari, Afnan Abou Elwafa, and Heba Kurdi. Impact of dataset size on classification performance: An empirical evaluation in the medical domain. *Applied Sciences*, 11(2), 2021. URL: <https://www.mdpi.com/2076-3417/11/2/796>, doi:10.3390/app11020796.
- [115] Konstantinos Sechidis, Grigorios Tsoumakas, and Ioannis Vlahavas. On the stratification of multi-label data. *Machine Learning and Knowledge Discovery in Databases*, pages 145–158, 2011.

- [116] Sheila Tatsumi Kimura-Medorima, Ana Paula Beppler Lazaro Lino, Marcel P. C. Almeida, Marcio J. O. Figueiredo, Lindemberg da Mota Silveira-Filho, Pedro Paulo Martins de Oliveira, Otavio Rizzi Coelho, José Roberto Matos Souza, Wilson Nadruz, Orlando Petrucci, and Andrei C. Sposito. P-wave duration is a predictor for long-term mortality in post-CABG patients. *PLOS ONE*, 13(7):e0199718, July 2018. doi:10.1371/journal.pone.0199718.
- [117] J S Steinberg, S Zelenkofske, S C Wong, M Gelernt, R Sciacca, and E Menchavez. Value of the p-wave signal-averaged ECG for predicting atrial fibrillation after cardiac surgery. *Circulation*, 88(6):2618–2622, December 1993. doi:10.1161/01.cir.88.6.2618.
- [118] Peter J. Stafford, Ian Turner, and Richard Vincent. Quantitative analysis of signal-averaged p waves in idiopathic paroxysmal atrial fibrillation. *The American Journal of Cardiology*, 68(8):751–755, September 1991. doi:10.1016/0002-9149(91)90648-5.
- [119] Dawood Darbar, Arshad Jahangir, Stephen C. Hammill, and Bernard J. Gersh. P wave signal-averaged electrocardiography to identify risk for atrial fibrillation. *Pacing and Clinical Electrophysiology*, 25(10):1447–1453, October 2002. doi:10.1046/j.1460-9592.2002.01447.x.
- [120] T. Thong, J. McNamers, M. Aboy, and B. Goldstein. Prediction of paroxysmal atrial fibrillation by analysis of atrial premature complexes. *IEEE Transactions on Biomedical Engineering*, 51(4):561–569, April 2004. doi:10.1109/tbme.2003.821030.
- [121] P Brugada, J Brugada, L Mont, J Smeets, and E W Andries. A new approach to the differential diagnosis of a regular tachycardia with a wide QRS complex. *Circulation*, 83(5):1649–1659, May 1991. doi:10.1161/01.cir.83.5.1649.
- [122] William J. Brady, Amal Mattu, Jeffrey Tabas, and John D. Ferguson. The differential diagnosis of wide QRS complex tachycardia. *The American Journal of Emergency Medicine*, 35(10):1525–1529, October 2017. doi:10.1016/j.ajem.2017.07.056.
- [123] Kurt C. Roberts-Thomson, Peter M. Kistler, and Jonathan M. Kalman. Focal atrial tachycardia i: Clinical features, diagnosis, mechanisms, and anatomic location. *Pacing and Clinical Electrophysiology*, 29(6):643–652, June 2006. doi:10.1111/j.1540-8159.2006.00413.x.
- [124] Lin Y. Chen, Mina K. Chung, Larry A. Allen, Michael Ezekowitz, Karen L. Furie, Pamela McCabe, Peter A. Noseworthy, Marco V. Perez, and Mintu P. Turakhia. Atrial fibrillation burden: Moving beyond atrial fibrillation as a binary entity: A scientific statement from the american heart association. *Circulation*, 137(20), May 2018. doi:10.1161/cir.0000000000000568.

- [125] A. I. Hernández, G. Carrault, and F. Mora. Improvement of a p-wave detector by a bivariate classification stage. *Transactions of the Institute of Measurement and Control*, 22(3):231–242, August 2000. doi:10.1177/014233120002200303.
- [126] François Portet. P wave detector with PP rhythm tracking: evaluation in different arrhythmia contexts. *Physiol. Meas.*, 29(1):141–155, January 2008.
- [127] Larisa G. Tereshchenko, Jonathan W. Waks, Muammar Kabir, Elyar Ghafoori, Alexei Shvilkin, and Mark E. Josephson. Analysis of speed, curvature, planarity and frequency characteristics of heart vector movement to evaluate the electrophysiological substrate associated with ventricular tachycardia. *Computers in Biology and Medicine*, 65:150–160, October 2015. doi:10.1016/j.compbiomed.2015.03.001.
- [128] Cuiwei Li, Chongxun Zheng, and Changfeng Tai. Detection of ECG characteristic points using wavelet transforms. *IEEE Transactions on Biomedical Engineering*, 42(1):21–28, 1995. doi:10.1109/10.362922.
- [129] J.S. Sahambi, S.N. Tandon, and R.K.P. Bhatt. Using wavelet transforms for ECG characterization. an on-line digital signal processing system. *IEEE Engineering in Medicine and Biology Magazine*, 16(1):77–83, 1997. doi:10.1109/51.566158.
- [130] Nicolai Spicher and Markus Kukuk. Delineation of electrocardiograms using multiscale parameter estimation. *IEEE Journal of Biomedical and Health Informatics*, 24(8):2216–2229, 2020. doi:10.1109/JBHI.2019.2963786.
- [131] Luigi Y Di Marco and Lorenzo Chiari. A wavelet-based ECG delineation algorithm for 32-bit integer online processing. *Biomed. Eng. Online*, 10(1):23, April 2011.
- [132] Hamed Danandeh Hesar and Maryam Mohebbi. A multi rate marginalized particle extended kalman filter for p and t wave segmentation in ecg signals. *IEEE Journal of Biomedical and Health Informatics*, 23(1):112–122, 2019. doi:10.1109/JBHI.2018.2794362.
- [133] Arturo Martínez, Raúl Alcaraz, and José Joaquín Rieta. Application of the phasor transform for automatic delineation of single-lead ECG fiducial points. *Physiological Measurement*, 31(11):1467–1485, September 2010. doi:10.1088/0967-3334/31/11/005.
- [134] Lucie Maršánová, Andrea Němcová, Radovan Smíšek, Martin Vitek, and Lukáš Smital. Advanced p wave detection in ecg signals during pathology: Evaluation in different arrhythmia contexts. *Scientific Reports*, 9(1), December 2019. doi:10.1038/s41598-019-55323-3.
- [135] Lucie Saclova, Andrea Nemcova, Radovan Smisek, Lukas Smital, Martin Vitek, and Marina Ronzhina. Reliable P wave detection in pathological ECG signals. *Sci. Rep.*, 12(1):6589, April 2022.

- [136] M. Stridh and L. Sommo. Spatiotemporal QRST cancellation techniques for analysis of atrial fibrillation. *IEEE Transactions on Biomedical Engineering*, 48(1):105–111, 2001. doi:10.1109/10.900266.
- [137] Arne Sippensgroenewegen, Michael D. Mlynash, Franz X. Roithinger, Yoshinari Goseki, and Michael D. Lesh. Electrocardiographic analysis of ectopic atrial activity obscured by ventricular repolarization: P wave isolation using an automatic 62-lead QRST subtraction algorithm. *Journal of Cardiovascular Electrophysiology*, 12(7):780–790, July 2001. doi:10.1046/j.1540-8167.2001.00780.x.
- [138] John P. Marengo, Hiroshi Nakagawa, Shawn Yang, David MacAdam, Lucien Xu, Ding S. He, Mark S. Link, Munther K. Homoud, N.A. Mark Estes III, and Paul J. Wang. Testing of a new t-wave subtraction algorithm as an aid to localizing ectopic atrial beats. *Annals of Noninvasive Electrocardiology*, 8(1):55–59, January 2003. doi:10.1046/j.1542-474x.2003.08109.x.
- [139] J.J. Rieta, F. Castells, C. Sanchez, V. Zarzoso, and J. Millet. Atrial activity extraction for atrial fibrillation analysis using blind source separation. *IEEE Transactions on Biomedical Engineering*, 51(7):1176–1186, July 2004. doi:10.1109/tbme.2004.827272.
- [140] C. Vaya, J.J. Rieta, C. Sanchez, and D. Moratal. Convolutional blind source separation algorithms applied to the electrocardiogram of atrial fibrillation: Study of performance. *IEEE Transactions on Biomedical Engineering*, 54(8):1530–1533, August 2007. doi:10.1109/tbme.2006.889778.
- [141] Awni Y. Hannun, Pranav Rajpurkar, Masoumeh Haghpanahi, Geoffrey H. Tison, Codie Bourn, Mintu P. Turakhia, and Andrew Y. Ng. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 25(1):65–69, January 2019. doi:10.1038/s41591-018-0268-3.
- [142] Shenda Hong, Yuxi Zhou, Junyuan Shang, Cao Xiao, and Jimeng Sun. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. *Computers in Biology and Medicine*, 122:103801, July 2020. doi:10.1016/j.combiomed.2020.103801.
- [143] Konrad Duraj, Natalia Piaseczna, Paweł Kostka, and Ewaryst Tkacz. Semantic segmentation of 12-lead ECG using 1d residual u-net with squeeze-excitation blocks. *Applied Sciences*, 12(7):3332, March 2022. doi:10.3390/app12073332.
- [144] Aman Malali, Srinidhi Hiriyannaiah, Siddesh G.M., Srinivasa K.G., and Sanjay N.T. Supervised ECG wave segmentation using convolutional LSTM. *ICT Express*, 6(3):166–169, September 2020. doi:10.1016/j.icte.2020.04.004.

- [145] Analyzing artificial intelligence systems for the prediction of atrial fibrillation from sinus-rhythm ECGs including demographics and feature visualization. *Scientific Reports*, 11(1):1–10, 2021. doi:10.1038/s41598-021-02179-1.
- [146] Shaan Khurshid, Samuel Friedman, Christopher Reeder, Paolo Di Achille, Nathaniel Diamant, Pulkit Singh, Lia X. Harrington, Xin Wang, Mostafa A. Al-Alusi, Gopal Sarma, Andrea S. Foulkes, Patrick T. Ellinor, Christopher D. Anderson, Jennifer E. Ho, Anthony A. Philippakis, Puneet Batra, and Steven A. Lubitz. Electrocardiogram-based deep learning and clinical risk factors to predict atrial fibrillation. *Circulation*, November 2021. doi:10.1161/circulationaha.121.057480.
- [147] Serhat Selcuk Bucak, Rong Jin, and Anil K. Jain. Multi-label learning with incomplete class assignments. In *CVPR 2011*. IEEE, June 2011. doi:10.1109/cvpr.2011.5995734.
- [148] Baoyuan Wu, Fan Jia, Wei Liu, Bernard Ghanem, and Siwei Lyu. Multi-label learning with missing labels using mixed dependency graphs. *International Journal of Computer Vision*, 126(8):875–896, April 2018. doi:10.1007/s11263-018-1085-3.
- [149] Anhui Tan, Xiaowan Ji, Jiye Liang, Yuzhi Tao, Wei-Zhi Wu, and Witold Pedrycz. Weak multi-label learning with missing labels via instance granular discrimination. *Information Sciences*, 594:200–216, May 2022. doi:10.1016/j.ins.2022.02.011.
- [150] Lijuan Sun, Gengyu Lyu, Songhe Feng, and Xiankai Huang. Beyond missing: weakly-supervised multi-label learning with incomplete and noisy labels. *Applied Intelligence*, 51(3):1552–1564, September 2020. doi:10.1007/s10489-020-01878-y.
- [151] Qiuqiang Kong, Changsong Yu, Yong Xu, Turab Iqbal, Wenwu Wang, and Mark D. Plumbley. Weakly labelled AudioSet tagging with attention neural networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(11):1791–1802, November 2019. doi:10.1109/taslp.2019.2930913.
- [152] Weiqiang Ren, Kaiqi Huang, Dacheng Tao, and Tieniu Tan. Weakly supervised large scale object localization with multiple instance learning and bag splitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):405–416, February 2016. doi:10.1109/tpami.2015.2456908.
- [153] Wu-Jun Li and Dit Yan yeung. MILD: Multiple-instance learning via disambiguation. *IEEE Transactions on Knowledge and Data Engineering*, 22(1):76–89, January 2010. doi:10.1109/tkde.2009.58.
- [154] Yixin Chen, Jinbo Bi, and J.Z. Wang. MILES: Multiple-instance learning via embedded instance selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):1931–1947, December 2006. doi:10.1109/tpami.2006.248.

- [155] Thomas G Dietterich, Richard H Lathrop, and Tomás Lozano-Pérez. Solving the multiple instance problem with axis-parallel rectangles. *Artificial intelligence*, 89(1-2):31–71, 1997.
- [156] Cha Zhang, John Platt, and Paul Viola. Multiple instance boosting for object detection. *Advances in neural information processing systems*, 18, 2005.
- [157] Nicolas Gonthier, Saïd Ladjal, and Yann Gousseau. Multiple instance learning on deep features for weakly supervised object detection with extreme domain shifts. *Computer Vision and Image Understanding*, 214:103299, January 2022. doi:10.1016/j.cviu.2021.103299.
- [158] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. *CoRR*, abs/1406.4729, 2014. URL: <http://arxiv.org/abs/1406.4729>, arXiv:1406.4729.
- [159] S. Chopra, R. Hadsell, and Y. LeCun. Learning a similarity metric discriminatively, with application to face verification. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 539–546 vol. 1, 2005. doi:10.1109/CVPR.2005.202.
- [160] Florian Schroff, Dmitry Kalenichenko, and James Philbin. FaceNet: A unified embedding for face recognition and clustering. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, jun 2015. doi:10.1109/cvpr.2015.7298682.
- [161] Kihyuk Sohn. Improved deep metric learning with multi-class n-pair loss objective. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. URL: <https://proceedings.neurips.cc/paper/2016/file/6b180037abbebea991d8b1232f8a8ca9-Paper.pdf>.
- [162] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations, 2020. URL: <https://arxiv.org/abs/2002.05709>, doi:10.48550/ARXIV.2002.05709.
- [163] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschiot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning, 2020. URL: <https://arxiv.org/abs/2004.11362>, doi:10.48550/ARXIV.2004.11362.
- [164] F. Wang and H. Liu. Understanding the behaviour of contrastive loss. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2495–2504, Los Alamitos, CA, USA, jun 2021. IEEE Computer Society. URL: <https://doi.ieeecomputersociety.org/10.1109/CVPR46437.2021.00252>, doi:10.1109/CVPR46437.2021.00252.

- [165] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, March 2004. doi:10.1017/cbo9780511804441.
- [166] Zhi-Hua Zhou and Min-Ling Zhang. Neural networks for multi-instance learning. In *Proceedings of the International Conference on Intelligent Information Technology, Beijing, China*, pages 455–459, 2002.
- [167] Maximilian Ilse, Jakub Tomczak, and Max Welling. Attention-based deep multiple instance learning. In *International conference on machine learning*, pages 2127–2136. PMLR, 2018.
- [168] Xinggong Wang, Yongluan Yan, Peng Tang, Xiang Bai, and Wenyu Liu. Revisiting multiple instance neural networks. *Pattern Recognition*, 74:15–24, February 2018. doi:10.1016/j.patcog.2017.08.026.
- [169] Anwer Dhala, Donald Underwood, Robert Leman, Ernest Madu, Dainia Baugh, Yukio Ozawa, Yuji Kasamaki, Qiuzhen Xue, and Shankara Reddy. Signal-averaged p-wave analysis of normal controls and patients with paroxysmal atrial fibrillation: A study in gender differences, age dependence, and reproducibility. *Clinical Cardiology*, 25(11):525–531, November 2002. doi:10.1002/c1c.4960251109.
- [170] Tomas Vicar, Petra Novotna, Jakub Hejc, Oto Janousek, and Marina Ronzhina. Cardiac abnormalities recognition in ecg using a convolutional network with attention and input with an adaptable number of leads. In *2021 Computing in Cardiology (CinC)*, volume 48, pages 1–4, 2021. doi:10.23919/CinC53138.2021.9662806.
- [171] Lucie Saclova, Andrea Nemcova, Radovan Smisek, Lukas Smital, Martin Vitek, and Marina Ronzhina. A new database with annotations of p waves in ECGs with various types of arrhythmias. *Physiological Measurement*, 43(10):10NT01, October 2022. doi:10.1088/1361-6579/ac944e.
- [172] Guido Van Rossum and Fred L Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.
- [173] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. URL: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- [174] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework.

CoRR, abs/1907.10902, 2019. URL: <http://arxiv.org/abs/1907.10902>, arXiv:1907.10902.

- [175] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi:10.1038/s41592-019-0686-2.
- [176] Skipper Seabold and Josef Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.

Seznam příloh

A Přílohy	125
A.1 Vybrané implementační detaily	125
A.1.1 Parametry augmentačních technik v Kapitole 2	125
A.1.2 Parametry augmentačních technik v Kapitole 3	126
A.1.3 Definice zahřívací fáze	126
A.1.4 Softwarové a hardwarové prostředky	127
B Curriculum vitae	128

A Přílohy

A.1 Vybrané implementační detaily

A.1.1 Parametry augmentačních technik v Kapitole 2

V Tabulce A.1 uvádíme přehled nastavení parametrů augmentačních technik. Ve čtvrtém sloupci je uvedena pravděpodobnost (p) s jakou bude transformace aplikována. Způsob aplikace značí, jestli je transformace či nastavení parametru společné pro celý signál či je prováděna nezávisle pro každý kanál.

Tab. A.1: Specifikace parametrů augmentačních technik.

Zkratka	parametr	hodnota/rozsah	p	způsob aplikace
T_{HP}	aktivace typ f_{mez}	ano/ne FIR 5 Hz	1.0	signál
T_{clip}	aktivace práh	ano/ne 12.46 mV	1.0	signál
T_{std}	aktivace std	ano/ne 0.79 mV	1.0	signál
$T_{stretch}$	aktivace limit	ano/ne $\pm 20\%$	0.5	signál
T_{shift}	aktivace limit	ano/ne $\pm 90\%$	0.8	signál
T_{swap}	aktivace	ano/ne	0.2	signál
T_{rev}	aktivace	ano/ne	0.1	kanál
T_{gain}	aktivace limit	ano/ne $\pm 20\%$	0.8	kanál
T_{zero}	aktivace	ano/ne	0.1	kanál
T_{stim}	aktivace šířka pulzu šířka přechodu délka cyklu amplituda	ano/ne 2–20 ms 5–30 ms 0.2–2 s 0.5–5.0 mV	0.1 0.1	signál kanál
T_{white}	aktivace SNR	ano/ne 20–0 dB	0.1	kanál kanál
T_{pwr}	aktivace SNR	ano/ne 20–0 dB	0.1	kanál kanál

A.1.2 Parametry augmentačních technik v Kapitole 3

V Tabulce A.2 uvádíme přehled nastavení parametrů augmentačních technik využitých v rámci segmentace P vln v Kapitole 3.

Tab. A.2: Specifikace parametrů augmentačních technik.

Zkratka	parametr	hodnota/rozsah	p	způsob aplikace
T _{HP}	aktivace	ano/ne	1.0	signál
	typ	FIR		
	f _{mez}	0.5 Hz		
T _{LP}	aktivace	ano/ne	1.0	signál
	typ	FIR		
	f _{mez}	40 Hz		
T _{clip}	aktivace práh	ano/ne 5.26 mV	1.0	signál
T _{std}	aktivace std	ano/ne 3.15 mV	1.0	signál
T _{stretch}	aktivace limit	ano/ne ±20%	0.5	signál
T _{swap}	aktivace	ano/ne	0.2	signál
T _{time}	aktivace	ano/ne	0.5	signál
T _{shift}	aktivace	ano/ne	0.8	signál
	limit	±90%		
T _{rev}	aktivace	ano/ne	0.1	kanál
T _{gain}	aktivace	ano/ne	0.8	kanál
	limit	±30%		
T _{zero}	aktivace	ano/ne	0.05	kanál
T _{white}	aktivace	ano/ne	0.2	kanál
	SNR	40–20 dB		kanál
T _{pwr}	aktivace	ano/ne	0.2	kanál
	SNR	40–20 dB		kanál

A.1.3 Definice zahřívací fáze

Velikost kroku učení v každé iteraci i až do epochy N , kde M_T je celkový počet vzorů v trénovací datové sadě a BS_T je počet vzorů v každé trénovací dávce, je určena lineárním vztahem:

$$\eta_{i,N} = iN \frac{M_T}{BS_T} \eta_0, \quad (\text{A.1})$$

A.1.4 Softwarové a hardwarové prostředky

Implementace hlubokých modelů, sekvenčních modulů pro předzpracování dat a dalších pomocných procesů byla provedena v prostředí Python v. 3.8.1 [172] s využitím knihovny PyTorch v. 1.12.1 [173]. Pro prohledávání prostoru parametrů byla využita knihovna Optuna v. 2.10.1 [174]. Statistické zpracování výsledků bylo provedeno pomocí balíčku SciPy v. 1.8.1 [175] a Statsmodels v. 0.13.2 [176]. Učení modelů bylo provedeno na grafické jednotce NVIDIA RTX 3060 12 GB.

B Curriculum vitae

Jméno:	Jakub Hejč
ORCID:	0000-0001-5743-9960
Web of Science:	Q-9341-2017
WOS publikace:	18
WOS citace:	61
H-index:	5

Vzdělání

2014–nyní	Doktorské studium Biomedicínské technologie a bioinformatika, Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií
2010–2013	Magisterské studium Biomedicínské inženýrství a bioinformatika, Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií
2007–10	Bakalářské studium Biomedicínská technika a bioinformatika, Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií

Pracovní praxe

2021–nyní	Mezinárodní centrum klinického výzkumu (ICRC) ve Fakultní nemocnici u sv. Anny Role: juniorní výzkumný pracovník, výzkumný tým Intervenční srdceční elektrofyziologie pod vedením doc. MUDr. Stárka, Ph.D.
2017–2021	Vysoké učení technické v Brně, Ústav biomedicínského inženýrství Role: lektor, výzkumný pracovník
2017–nyní	Fakultní nemocnice Brno – Dětská nemocnice Role: biomedicínský inženýr v elektrofyziologické laboratoři
2008–2013	Fakultní nemocnice Brno – Dětská nemocnice Role: biomedicínský technik v elektrofyziologické laboratoři
2005–2007	Fakultní nemocnice u sv. Anny Role: biomedicínský technik v elektrofyziologické laboratoři

Lektorská činnost

2017–2021	Algoritmizace a programování
2018–2021	Umělá inteligence v medicíně
2014–2018	Laboratorní technika v genomice a proteomice

Projekty

2019–2021	Development of Deep Learning Models for Multi-stage Cardiac Arrhythmia Recognition (TAČR TJ02000331) Role: hlavní řešitel, vedoucí týmu
2019–2021	Computing Methods for Objective Evaluation of Nonverbal Communication Skills (TAČR TJ02000293) Role: člen týmu
2018–2019	Development of Machine Learning Models for the Detection of Atrial Fibrillation (TAČR TJ02000293) Role: člen týmu

Kurzy a další vzdělání

2020	Advanced Machine Learning with TensorFlow on Google Cloud Platform Specialization On-line kurz certifikovaný společností Google
2019	Machine Learning On-line kurz certifikovaný Stanford University
2019	Erasmus+ Kaplan International Edinburgh
2018	Humusoft Technical Computing Camp
2017–2022	České a slovenské sympóziium o arytmiích a kardiostimulaci

Soutěže a ocenění

2021	Moody PhysioNet Challenge <i>Will Two Do? Varying Dimensions in Electrocardiography</i>
2020	Moody PhysioNet Challenge <i>Classification of 12-lead ECGs</i> , 1. místo v předběžném kole
2017	Moody PhysioNet Challenge <i>Atrial Fibrillation Classification from a Short Single Lead ECG Recording</i> , 9. místo ve finálovém kole
2014	Student EEICT 3. místo v soutěži pro doktorskou formu studia

Vybrané publikace

LEHAR, F.; SZEGEDI, N.; **HEJČ, J.**; JEZ, J.; SOUCEK, F.; KULÍK, T.; ŠIRŮČKOVÁ, A.; SALO, Z.; NAGY, K.; MERKELY, B.; GELLER, L.; STÁREK, Z. Randomized comparison of atrioventricular node re-entry tachycardia and atrial flutter catheter ablation with and without fluoroscopic guidance: ZeroFluoro study. *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*, 2022, roč. 24, č. 10, s. 1636-1644. ISSN: 1532-2092.

HEJČ, J.; POSPÍŠIL, D.; NOVOTNÁ, P.; PEŠL, M.; JANOUŠEK, O.; RONZHINA, M.; STÁREK, Z. Segmentation of Atrial Activity in Intracardiac Electrograms (EGMs) Using Convolutional Neural Network (CNN) Trained on Small Imbalanced Dataset. In *Computing in Cardiology 2021. Computers in Cardiology. Computing in Cardiology 2021*, 2021. s. 1-4. ISSN: 0276-6574.

JANOUŠEK, O.; **HEJČ, J.**; POSPÍŠIL, D. A new filtering method for smoothing intracardiac records preserving the steepness of A, V, H waves. In *Computing in Cardiology 2021. Computers in Cardiology. Computing in Cardiology 2021*, 2021. s. 1-4. ISSN: 0276-6574.

NOVOTNÁ, P.; VIČAR, T.; **HEJČ**, J.; RONZHINA, M. Deep-Learning Premature Contraction Localization Using Gaussian Based Predicted Data. In Computing in Cardiology 2021. Computing in Cardiology. Computing in Cardiology 2021, 2021. s. 1-4. ISSN: 2325-887X.

PEŠL, M.; **HEJČ**, J.; KULIK, T.; VIČAR, T.; NOVOTNÁ, P.; RONZHINA, M.; JAKUBÍK, J.; LEINVEBER, P.; GONZALEZ RIVAS, J.; STÁREK, Z. Cardiac arrhythmias classification in Kardiovizie population study. In Computing in Cardiology 2021. Computing in Cardiology. IEEE Computer Society, 2021. s. 1-4. ISSN: 2325-887X.

VIČAR, T.; NOVOTNÁ, P.; **HEJČ**, J.; JANOUŠEK, O.; RONZHINA, M. Cardiac Abnormalities Recognition in ECG Using a Convolutional Network with Attention and Input with an Adaptable Number of Leads. In Computing in Cardiology 2021. Computing in Cardiology. Brno: Computing in Cardiology 2021, 2021. s. 1-4. ISSN: 2325-887X.

VIČAR, T.; **HEJČ**, J.; NOVOTNÁ, P.; RONZHINA, M.; JANOUŠEK, O. ECG Abnormalities Recognition Using Convolutional Network With Global Skip Connections and Custom Loss Function. In Computing in Cardiology 2020. Computing in Cardiology 2013. 47. NEW YORK: IEEE, 2020. s. 1-4. ISBN: 978-1-7281-7382-5. ISSN: 2325-8861.

SMÍŠEK, R.; **HEJČ**, J.; RONZHINA, M.; NĚMCOVÁ, A.; MARŠÁNOVÁ, L.; KOLÁŘOVÁ, J.; SMITAL, L.; VÍTEK, M. Multi-stage SVM approach for cardiac arrhythmias detection in short single-lead ECG recorded by a wearable device. PHYSIOLOGICAL MEASUREMENT, 2018, roč. 39, č. 9, s. 1-14. ISSN: 0967-3334.

SMÍŠEK, R.; **HEJČ**, J.; RONZHINA, M.; NĚMCOVÁ, A.; MARŠÁNOVÁ, L.; CHMELÍK, J.; KOLÁŘOVÁ, J.; PROVAZNÍK, I.; SMITAL, L.; VÍTEK, M. SVM Based ECG Classification Using Rhythm and Morphology Features, Cluster Analysis and Multilevel Noise Estimation. In Computing in Cardiology 2017. Computers in Cardiology. Rennes, France: 2017. s. 1-4. ISBN: 978-1-5090-0684-7. ISSN: 0276-6574.

RONZHINA, M.; OLEJNÍČKOVÁ, V.; STRAČINA, T.; NOVÁKOVÁ, M.; JANOUŠEK, O.; **HEJČ**, J.; KOLÁŘOVÁ, J.; HLAVÁČOVÁ, M.; PAULOVÁ, H. Effect of increased left ventricle mass on ischemia assessment in electrocardiographic signals: rabbit isolated heart study. BMC Cardiovascular Disorders, 2017, roč. 17, č. 216, s. 1-11. ISSN: 1471-2261.

HEJČ, J.; VÍTEK, M.; RONZHINA, M.; NOVÁKOVÁ, M.; KOLÁŘOVÁ, J. A Wavelet-Based ECG Delineation Method: Adaptation to an Experimental Electrograms with Manifested Global Ischemia. Cardiovascular Engineering and Technology, 2015, roč. 2015, č. 6, s. 364-375. ISSN: 1869-408X.