

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

DEPARTMENT OF BIOMEDICAL ENGINEERING

# METODY HLUBOKÉHO UČENÍ PRO SEGMENTACI ELEKTROKARDIOGRAFICKÝCH SIGNÁLŮ

SEGMENTATION OF ELECTROCARDIOGRAPHIC SIGNALS  
USING DEEP LEARNING METHODS

ZKRÁCENÁ VERZE DIZERTAČNÍ PRÁCE  
SHORTENED DOCTORAL THESIS

BIOMEDICÍNSKÉ TECHNOLOGIE A BIOINFORMATIKA  
BIOMEDICAL TECHNOLOGY AND BIOINFORMATICS

AUTOR PRÁCE  
AUTHOR

Ing. JAKUB HEJČ

ŠKOLITEL  
SUPERVISOR

doc. Ing. JANA KOLÁŘOVÁ, Ph.D.

OPONENTI  
OPONENTS

BRNO 2023

## **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.

## **KEYWORDS**

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

## **MÍSTO ULOŽENÍ**

Areálová knihovna VUT, Technická 12

# Obsah

Úvod	5
Cíle práce	6
<b>1 Zpracování klinických dat</b>	<b>7</b>
1.1 Úvod	7
1.2 Zpětná analýza dat	7
1.3 Učební datové sady	8
1.3.1 Charakteristika zařazených pacientů	8
1.3.2 Selektce zájmových úseků	9
1.3.3 Expertně anotovaná datová sada a (DS-A)	9
1.3.4 Automaticky anotovaná datová sada B (DS-B)	10
1.4 Diskuze a limitace	11
<b>2 Segmentace síňových elektrogramů pomocí modelů hlubokého učení</b>	<b>12</b>
2.1 Současný stav	12
2.2 Formální definice problému	12
2.3 Architektura modelu	13
2.3.1 Blok dilatovaných konvolucí	13
2.4 Stratifikace DS-A	14
2.5 Kriteriaální funkce	15
2.6 Optimalizace modelu	15
2.7 Hodnocení kvality modelu	16
2.8 Výsledky	17
2.8.1 Optimalizace architektury	17
2.8.2 Optimalizace kriteriaální funkce	18
2.8.3 Srovnání modelů $M_{\text{aopt}}$ a $M_{\text{lopt}}$	19
2.8.4 Segmentace fibrilačních záznamů	21
2.8.5 Hodnocení vlivu šumu	21
2.8.6 Vliv stratifikace datové sady	24
2.8.7 Ukázka segmentace	25
2.9 Diskuze a závěr	25
<b>3 Slabě dozorovaná segmentace P vln</b>	<b>26</b>
3.1 Současný stav	26
3.2 Formulace problému segmentace s neúplnou informací	26
3.3 Architektura modelu	27
3.3.1 Regionální příznakové mapy	27
3.4 Normalizovaná křížová entropie pro kontrastní učení	27

3.4.1 Kontrastní učení v lokalizované oblasti signálu . . . . .	28
3.5 MIL vrstva . . . . .	29
3.6 Popis experimentů . . . . .	29
3.7 Trénovací data . . . . .	29
3.8 Předzpracování dat . . . . .	30
3.9 Hodnocení kvality segmentace . . . . .	30
3.10 Výsledky . . . . .	30
3.11 Ukázky segmentace . . . . .	32
3.12 Diskuze . . . . .	33
3.12.1 Jak věrohodná je segmentace P vlny hlubokým modelem? . . . . .	33
3.12.2 Limitace . . . . .	34
<b>Závěr</b>	<b>35</b>
<b>Literatura</b>	<b>36</b>
<b>A Curriculum vitae</b>	<b>43</b>

# Úvod

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

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 a morfologie těchto vln patří mezi významné klinické indikátory. Jejich kvantitativní hodnocení je základním krokem při diagnostické interpretaci EKG záznamu [2]. Časovou lokalizaci, stanovení počátku a konce zmíněných útvarů označujeme obecně jako *segmentaci* EKG signálu.

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 [2] je vizualní interpretace zatížena lidskou chybou, která může v krajním případě vést k mylnému závěru.

Tyto důvody vedly k rozvoji výpočetních metod, jejichž cílem je automatizace analýzy EKG a extrakce klinicky konzistentních parametrů bez ohledu na komplexnost signálu. U konvenčních algoritmů je návrh celého řetězce založen na expertních znalostech o vlastnostech EKG signálu [3]. Extrahované příznaky tak obvykle mají snadno interpretovatelný vztah s klinickými projevy. Neúplné lidské poznání v oblasti elektrokardiografie však v důsledku zanáší do analýzy další zkreslení, Výsledkem je vyšší míra nepřesností než u lidských hodnotitelů.

Metody hlubokého učení jsou schopny získat vhodnou reprezentaci komplexních dat na různých úrovních abstrakce. Hluboké neuronální modely (DNN<sup>1</sup>) patří mezi potenciální univerzální aproximátory [4] a v poslední dekádě dosáhly významných úspěchů v řadě oblastí [5, 6, 7] včetně analýzy EKG signálů [8]. Metody umožňují automaticky extrahovat relevantní příznaky, čímž lze snížit vliv nevhodného předzpracování a případně identifikovat zcela nové vztahy mezi daty a klinickým obrazem pacienta. DNN modely reprezentují rozhodovací proces založený řádově na milionech parametrů. Přes dosavadní pokrok [9] je porozumění komplexním vztahům uvnitř hlubokých modelů stále limitováno, čímž je limitována i důvěra v tyto systémy. Přeparametrizování modelu dále souvisí s problematikou aproximace šumu a dosažení robustní generalizace. DNN metody proto v současnosti vyžadují dostupnost adekvátního množství reprezentativních dat [10]. V oblasti segmentace a klasifikace EKG jsou využívány standardní databáze signálů. Ty však trpí některými nedostatky jako je nedostatečný počet subjektů, asymetrické zastoupení diagnostických tříd, či referenční hodnoty s nízkou věrohodností.

---

<sup>1</sup>Z 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.

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.

# 1 Zpracování klinických dat

## 1.1 Úvod

Převážná většina prací [11, 12, 13] zaměřených na automatizovanou interpretaci EKG využívá některou ze standardních databází. Databáze určené pro vývoj segmentačních algoritmů [14, 15, 16, 17] obsahují primárně fyziologické záznamy nebo záznamy, které mají minimální vliv na morfologii a časové uspořádání P vln. Potenciálně závažné choroby, jako jsou např. tachykardie s širokým QRS komplexem, jsou reprezentovány spíše sporadicky. Nerovnoměrnost zastoupení diagnostických tříd přináší řadu problémů z pohledu optimalizace [18] a konzistence výsledků napříč různými kohortami pacientů [19]. Sporadicky je ve standardních databázích provedena nezávislá křížová validace více hodnotitelů. Ani mezi zkušenými kardiology přitom nepanuje shoda na lokalizaci hranic EKG segmentů či pravidlech pro stanovení diagnostické třídy [20], což se odráží na kvalitě anotací. Konzervativní odhad chybovosti diagnostických tříd činí ve standardních databázích kolem 13 % [21]. 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. Anotační šum je přitom považován za potenciálně vyšší riziko pro model než konvenční šum na straně extrahovaných příznaků [22].

V průběhu EFV je snímáno nejen povrchové 12-svodové EKG, ale také intrakardiální elektrogramy (EGM). EGM poskytují informaci o časově-prostorovém uspořádání elektrické aktivity lokálních úseků myokardu. Jedná se o významné parametry pro charakterizaci elektrických vlastností tkáně a jejich interpretace umožňuje ve většině případů přímé klinické potvrzení původce obtíží. EFV podstupuje široká skupina pacientů jejichž EKG data by mohla významnou měrou přispět ke zvýšení kvality současných modelů strojového učení. EFV data byla dosud u zmíněných úloh využívána zcela minimálně [23].

## 1.2 Zpětná analýza dat

EFV data byla snímána pomocí komerčního systému WorkMate 4.2 a 4.3 (Abbott). Systém WorkMate nativně ukládá data do proprietárních binárních formátů s nedokumentovanou strukturou. Z toho důvodu byly aplikovány postupy reverzního inženýrství s cílem identifikovat datovou strukturu a implementovat softwarové nástroje pro konverzi dat do otevřených formátů. Navržené nástroje jsou veřejně dostupné a kromě naplnění cílů této práce by mohli k širšímu využití EFV záznamů vědeckou komunitou. Obecný postup analýzy je znázorněn pomocí Pseudokódu 2. Výsledek identifikace hlavního datového souboru *\*.log* se záznamem surových elektrokardiografických signálů je uveden na Obrázku 1.1.

---

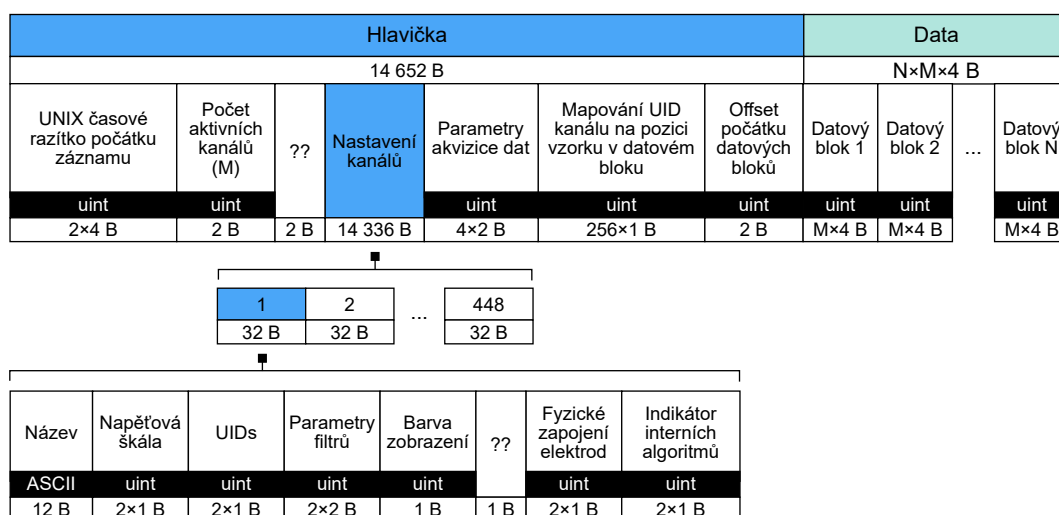
**Algorithm 1** Postup zpětné analýzy

---

**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 a porovnej datovou strukturu souborů v kroku  $p_i$  a  $p_{i-1}$
  - 4:     **if** existuje rozdíl mezi  $p_i$  a  $p_{i-1}$  **then**
  - 5:         Identifikuj rozsah adres  $\{a_i, b_i\}$  odpovídající nalezenému rozdílu
  - 6:          $A \leftarrow \{a_i, b_i\}$
  - 7:     **end if**
  - 8: **end for**
  - 9: Proveď agregaci potenciálních adres pro každý parametr
- 



Obr. 1.1: Struktura nativního datového formátu pro archivaci signálů \*.log.

## 1.3 Učební datové sady

### 1.3.1 Charakteristika zařazených pacientů

Jedná se o retrospektivní elektrokardiografická data, která byla měřena u pacientů indikovaných k EFV ve FN Brno. Pacienti podepsali Informovaný souhlas se zpracováním klinických dat a využití dat 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: a) chybějící EGM z CS; b) chybějící záznam arytmie > 5 s.; c) opakovaná EFV bez nového nálezu.

### 1.3.2 Selekcce zájmových úseků

U EFV studií byla provedena vizuální inspekce EKG a EGM záznamů s následným výběrem zájmových úseků. Klinická interpretace úseků vycházela z EFV nálezů potvrzených operujícím kardiologem. V každé EFV studii byl proveden výběr 1–5 nepřekrývajících se segmentů s orientační doporučenou délkou 10 s. Cílem výběru bylo maximalizovat klinické zastoupení variabilních typů poruch srdečního rytmu. Pro každý segment byl exportován 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.

### 1.3.3 Expertně anotovaná datová sada a (DS-A)

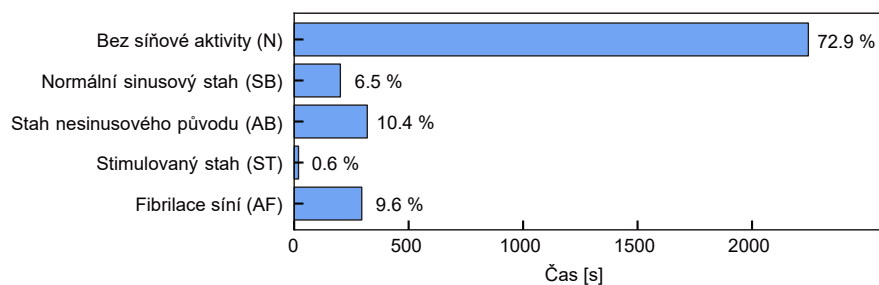
Účelem DS-A je podpora při učení systémů pro rozpoznání síňové aktivity v EGM záznamech. Do DS-A byly zařazeny EFV záznamy od 100 po sobě jdoucích pacientů z kohorty dětí a mladých dospělých na pracovišti FNB-PeK. Počet byl limitován časovou náročností manuální anotace. Z důvodu nedostupnosti dat v době tvorby DS-A nebyl zařazen žádný pacient středního věku. Celkem bylo vybráno 326 zájmových segmentů. Základní charakteristiku DS-A uvádíme v Tabulce 1.1.

Tab. 1.1: 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)

V DS-A byla provedena manuální anotace aktivačního času síňové aktivity v 5-kanálových bipolárních signálech snímaných v CS. Za počátek a konec síňové aktivity byl považován nejčasnější, resp. poslední lokální aktivační čas v rámci všech 5 kanálů. Výjimku tvořili záznamy s fibrilací síní, u kterých nelze přesný počátek a konec v rámci více svodů stanovit. Z tohoto důvodu byl označen počátek a konec samotné arytmie. V případě sloučených síňových a komorových stahů byly hranice síňové aktivity odhadnuty na základě morfologie elektrogramů při sinusovém rytmu.

Každý anotovaný úsek síňové aktivity byl zařazen do jedné ze 4 tříd: a) normální sinusový stah (SB); b) abnormální stah (AB); c) fibrilace síní (AF) a d) stimulovaný stah (ST). Časové zastoupení jednotlivých typů EGM uvádíme na Obrázku 1.2.



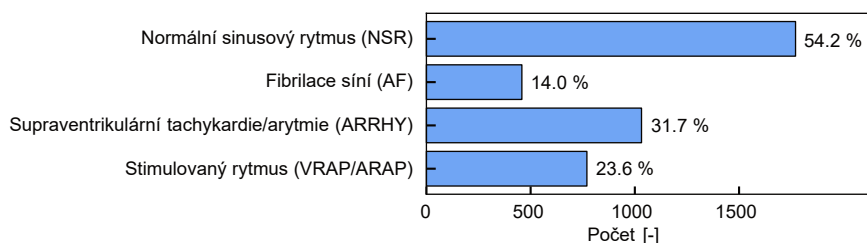
Obr. 1.2: Časové zastoupení jednotlivých typů síňových stahů v DS-A. Hodnoty vychází z referenčních úseků stanovených manuální anotací.

### 1.3.4 Automaticky anotovaná datová sada B (DS-B)

DS-B vznikla jako databáze arytmiických záznamů s anotovanou elektrickou aktivitou síní. Selekcce proběhla stejným způsobem jako v Sekci 1.3.3. Na rozdíl od DS-A byla síňová aktivita označena automaticky pomocí DNN modelu, který je popsán v Kapitole 2. Důvodem je sestavení obsáhlé datové sady s minimálními časovými náklady. Ze 2322 studií se podařilo do současné doby do DS-B zařadit celkem 708 pacientů a 3265 individuálních EKG a EGM segmentů. Jedná se o kohortu mladých dospělých i pacientů středního věku. Základní charakteristiku DS-B uvádíme v Tabulce 1.2. Zastoupení obecných diagnostických tříd v DS-B je uvedeno na Obrázku 1.3. Záznamy NSR mohou obsahovat lokalizované arytmie jako jsou síňové extrasystoly, které však nebyly zatím anotovány.

Tab. 1.2: 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)



Obr. 1.3: Zastoupení obecných diagnostických tříd v DS-B.

## 1.4 Diskuze a limitace

Datové sady obsahují záznamy s širokou škálou poruch srdečního rytmu a relativně vysoký počet individuálních subjektů. 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.

Přes nespornou informační výtěžnost trpí data několika nedostatky. Jedná o selektivní kohortu pacientů, která nereprezentuje skutečné zastoupení diagnóz v populaci. Algoritmy na nich učené proto mohou nadhodnocovat skutečné riziko arytmií u zdravé populace. Některé dokumentované nálezy nemají původ ve strukturálním postižení myokardu. Morfologické příznaky EKG i EGM signálu z toho důvodu nemusí zcela odpovídat nálezům u pacientů s primárním substrátem. Z technických důvodů se v průběhu EFV využívá odlišné anatomické umístění povrchových EKG elektrod než u abmulatorních záznamníků. To může být zdrojem morfologických změn v signálu s dopadem na elektrokardiografické markery některých srdečních onemocnění [24]. U DNN modelů je možné tento efekt do určité míry eliminovat doučením modelu na datech s požadovanou charakteristikou.

Některé limitace DS je třeba zohlednit při návrhu automatických algoritmů. EGM z oblasti CS představuje při sinusovém rytmu přibližně druhou polovinu trvání P vlny. V případě ostatních rytmů může být anotace umístěna jak v počáteční, tak konečné fázi síňového cyklu. Z pohledu detekce síňové aktivity představuje využití těchto intervalů jako referenčních anotací problém učení s neúplnou informací. S ohledem na zastoupení diagnostických tříd je zřejmé, že jsou obě datové sady zatíženy nerovnoměrným zastoupením vstupní a výstupní informace. Vstupní imbalance je dána krátkými časovými intervaly, které odpovídají aktivaci síní. 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 automatických algoritmů.

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.

## 2 Segmentace síňových elektrogramů pomocí modelů hlubokého učení

### 2.1 Současný stav

Mezi základní techniky identifikace arytmogenního substrátu při EFV patří měření lokální časové aktivace (LAT). Jedná se o počátek elektrické aktivace tkáně v oblasti snímací elektrody. Analýza LAT u multipolárních signálů umožňuje odvodit další parametry jako jsou rychlost a směr šíření aktivační vlny, post-stimulační intervaly a restituční křivky tkáně. Ty nesou informaci o vlastnostech myokardu a jsou základním vodítkem pro lokalizaci arytmogenních ložisek [25, 26]. Reprodukovatelné stanovení počátku LAT je základním požadavkem při vytváření aktivačních map pomocí elektroanatomického mapování.

Při unipolárním snímání je LAT spojován s časem minima první derivace lokálního EGM [27]. Limitací unipolárního záznamu je citlivost na vzdálené elektromagnetické pole odlehlé srdeční svaloviny. Bipolární snímání umožňuje tento efekt částečně potlačit, avšak za cenu zkreslení EGM díky prostorové geometrii a orientaci elektrod [28]. Derivační metoda je přesto využívána i v případě bipolárních EGM.

Morfologické metody jsou založeny na nalezení významného vzorku signálu jako je lokální maximum či barycentrum [29]. Tyto metody dosahují lepší reprodukovatelnosti ve srovnání s derivační metodou [29]. V obou případech je nutné předchozí vymezení segmentu s cílem eliminovat komplexy, které nejsou předmětem zájmu. Obvykle se jedná o manuální vymezení časového okna vázaného např. na komorový komplex v povrchovém EKG [30]. Méně časté je využití adaptivního okna na základě referenčního kanálu [29]. Automatické metody zahrnují základní filtrační založené na korelaci s bankou vzorů [31], nelineárních operátorech pro odhad okamžité energie [30], vlnkové transformaci [32] či derivaci a umocnění signálu [29]. Anotace LAT je následně provedena pomocí předdefinovaného napětového prahu.

Zdrojem selhání automatické segmentace bývá přítomnost stimulačních artefaktů či překrytí síňové a komorové aktivity při změně srdečního rytmu. Síňové a komorové EGM mohou v závislosti na vlastnostech tkáně vykazovat vysokou míru morfologické a spektrální variability. V takovém případě lze u současných metod sledovat neúměrné odstranění síňové složky nebo neschopnost diferencovat sdružený síňo-komorový komplex. Vizuální hodnocení LAT je proto stále nezbytnou součástí EFV.

### 2.2 Formální definice problému

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

## 2.3 Architektura modelu

Model byl založen na vlastní implementaci plně konvoluční neuronové sítě typu ResNet [33]. U modelu byly aplikovány dílčí modifikace reziduálního bloku RB. Vzhledem k dimenzionalitě vstupních dat byly původní dvourozměrné konvoluční masky nahrazeny jejich 1D ekvivalentem. Oproti původnímu řazení vrstev bylo využito schéma s tzv. předaktivací [33] a úzkým hrdlem. Za identickým zobrazením byla odstraněna výstupní aktivací vrstva, což umožňuje snížení rizika mizejícího gradientu a dle He et al. [34] nižší dosažené chyby. Úzké hrdlo efektivně snižuje počet parametrů. Tím je snížena i výpočetní náročnost konvolučních filtrů pro receptivní snímání.

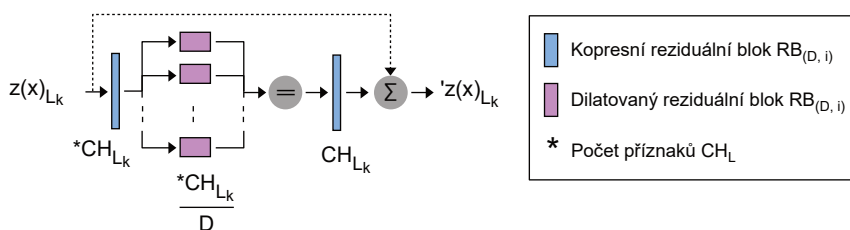
V dekodérové části byla použita hybridní struktura modelu vycházející z architektury U-net [35] a architektury pyramidových příznaků [36]. Převzorkování časové škály bylo na rozdíl od [35] provedeno pomocí lineární interpolace, která nevyžaduje dodatečné učení a umožňuje snížit výpočetní náročnost modelu. Oproti Lin et al. [36] nebylo využito nezávislé učení jednotlivých vrstev dekodéru, a to s ohledem na řídký výskyt referenčních segmentů v nejnižších úrovních hierarchie modelu.

### 2.3.1 Blok dilatovaných konvolucí

Fibrilace síní vykazuje vysokou míru neuspořádanosti a projevuje se jako polymorfní signál s dlouhou dobou trvání. Pro navýšení receptivního pole s ohledem na dlouhodobé variace signálu byla navržena vrstva reziduálních bloků s dilatovanými konvolučními maskami  $RB_D$ . Vrstva umožňuje exponenciálně navýšit receptivní oblast bez ztráty rozlišení a zvýšení komplexity modelu [37]. Přibližnou velikost receptivní oblasti filtru v ms lze definovat jako:

$$t_{rec} = \frac{1000}{f_s} \cdot (1 + d_i \cdot (f_{RB} - 1)), \quad (2.1)$$

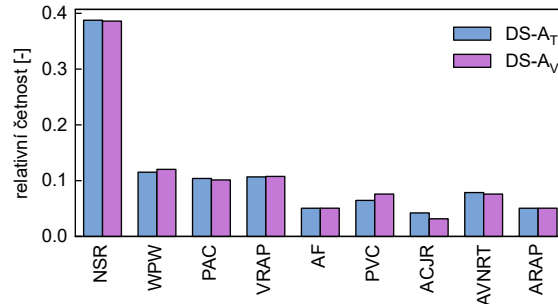
kde  $d_i$  je dilatační faktor ve vrstvě  $RB_{D,i}$ ,  $f_s$  je vzorkovací frekvence vstupního signálu a  $f_{RB}$  je počet aktivních prvků masky. Schéma vrstvy je uvedeno na Obrázku 2.1.



Obr. 2.1: Schéma bloku s dilatovanými konvolučními jádery.  $D$  – počet paralelních větví a kompresní faktor vstupních konvolučních masek;  $CH_{L_k}$  – počet příznaků vstupní mapy.

## 2.4 Stratifikace DS-A

Dataset byl pomocí metody hold-out rozdělen na trénovací DS-A<sub>T</sub> a validační DS-A<sub>V</sub> množinu v poměru 70:30. Data byla rozdělena pomocí stratifikované randomizace tak, aby každá sada obsahovala pouze unikátní pacienty. Cílem stratifikace bylo dosažení homogenity obou sad z pohledu zastoupení vybraných diagnostických tříd. Kvalita stratifikace byla stanovena na základě hodnoty relativní entropie a statistické shody diskrétního rozdělení (Pearsonův  $\chi^2$  test). Zastoupení jednotlivých diagnostických tříd v DS-A<sub>T</sub> a validační DS-A<sub>V</sub> po stratifikaci je uvedeno na Obrázku 2.2. Princip metody je popsán Algoritmem 2.



Obr. 2.2: Relativní rozložení diagnostických tříd v DS-A<sub>T</sub> a DS-A<sub>V</sub> 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<sub>T</sub>.
  - 4:     Zařaď zbylých 30 % záznamů do DS-A<sub>V</sub>.
  - 5:     Urči pravděpodobnostního rozložení  $P(x_T)$  a  $Q(x_V)$  diag. tříd v DS-A<sub>T</sub> a DS-A<sub>V</sub>.
  - 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**
-

## 2.5 Kriteriační funkce

V práci byla využita kriteriační funkce aproximující Tverskyho index (FTL) [38]. Z principu tato metrika nabývá totožné penalizace pro stejnou míru překryvu v případě rozměrných i malých objektů. Z tohoto vyplývá i její schopnost snížit vliv nerovnoměrného zastoupení klasifikačních tříd. Funkce obsahuje tři hyperparametry pro váhování vlivu falešně pozitivních, falešně negativních a obtížně rozlišitelných vzorků k celkové penalizaci:

$$\text{FTL} = \sum_n \left\{ 1 - \frac{\hat{Y} \odot T}{\hat{Y} \odot T + \alpha[(1 - \hat{Y}) \odot T] + \beta[\hat{Y} \odot (1 - T)]} \right\}^\gamma, \quad (2.2)$$

kde  $\odot$  značí Hadamardův součin tenzorů;  $\hat{Y}$  a  $T$  jsou tenzory představující surová skóre vrácená modelem, respektive binární referenční hodnoty;  $\alpha$  a  $\beta$  jsou váhy upravující vztah mezi Se a  $P^+$ ; exponent  $\gamma$  moduluje příspěvek snadno a obtížně trénovatelných vzorků k celkové hodnotě kriteriační funkce. Pro  $\alpha = \beta = 0.5$  and  $\gamma = 1$  se funkce zjednodušuje na Sørensen–Dice koeficient (SDC) [39]. To umožňuje nalezení optimálního nastavení v široké skupině kriteriačních funkcí běžně užívaných v oblasti sémantické segmentace.

## 2.6 Optimalizace modelu

Optimalizace modelu byla provedena pomocí algoritmu AdamW se separovanou  $L_2$  regularizací [40]. Optimální kombinace hyperparametrů s ohledem na kvalitu segmentace byla nalezena pomocí pravděpodobnostní optimalizace založené na odhadu Parzenových oken [41].

Optimalizace byla provedena nezávisle pro dvě skupiny hyperparametrů. První sada experimentů byla zaměřena na optimalizaci vybraných hyperparametrů architektury a procesu učení. Parametry FTL byly nastaveny konstantně na hodnotu  $\alpha_{\text{FTL}} = \beta_{\text{FTL}} = 0.5$  a  $\gamma_{\text{FTL}} = 1.0$ , což odpovídá aproximaci SDC. Výsledkem je segmentační model označený jako  $M_{\text{aopt}}$ . Druhá sada experimentů měla za úkol najít nejvhodnější parametry kriteriační funkce z rovnice 2.2 pro optimalizovaný model  $M_{\text{aopt}}$ . Výsledkem druhého kola optimalizace parametrů je model  $M_{\text{lopt}}$ . Optimální výchozí hodnoty hyperparametrů algoritmu AdamW jsou uvedeny v Tabulce 2.1.

Cílem druhé sady experimentů bylo nalezení optimálních parametrů penalizační funkce z Rovnice 2.2 pro architekturu  $M_{\text{aopt}}$ . Výstupem je model  $M_{\text{lopt}}$ . Prostor parametrů pro optimalizaci modelu  $M_{\text{aopt}}$  a parametrů FTL u  $M_{\text{lopt}}$  je uveden v Tabulce 2.2.

Tab. 2.1: Nastavení hyperparametrů algoritmu AdamW.

Hyperparametr	$\eta_0$	$\beta_1$	$\beta_2$	$\epsilon$
Hodnota	0.005	0.9	0.999	0

Tab. 2.2: Specifikace prostoru hyperparametrů modelu pro parametry architektury a FTL

Název hyperparametru	Symbol	Definiční obor
Výchozí krok učení	$\eta_0$	$\{1^{-5}, 5^{-5}, 1^{-4}, \dots, 1^{-1}\}$
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	$D_d$	$\{x \in \mathbb{N} \mid 2 \leq x \leq 5\}$
Váhování falešně negativních vzorků	$\alpha_{FTL}$	$\{x \in \mathbb{R} \mid 0.1 \leq x \leq 0.9\}$
Váhování falešně pozitivních vzorků	$\beta_{FTL}$	$x = 1 - \alpha_{FTL}$
Exponenciální váha celkové chyby	$\gamma_{FTL}$	$\{\frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4\}$

## 2.7 Hodnocení kvality modelu

V rámci hodnocení kvality modelu byl sledován 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. Pro významové odlišení spojitě aproximace SDC využité při optimalizaci modelu a samotné metriky kvality bude v rámci sekce výsledků uváděn alternativní název  $F_1$  skóre, které lze definovat jako:

$$F_1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}, \quad (2.3)$$

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

Dále jsou reportovány senzitivita (Se) a pozitivní prediktivní hodnota ( $P^+$ ), které lze na základě správnosti zařazení vzorků definovat následovně:

$$Se = \frac{TP}{TP + FN}, \quad (2.4)$$

$$P^+ = \frac{TP}{TP + FP}, \quad (2.5)$$



Pro hodnocení lokálních aktivačních časů byl každý spojitý segment stanovený modelem převeden na individuální instanci. Přesnost stanovení pozic počátku a konce síňové aktivity byla hodnocena na základě absolutní (AE) a střední kvadratické chyby (RMSE):

$$AE = |\hat{y}_i - y_i|, \quad (2.6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}, \quad (2.7)$$

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 vyskytnou v tolerančním intervalu  $\tau_{tol} = 35$  ms v okolí reference, což odpovídá dvojnásobku směrodatné odchylky délky síňových EGM v DS-A.

Pro hodnocení nespojitosti nalezených oblastí byla zavedena doplňková metrika založená na celkové procentuální odchylce (PE) mezi počtem referenčních anotací a modelem identifikovaných instancí:

$$PE = 100 \cdot \frac{(P_t - A_t)}{A_t}, \quad (2.8)$$

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.

## 2.8 Výsledky

### 2.8.1 Optimalizace architektury

Výsledky pro trojici nastavení, při kterém bylo dosaženo nejlepší hodnoty metriky  $F_1$ , jsou uvedeny v Tabulkách 2.3 a 2.4. Všechny tři varianty dosahují obdobné hodnoty  $F_1$  skóre okolo 0.887, a to pro relativně široký rozsah nastavení a komplexnost modelu, viz Tabulka 2.4. Na Obrázku ?? jsou pak vyobrazeny obrysové grafy závislostí mezi dvojicí hyperparametrů a výsledného  $F_1$  skóre.

Tab. 2.3: Modely s nejvyšším dosaženým  $F_1$  skóre po Bayesovské optimalizaci architektury. DS-A<sub>T</sub> – trénovací sada; DS-A<sub>V</sub> – validační sada; Exp. – pořadí optimalizačního kroku.

Exp.	DS-A <sub>T</sub>			DS-A <sub>V</sub>		
	F <sub>1</sub>	Se	P <sup>+</sup>	F <sub>1</sub>	Se	P <sup>+</sup>
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. 2.4: Hyperparametry modelů s nejvyšším dosaženým  $F_1$  skóre po Bayesovské optimalizaci architektury.

Exp.	Hyperparametry								Počet p. modelu
	$\eta_0$	$BS_T$	$f_{RB}$	$f_{L0}$	$K_e$	$N_{RB}$	$CH_C$	$D_d$	
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

Největší vliv na dosažené  $F_1$  měl výchozí krok učení  $\eta_0$ . Při zohlednění pouze hyperparametrů architektury, se nejvýznamněji (relativní skóre 0.65) podílel počet v sérii zařazených vrstev enkodéru  $K_e$ . Jedná se o parametr, který nejvíce ovlivňuje kapacitu modelu extrahovat relevantní příznaky. Současně však navyšuje počet trénovatelných koeficientů a tím i výpočetní náročnost modelu, což je patrné při srovnání variant s  $K_e = 5$  a  $K_e = 6$  vrstvami. Z tohoto důvodu byla za optimální model  $M_{aopt}$  vybrán experiment č. 70.

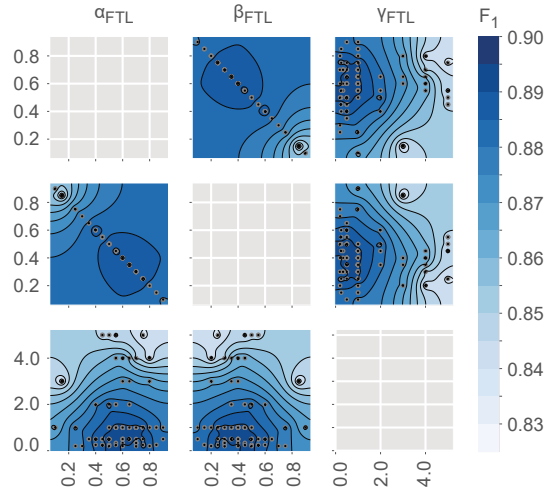
Mezi další významné parametry patří počet filtrů  $CH_C$  v dekodérové části. Při navýšení nad hodnotu 100 filtrů docházelo k systematickému poklesu  $F_1$ . Model tedy více těžší z komprimované reprezentace příznakových map než z přeparametrizování. Optimální hodnota se pohybovala v oblasti 32–48 filtrů. To představuje přibližně průměr počtu filtrů v jednotlivých vrstvách. Z pohledu konstrukce reziduálních bloků jsou preferovány varianty s užší receptivní oblastí filtrů  $f_{RB}$  v rozsahu 5–7 vzorků, avšak s navýšením počtu jejich sériového řazení  $N_{RB}$ .

## 2.8.2 Optimalizace kriteriální funkce

V Tabulce 2.5 uvádíme výsledky 3 nejlepších modelů po optimalizaci FTL. Nejlepší dosažené skóre se pohybuje v okolí hodnoty 0.888, a to i pro hyperparametry  $\alpha_{FTL} = \beta_{FTL} = 0.5$  a  $\gamma_{FTL} = 1$ . Zásadní vliv na skóre se projevil u exponenciální váhy  $\gamma_{FTL}$  s optimálním nastavením v rozmezí 0.25–1.00. Z Obrázku 2.3 je patrné, že jsou více penalizovány obtížně klasifikovatelné vzorky signálu. V průměrném případě naopak model dosahoval dostatečné míry jistoty na to, aby koeficient  $\gamma_{FTL}$  mohl snížit jejich příspěvek k penalizaci.

Tab. 2.5: Hyperparametry kriteriální funkce FTL získané pomocí Bayesovské 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	$P^+$	$\alpha_{FTL}$	$\beta_{FTL}$	$\gamma_{FTL}$
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

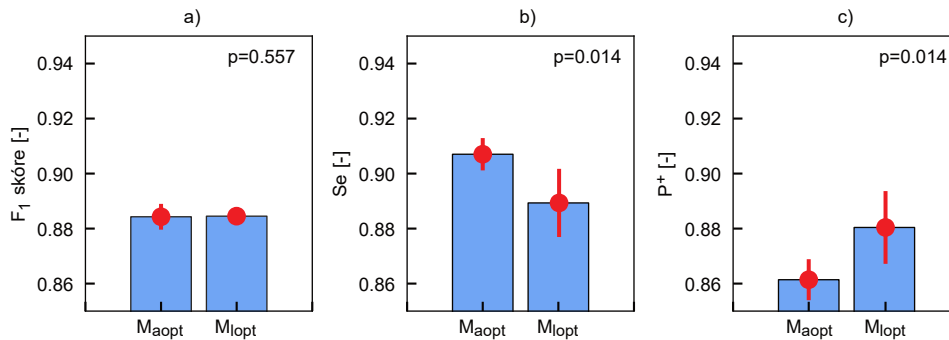


Obr. 2.3: Obrysové grafy znázorňující vztah mezi hyperparametry kritériální funkce  $\alpha_{\text{FTL}}$ ,  $\beta_{\text{FTL}}$  a  $\gamma_{\text{FTL}}$  a  $F_1$  skóre na DS- $A_V$ . Pořadí řádků v matici grafů odpovídá pořadí sloupců.

Pro výchozí hodnoty  $\alpha_{\text{FTL}} = \beta_{\text{FTL}} = 0.5$  dosahuje model významně vyšší hodnoty Se (0.902) na úkor chybné klasifikace negativní třídy. Zvýšením parametru  $\beta_{\text{FTL}}$  dochází ke změně poměru mezi oběma metrikami směrem k vyšším hodnotám  $P^+$ . Konvenčně je podobných cílů dosahováno pomocí změny rozhodovacího prahu modelu. Tato metoda u modelů  $M_{\text{aopt}}$  i  $M_{\text{llopt}}$  nemá v rozmezí hodnot 0.1–0.9 zásadní vliv ani na jednu z metrik.

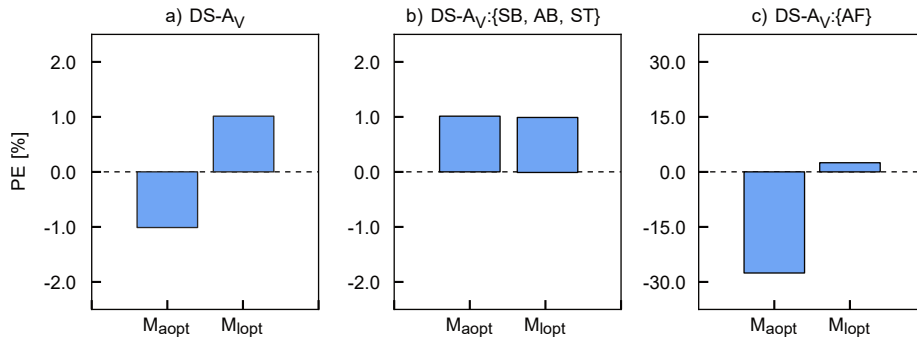
### 2.8.3 Srovnání modelů $M_{\text{aopt}}$ a $M_{\text{llopt}}$

Optimalizace kritériální funkce nemá významný vliv na hodnotu  $F_1$  ( $p=0.557$ ). Naopak dosažené skóre v ostatních metrikách se pro oba modely významně liší.  $M_{\text{aopt}}$  dosahuje průměrně o 0.021 bodů vyšší Se ( $p=0.014$ ), avšak za cenu systematického snížení  $P^+$  v průměru o 0.019 bodů ( $p=0.014$ ). Optimalizovaná FTL poskytuje více vyvážený poměr mezi Se a  $P^+$ . Zvýšená míra rozptylu u modelu  $M_{\text{llopt}}$  u Se a  $P^+$  byl pravděpodobně způsoben nižší numerickou stabilitou výpočtu při  $\gamma_{\text{FTL}} \neq 1$ .



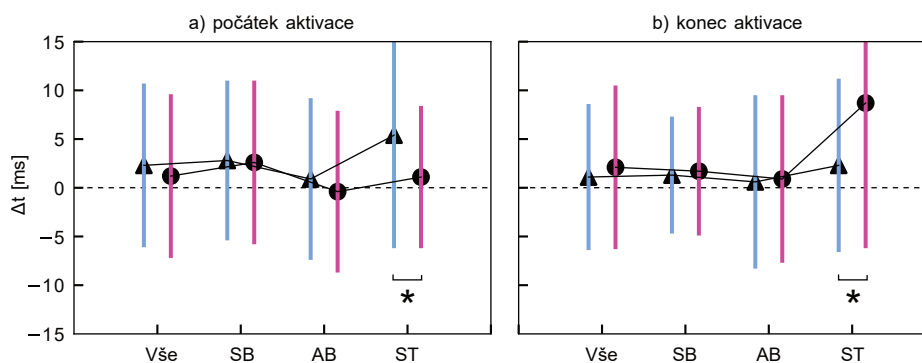
Obr. 2.4: Srovnání metrik kvality (a)  $F_1$  (b) Se a (c)  $P^+$  na DS- $A_V$  pro dvojici optimalizovaných modelů  $M_{\text{aopt}}$  a  $M_{\text{llopt}}$ .

Na Obrázku 2.5 je srovnána kontinuita klasifikovaných segmentů pro oba modely (blíže nule=lepší). V případě celé SB, AB a ST dosahují modely prakticky shodné chyby kolem 1 %. Zhoršení kvality segmentace až o 27.3 % lze spatřit v případě záznamů s fibrilací síní u modelu  $M_{\text{aopt}}$ , Obrázek 2.5 (c).



Obr. 2.5: Srovnání PE na (a) kompletní DS-A<sub>V</sub>, (b) záznamech bez fibrilace síní a (c) pouze záznamech s fibrilací síní pro dvojici optimalizovaných modelů  $M_{\text{aopt}}$  a  $M_{\text{lopt}}$ . Bodové odhady provedeny pro nejlepší modely z obou typů.

Chyba stanovení hranic segmentů je zobrazena na Obrázku 2.6. 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, kde docházelo k systematickému nadhodnocení pozice. Z pohledu distribuce absolutních chyb bylo v případě  $M_{\text{lopt}}$  celkově 54.0 % a 81.0 % odchylek  $\leq 5$  ms, respektive  $\leq 10$  ms. U modelu  $M_{\text{aopt}}$  bylo pozorováno mírné zhoršení s 50.2 % absolutních odchylek v kategorii  $\leq 5$  ms a 80.6 % odchylek bylo menších než 10 ms.



Obr. 2.6: Srovnání přesnosti stanovení počátku a konce síňové aktivity typu SB, AB a ST pro dvojici optimalizovaných modelů  $M_{\text{aopt}}$  (plný kruh s purpurovými vousy) a  $M_{\text{lopt}}$  (plný trojúhelník s modrými vousy). Symbol \* značí statisticky významný rozdíl mezi skupinami.

## 2.8.4 Segmentace fibrilačních záznamů

Testování bylo na základě předchozích výsledků provedeno pro model  $M_{\text{lopt}}$  pomocí ablační studie s 3 variantami: a) referenční model; b) model bez dilatovaných konvolučních masek (experiment  $M_{\text{lopt}}$  w/o dil); c) model učený bez přítomnosti anotací z kategorie AF (experiment  $M_{\text{lopt}}$  w/o AF). Výsledky pro podmnožinu  $DS-A_V:\{SB, AB, ST\}$  a podmnožinu fibrilačních záznamů  $DS-A_V:\{AF\}$  jsou uvedeny v Tabulce 2.6.

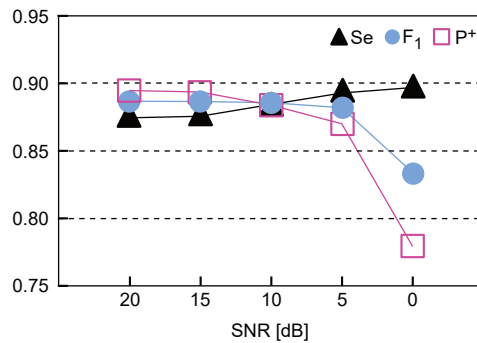
Tab. 2.6: Výsledky ablační studie u modelu  $M_{\text{lopt}}$  pro různé AF scénáře.

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

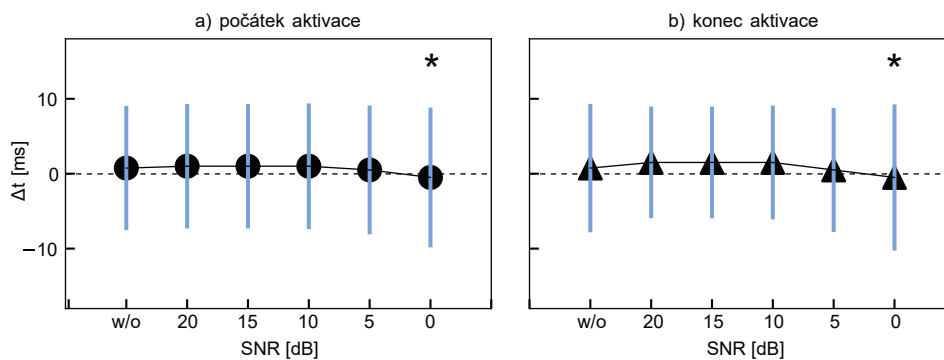
## 2.8.5 Hodnocení vlivu šumu

V této sekci jsou prezentovány výsledky segmentace pro tři typy uměle generovaného rušení s hladinou poměru signál-šum (SNR) v rozmezí 20–0 dB. Testované varianty byly: a) aditivní bílý šum Gaussovského typu; b) 50 Hz síťový brum; c) nulování kanálů. Kritickou hladinou se ve většině případů ukázal SNR 5 dB. V případě bílého šumu byla příčinou zvýšené falešné pozitivivity tendence modelu interpretovat šum jako fibrilační EGM.

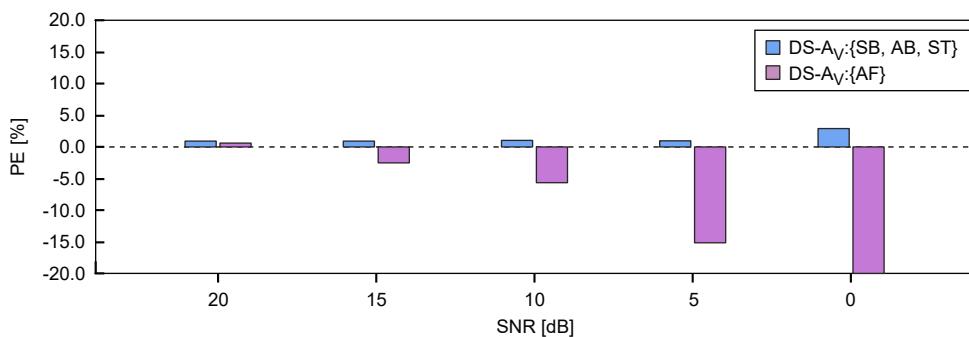
Bílý šum má negativní dopad na segmentaci izolovaných síňových stahů, které se nachází v záznamech s přechodem fibrilace síní do jiného, více organizovaného rytmu. Od hladiny 10 dB dochází k poklesu počtu identifikovaných EGM komplexů, a to o 5.8 % až o výsledných 27.3 % v případě 0 dB, viz Obrázek 2.9. Takový výsledek naznačuje buď úplné selhání identifikace izolovaných komplexů nebo naopak jejich zahrnutí do AF segmentu.



Obr. 2.7: Srovnání  $F_1$ , Se a  $P^+$  v závislosti na hladině bílého šumu u DS- $A_V$ .

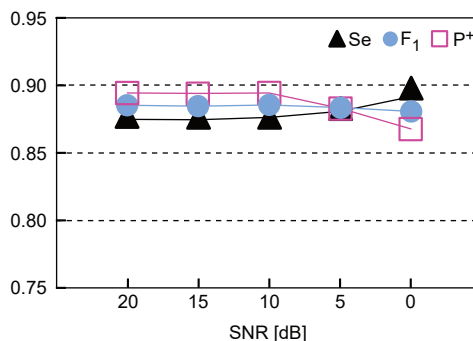


Obr. 2.8: Vliv hladiny bílého šumu na přesnost nalezení počátku a konce síňové aktivity v celé DS- $A_V$ . Zkratka w/o označuje referenční model na signálech bez přidaného šumu. Symbol \* značí statisticky významný rozdíl oproti w/o.

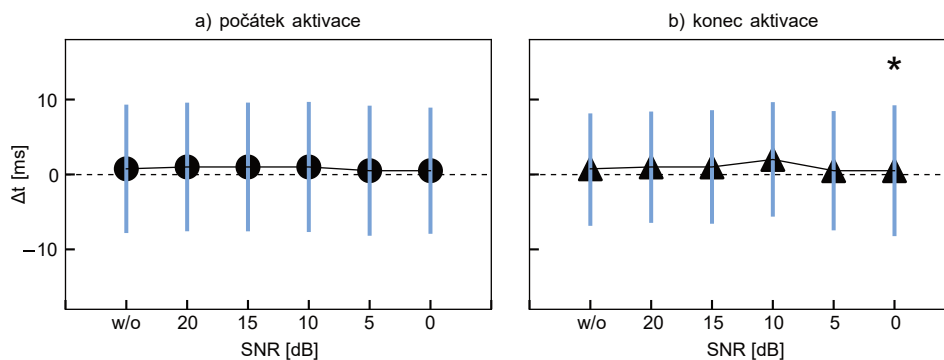


Obr. 2.9: Vliv hladiny bílého šumu na přesnost určení počtu izolovaných EGM.

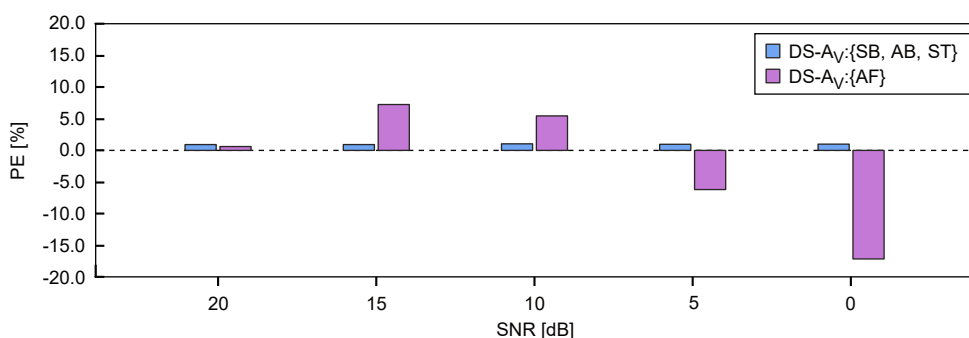
V případě síňového brumu byl zaznamenán mírný pokles kvality segmentace od hranice SNR 5 dB. Ovlivněna je primárně  $P^+$ , která při 0 dB klesla až o 0.024 bodů. Nárůst senzitivity byl přítomen primárně u záznamů s fibrilací síní, a to z hodnoty 0.877 na 0.930. Přesnost nalezení hranic v případě správné identifikace síňového komplexu zůstala relativně zachována i při hladině šumu 0 dB, viz Obrázek 2.11.



Obr. 2.10: Srovnání  $F_1$ , Se a  $P^+$  v závislosti na hladině síňového brumu u DS- $A_V$ .

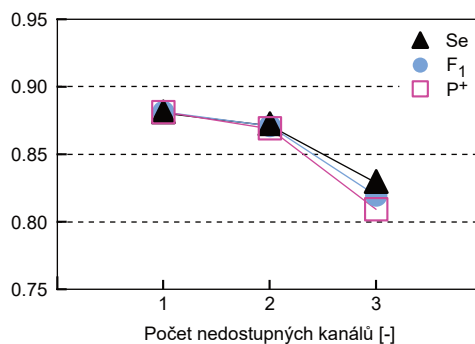


Obr. 2.11: Vliv hladiny síťového brumu na přesnost nalezení počátku a konce síňové aktivity v celé DS- $A_V$ . Zkratka w/o označuje referenční model na signálech bez šumu.



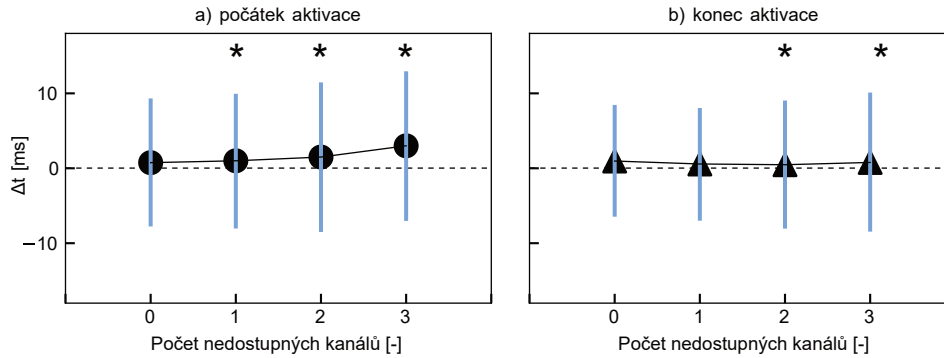
Obr. 2.12: Vliv hladiny síťového brumu na přesnost stanovení počtu izolovaných EGM.

V případě nulování kanálů došlo systematicky k poklesu kvality segmentace již při odstranění 1 z kanálů, jak je patrné z Obrázků 2.13 a 2.14. Ovlivněna byla i přesnost stanovení hranic jednotlivých komplexů. Hodnota RMSE postupně narostla 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. Selektce nulovaného kanálu pro každý záznam byla provedena náhodně.



Obr. 2.13: Srovnání  $F_1$ , Se a  $P^+$  pro model  $M_{1opt}$  v závislosti na počtu dostupných kanálů.

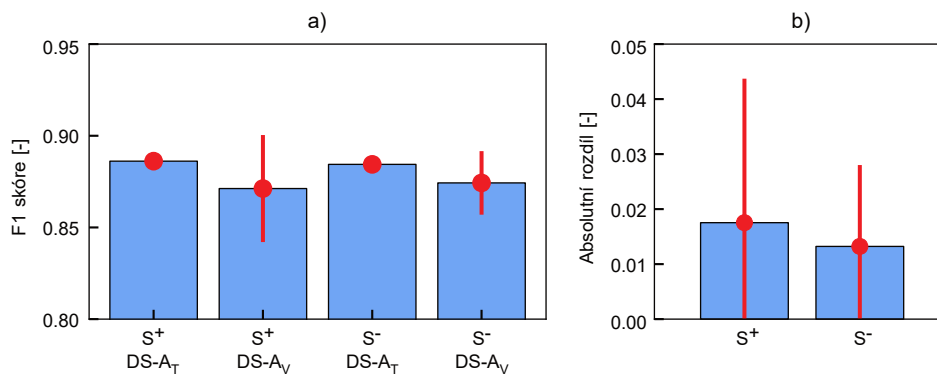
Rozdíly v rozptylu reziduí oproti segmentaci záznamů bez odstraněných kanálů byly statisticky významné již po odstranění prvního kanálu v případě počátku ( $p=0.011$ , následně  $p<0.0001$ ), a od druhého kanálů u stanovení konce (v obou případech  $p<0.0001$ ). Graf (a) na Obrázku 2.14 dále ukazuje rostoucí nadhodnocení detekce počátku v průměru až o 3.6 ms v nejhorsím případě. Hladina statistické významnosti  $\alpha = 0.0166$  byla korigována dle Bonferonniho.



Obr. 2.14: Vliv počtu kanálů na stanovení (a) počátku a (b) konce síňové aktivity pro model  $M_{1opt}$  v DS- $A_V$ . \* značí statisticky významný rozdíl proti výchozímu stavu.

## 2.8.6 Vliv stratifikace datové sady

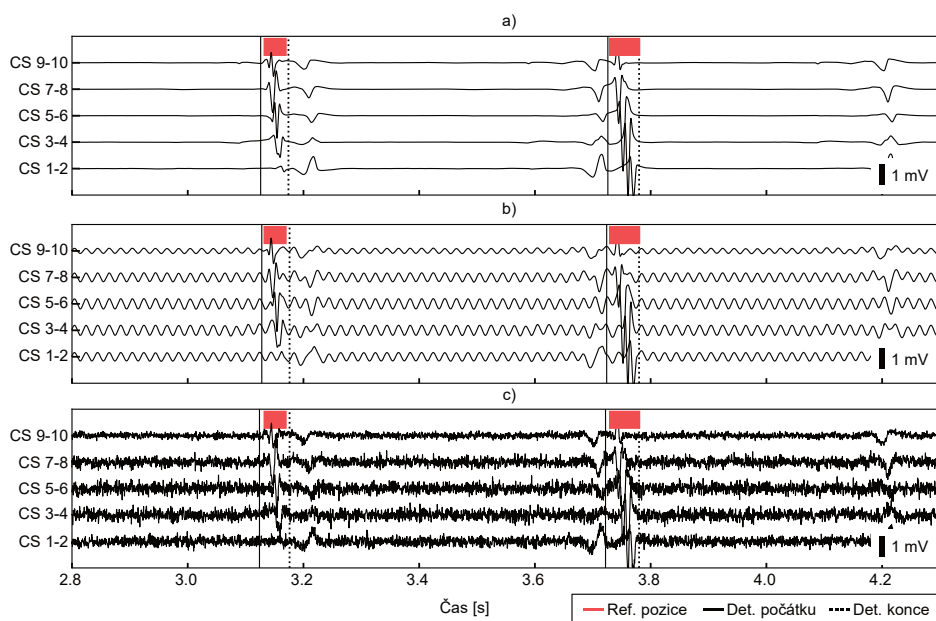
Náhodného rozdělení bez stratifikace vedlo jen v 24.4 % případů k vygenerování dvou sad se obdobným rozložením ( $S^-$ ) diagnostických tříd (Pearsonův  $\chi^2$ ,  $p=0.994$ ). Trénování s nehomogenními sety  $S^+$  vedlo k významnému zvýšení rozptylu  $F_1$  skóre (F-test,  $p<0.001$ ) na DS- $A_T$  i DS- $A_V$ , což snižuje věrohodnost výsledků. Po stratifikaci došlo ke snížení rozptylu  $F_1$  i rozdílů mezi DS- $A_T$  a validační DS- $A_V$ , byť na hranici statistické významnosti (Wilcoxonův test,  $p=0.022$ ). Vliv stratifikace je zobrazen na Obrázku 2.15.



Obr. 2.15: Porovnání  $F_1$  pro stratifikované a nestratifikované rozdělení do DS- $A_T$  a DS- $A_V$ .



## 2.8.7 Ukázka segmentace



Obr. 2.16: Ukázka segmentace modelem  $M_{\text{lopt}}$  v případě a) bez přidaného šumu; b) generovaný střídavý signál o frekvenci 50 Hz, SNR = 0 dB; c) generovaný bílý šum, SNR=0 dB.

## 2.9 Diskuze a závěr

U elektrofyziologických systémů určených pro akvizici a analýzu intrakardiálních dat stále dominují anotační nástroje založené na jednoduchých filtračních technikách a manuálním měření. V této části disertační práce byl realizován návrh modelů hlubokého učení pro segmentaci síňových EGM. Autoři El Haddad et al. [29] pro měření LAT definují tři úrovně přesnosti: a) vysoce přesné s chybou do  $\pm 5$  ms; b) přesné s chybou do  $\pm 10$  ms a c) nepřesné s chybou nad  $\pm 10$  ms. V tomto kontextu dokáže model  $M_{\text{lopt}}$  ve vícesvodovém EGM záznamu určit přesnost počátku či konce aktivace s chybou  $< 5$  ms v 54 %. V 81 % případů dokáže stanovit hranici segmentu odchylkou  $< 10$  ms.

Specifickou roli ve vývoji modelu hrály záznamy s fibrilací síní a způsob jejich anotace. Anotace AF v podobě kontinuálního segmentu významně komplikuje optimalizační cíle. V případě zařazení AF do negativní třídy naproti tomu došlo ke snížení všech metrik kvality u AF záznamů i izolovaných síňových stahů. Variabilní morfologie fibrilačních EGM by proto mohla přispívat k lepší regularizaci modelu. Rozšíření receptivní oblasti sice nepřispělo k identifikaci fibrilačních vln, ale umožnilo přesnější odlišení mezi organizovanou a neorganizovanou síňovou aktivitou v oblasti přechodu do jiného rytmu.

Dle vědomí autora jde o první práci zaměřenou na segmentaci EGM pomocí DNN modelů. Výstup byl navržen jako součást řetězce pro automatizované zpracování EFV dat. Získané poznatky mohou pomoci při vývoji modelů pro klinické využití.

## 3 Slabě dozorovaná segmentace P vln

### 3.1 Současný stav

Přesnost detekce a segmentace P vln se významně snižuje s výskytem patofyziologických rytmů v EKG signálu. Metody pro identifikaci síňové aktivity v EKG záznamu lze rozdělit do tří základních kategorií. První skupina je založena na filtračních technikách pro zvýraznění P vlny. Mezi typické reprezentanty patří metody založené na vlnkové transformaci [12, 42, 43], empirickém rozkladu [44], fázorové transformaci [45, 46] či aproximaci parametrických modelů [47]. Následně je provedeno prohledávání v oblasti mezi koncem T vlny a QRS komplexem. Předpoklad výskytu P vlny v oblasti před QRS komplexem diskvalifikuje tento přístup v analýze značné části patologických signálů.

Druhá skupina metod je založena na potlačení komorové aktivity. To může být provedeno buď naivní subtrakcí QRS-T komplexu z původního signálu [48] nebo s využitím analýzy nezávislých komponent (ICA) v případě vícesvodových záznamů [49]. Subtrakce QRS-T je z principu citlivá na chybnou detekci QRS a nestacionární charakter signálu. ICA nalézá uplatnění primárně v analýze fibrilačních vln [48], kde je možné naplnit statistické předpoklady metody.

Do poslední kategorie spadají metody využívající hluboké učení. Typicky jsou využívány konvoluční síť U-Net [50, 51] a rekurentní neuronové sítě [52, 13]. Kvalita DNN je obecně odvozena od kvality použité datové sady, kterých je oblastí segmentace P vln nedostatek. DNN modely stále nesou memento černé skříňky, jejich rozhodovací proces je obtížné interpretovat.

### 3.2 Formulace problému segmentace s neúplnou informací

V případě učení s neúplnou informací jsou referenční hodnoty poskytnuty pouze pro: a) malou podmnožinu objektů; b) malou podmnožinu vzorků reprezentujících objekt. Formálně je cílem nalézt transformační funkci  $g(x)$  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 vzorů,  $j$  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.

V této části práce byla využita technika víceinstančního učení (MIL<sup>1</sup>), u které není vyžadována klasifikace individuálních instancí. Místo toho jsou instance sdruženy do množin a referenční třída je následně přiřazena celé množině. Standardním předpokladem je označení množiny pozitivní klasifikační třídou, pokud do ní spadá alespoň jedna zařazená instance [53].

---

<sup>1</sup>Z anglického Multiple-instance learning

### 3.3 Architektura modelu

Vstupní část modelu byla tvořena 50-vrstvým reziduálním modelem (ResNet-50) se shodnou implementací jako v Kapitole 2. Hodnoty 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 $L_0$	$f_{L_0}$	17
Počet vrstev v enkodéru	$K_e$	4
Počet RB ve vrstvách	$N_{RB}$	{9, 12, 18, 9}
Počet filtrů v RB	$PRB$	{128, 256, 512, 1024}
Počet filtrů dekodéru	$CH_C$	256

#### 3.3.1 Regionální příznakové mapy

Příznakové mapy  $z(x)_{L_k}$  jsou vyvedeny z prvních tří hierarchických vrstev dekodéru. Jedná se o časově agregované příznaky, které 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  $z(x)_{L_k}$  jsou následně rozděleny na ekvidistantní regiony o konstantní šířce. Každý z nich reprezentuje jednu množinu instancí a je mu přidělena anotace na základě předpokladů MIL zmíněných v předchozím textu. Šířka jednotlivých segmentů byla empiricky stanovena na 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).

### 3.4 Normalizovaná křížová entropie pro kontrastní učení

V rámci této části byla adoptována penalizační funkce *SimCLR* z práce autorů Khosla et al. [54]. Princip optimalizace pomocí *SimCLR* spočívá v maximalizaci mezishlukové podobnosti pro dvě odlišné klasifikační třídy. Pro shlukování jednotlivých úseků EKG byla využita kosinová podobnost:

$$\text{sim}(z_i, z_k) = \frac{z_i^T \cdot z_k}{\|z_i\| \cdot \|z_k\|}, \quad (3.1)$$

kde proměnné  $z_i$  a  $z_k$  značí dvojici příznakových vektorů a symbol  $\|\cdot\|$  reprezentuje jejich  $L^2$  vektorovou normu.

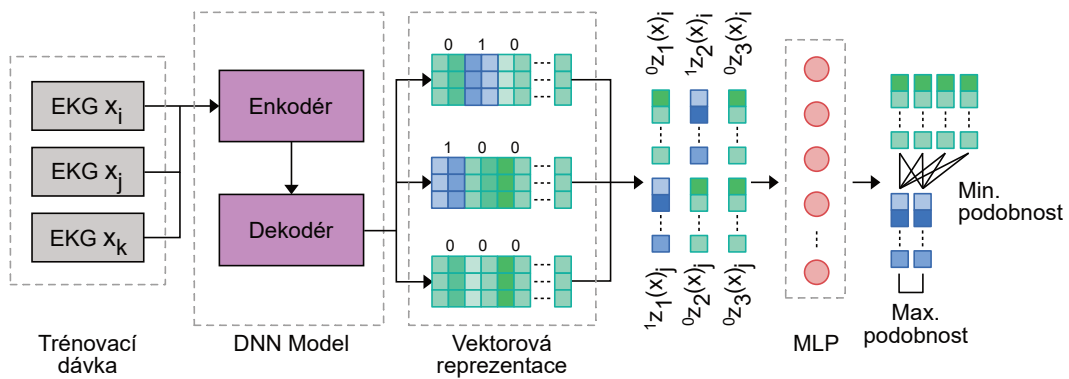
Penalizace pro dvojici vektorů s přítomností  $P$  vlny je následně určena pomocí standardní křížové entropie:

$$\text{SimCLR} = -\log \frac{e^{\text{sim}(z_i, z_j)/\tau_S}}{\sum_{k=1}^{2N} \mathbb{P}[i \neq k] e^{\text{sim}(z_i, z_k)/\tau_S}}. \quad (3.2)$$

Parametr  $p \in \{0, 1\}$  nabývá hodnoty 1 pokud  $i \neq k$ , čímž je v děliteli zanedbána sobě-podobnost vektoru  $z_i$ . Koefficient  $\tau_S$  slouží k váhování vlivu obtížně klasifikovatelných segmentů z negativní třídy na penalizaci.

### 3.4.1 Kontrastní učení v lokalizované oblasti signálu

Trénovací vzory jsou v našem případě tvořeny individuálními EKG záznamy. Tato situace je v rámci principu *SimCLR* ilustrována na Obrázku 3.1 (a). Anotované úseky síňové aktivity jsou ovšem lokalizovány do mnoha relativně úzce vymezených oblastí signálu. Proto byla navržena modifikace *SimCLR* tak, aby ke stanovení podobnosti docházelo mezi lokalizovanými oblastmi EKG signálu. Základní princip je znázorněn na Obrázku 3.1 (b).



Obr. 3.1: 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. Instance jsou na rozdíl od konvenčního MIL tvořeny vzorky hluboké příznakové mapy podobně jako v práci [55]. Před výpočtem penalizace v *SimCLR* byla provedena nelineární transformace

lokální vektorové reprezentace pomocí série plně propojených vrstev (MLP<sup>2</sup>) a ReLU aktivizační funkce.

### 3.5 MIL vrstva

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 [56]. Agregace je provedena na normalizované mapě a lze ji formálně vyjádřit jako:

$$z_{ij} = \max_{i=\{k,\dots,s\},j=\{0,\dots,m\}} \left( \frac{1}{1 + e^{-z_{i,j}}} \right) \quad (3.3)$$

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  $\tau_i$ .

Pro výstupní vrstvu  $L_{\text{end}}$  byl rozsah  $\tau_i$  stanoven na konstantní hodnotu 160 ms, aby bylo možné potenciálně pokrýt celou šířku  $P$  vlny sinusového původu.

### 3.6 Popis experimentů

Ověření navržených metod bylo provedeno pro 3 různé způsoby učení, jejichž princip je schématicky znázorněn na Obrázku ?? na následující straně. Jemná segmentace byla u všech variant provedena za výstupní vrstvou dekodéru  $L_{\text{end}}$ , kde dochází k agregaci příznaků pomocí MIL vrstvy. Penalizace MIL příznakových map byla realizována pomocí váhované verze binární křížové entropie (BCE<sup>3</sup>). Nejjednodušší scénář v (a) odpovídal standardnímu scénáři s učením přes finální klasifikační vrstvu. Výsledkem byl model  $M_{\text{MIL}}$ . 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  $M_{\text{SIM}}$  učeny s využitím kontrastní křížové entropie (*SimCLR*) aplikované na lokální oblasti příznaků.

### 3.7 Trénovací data

Pro účely této části práce je využita datová sada DS-B. Z dostupných záznamů bylo vybráno 8 nezávislých kanálů: a) dva bipolární povrchové svody I, II a b) unipolární hrudní svody V1–V6. Dataset byl rozdělen na trénovací DS-B<sub>T</sub> a validační DS-B<sub>V</sub> množinu v poměru 70:30 metodou popsanou v Sekci 2.4.

<sup>2</sup>Z anglického Multi-layer Perceptron

<sup>3</sup>Z anglického Binary Cross Entropy

### 3.8 Předzpracování dat

Signál byl podvzorkován z 2000 Hz na 250 Hz pomocí decimace a anti-aliasingového FIR filtru. Dále byly aplikovány FIR filtr typu horní a dolní propust s mezní frekvencí 0.5 Hz, respektive 40 Hz [57]. Standardizace napěťové škály byla provedena do jednotkového mezi-percentilového rozpětí, a to v rozsahu napětí  $[-1.90, 1.90]$  mV. Augmentační techniky byly obdobné jako v Kapitole 2. Záměna dvou po sobě jdoucích kanálů  $T_{\text{swap}}$  byla aplikována oproti EGM pouze na hrudní svody V1-V6. Do řetězce byl navíc přidán blok pro náhodné převrácení časové posloupnosti EKG signálu.

### 3.9 Hodnocení kvality segmentace

U prahované příznakové mapy z výstupní klasifikační vrstvy  $L_{\text{end}}$  byla hodnocena Se a  $P^+$  z rovnic 2.4 a 2.5. Toleranční pásmo správné detekce  $\tau_{\text{tol}} = 60$  bylo v tomto případě nastaveno tak, aby byl zohledněn vztah mezi reálnou šířkou P vlny a referencí. U vrstvy MIL byla hodnocena úspěšnost zařazení celého lokálního agregovaného segmentu do klasifikační třídy, a to na základě  $F_1$  skóre uvedeného v Rovnici 2.3.

### 3.10 Výsledky

V Tabulce 3.2 jsou uvedeny výsledky segmentace pro výstupní klasifikační vrstvu modelu.  $F_1$  skóre reprezentuje kvalitu klasifikace po agregaci příznaků pomocí MIL.  $F_1$  přesáhlo pro všechny tři varianty hodnotu 0.8. Nejlepšího výsledku na validačních datech  $F_1 = 0.819$  bylo dosaženo u kontrastního modelu  $M_{\text{SIM}}$ .

Tab. 3.2: Výsledky segmentace po MIL agregaci příznaků z klasifikační vrstvy  $L_{\text{end}}$ . DS- $A_T$  – trénovací sada; DS- $A_V$  – validační sada.

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

V Tabulce 3.3 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.  $P_{\text{on}}$  značí počátek P vlny,  $P_{\text{end}}$  její konec a  $P_{\text{peak}}$  reprezentuje hypotetický vrchol.  $P_{\text{peak}}$  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. 3.3: 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 <sub>T</sub>			DS-B <sub>V</sub>		
Model		P <sub>on</sub>	P <sub>peak</sub>	P <sub>end</sub>	P <sub>on</sub>	P <sub>peak</sub>	P <sub>end</sub>
M <sub>MIL</sub>	Se	0.66	0.67	0.68	0.59	0.60	0.59
	P <sup>+</sup>	<b>0.71</b>	<b>0.73</b>	<b>0.71</b>	0.62	0.62	0.60
M <sub>3MIL</sub>	Se	0.69	0.71	0.71	0.63	0.63	0.64
	P <sup>+</sup>	<b>0.79</b>	<b>0.81</b>	<b>0.78</b>	0.69	0.69	0.66
M <sub>SIM</sub>	Se	0.78	0.79	0.78	<b>0.71</b>	<b>0.72</b>	<b>0.69</b>
	P <sup>+</sup>	<b>0.83</b>	<b>0.87</b>	<b>0.85</b>	<b>0.80</b>	<b>0.82</b>	<b>0.80</b>

Modely M<sub>MIL</sub> i M<sub>3MIL</sub> dosáhly v případě P<sub>peak</sub> na celé DS-B<sub>V</sub> podprůměrných výsledků se sníženými hodnotami Se i P<sup>+</sup> do rozsahu 0.60–0.63, respektive 0.62–0.69. Při srovnání vedla navržená metoda kontrastního učení k významnému zlepšení výsledků. Model M<sub>SIM</sub> dosáhl ve srovnání s M<sub>MIL</sub> a M<sub>3MIL</sub> zlepšení senzitivity o 12 %, respektive 9 %. K významnému nárůstu došlo u pozitivní prediktivní hodnoty, a to o 20 %, respektive 13 %.

V Tabulce 3.4 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 tedy vybrán pro demonstraci chování modelů na datové sadě, která se složením nejvíce blíží standardním databázím.

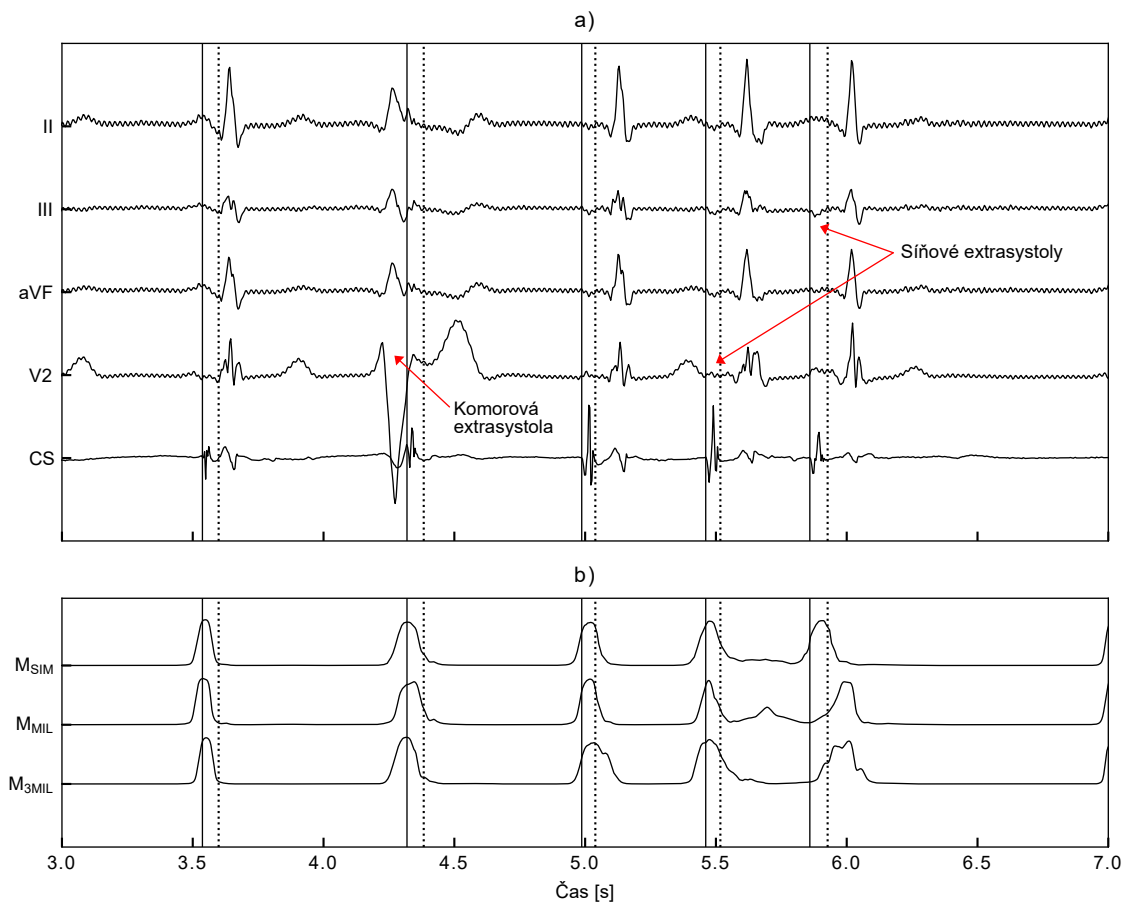
Tab. 3.4: Hodnoty Se a  $P^+$  významných bodů po převodu příznakové mapy na instance P vln. DS-B<sub>V</sub>:{NSR} – podmnožina validační sady se záznamy se sinusovým rytmem. Tučně vyznačena nejvyšší skóre.

		DS-B <sub>V</sub> :{NSR}		
Model		P <sub>on</sub>	P <sub>peak</sub>	P <sub>end</sub>
M <sub>MIL</sub>	Se	0.83	0.83	0.83
	P <sup>+</sup>	<b>0.84</b>	0.84	0.84
M <sub>3MIL</sub>	Se	0.85	0.84	0.85
	P <sup>+</sup>	<b>0.84</b>	0.85	0.84
M <sub>SIM</sub>	Se	<b>0.88</b>	<b>0.88</b>	<b>0.88</b>
	P <sup>+</sup>	0.82	<b>0.86</b>	<b>0.86</b>

V této skupině signálů dosáhly všechny tři modely hodnot Se a P<sup>+</sup> nad 80 %. To lze považovat za dobrý výsledek i v kontextu dalších prací validovaných na standardních databázích. Modely M<sub>MIL</sub> i M<sub>3MIL</sub> dosáhly velmi podobných hodnot Se a P<sup>+</sup>, a to v případě P<sub>peak</sub> 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<sub>SIM</sub>.

### 3.11 Ukázky segmentace

V ukázce na Obrázku 3.2 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. 3.2: Ukázka segmentace EKG pomocí modelů  $M_{MIL}$ ,  $M_{3MIL}$  a  $M_{SIM}$  za přítomnosti síňových a komorové extrasystoly: 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.



## 3.12 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 arytmií nelze považovat za triviální problém. Nejlepšího výsledku bylo dosaženo pomocí navržené modifikace kontrastního učení s hodnotou senzitivity a pozitivní prediktivní hodnoty v případě  $P_{\text{peak}}$  72 %, respektive 82 %. Post-hoc vizuální analýza ukázala, že největší obtíže činilo nalezení P vln zcela překrytých QRS komplexem. Tento typ záznamů měl vliv na 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. Modely tedy 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. Bez zásadních zásahů do metodologie učení nelze u  $M_{3MIL}$  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.

### 3.12.1 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 [14, 15, 58, 59]. Nejčastěji [11, 12, 13, 50, 52, 60, 44] za tímto účelem bývá využívána databáze QT-DB. Publikované algoritmy hlubokého učení [13, 50] reportují na QT-DB senzitivitu a specificitu shodně nad 98 % v případě obou metrik. Práce autorů Saclova et al [61, 46] 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. Také při validaci na EKG záznamech s primárně sinusovým rytmem lze ovšem jen stěží odhalit, jestli je detekce podložena smysluplnou

extrakcí příznaků nebo jen pravděpodobnostním očekáváním. 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.

### 3.12.2 Limitace

S ohledem na neúplné anotace jsou hodnoty metrik kvality pouze orientační. Výsledky tedy 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 [61].

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, že nebylo možné pro testovanou variantu dosáhnout stabilního procesu učení. Některé metody by tak mohli při optimální volbě hyperparametrů teoreticky dosáhnout stejných výsledků jak 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 elektrofyzilogického 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.

V práci byly navrženy softwarové nástroje pro export dat do otevřeného formátu. Díky tomu bylo možné sestavit trénovací datové sady s širokým spektrem poruch srdečního rytmu. U datasetů byla provedena expertní anotace a anotace navrženým automatickým algoritmem.

V Kapitole 2 byl navržen algoritmus pro automatickou segmentaci síňových elektrogramů (EGM). Algoritmus využívá vlastní implementaci 1D konvoluční neuronové sítě s reziduálním spojením. Pro rozšíření receptivní oblasti modelu byl do modelu implementován návrh vrstvy s kompresními a dilatovanými konvolučními jádry. 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 s zde prezentovaným modelem. 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 P vlnou od zbývajících částí signálu. Navržená metoda umožňuje nalézt množství P vln, které jsou zcela skryté v QRST komplexu. 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 tak teoreticky možné přenést do dalších oblastí analýzy EKG se zachováním vnořené informace.

## Literatura

- [1] Paul Kligfield. The centennial of the einthoven electrocardiogram. *Journal of electrocardiology*, 35:123, 2002.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] 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>.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.

- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] Pablo Laguna, Roger G Mark, Ary L Goldberger, and George B Moody. The qt database, 1997. URL: <https://physionet.org/content/qtddb/>, doi:10.13026/C24K53.
- [15] JL Willems, P Arnaud, JH van Bommel, R Degani, PW Macfarlane, and C Zywiets. 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>.
- [16] Alena Kalyakulina, Igor Yusipov, Viktor Moskalenko, Alexander Nikolskiy, Konstantin Kosonogov, Nikolai Zolotykh, and Mikhail Ivanchenko. Lobachevsky university electrocardiography database, 2021.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] 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.

- [21] 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.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] Roger S. Damle, Willie Choe, Nabil M. Kanaan, Frederick A. Ehlert, Jeffrey J. Golberger, 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.
- [26] 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.
- [27] 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.
- [28] GJIN NDREPEPA, EDWARD B. CAREF, HONG YIN, NABIL EL-SHERIF, and MARK RESTIVO. 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.
- [29] 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.

- [30] 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.
- [31] 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.
- [32] 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.
- [33] 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.
- [34] 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.
- [35] 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.
- [36] 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.
- [37] 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.
- [38] 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.
- [39] 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.

- [40] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019. URL: <https://openreview.net/forum?id=Bkg6RiCqY7>.
- [41] 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>.
- [42] 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.
- [43] 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.
- [44] 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.
- [45] 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.
- [46] 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.
- [47] 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.
- [48] 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.
- [49] C. Vaya, J.J. Rieta, C. Sanchez, and D. Moratal. Convolutive 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.



- [50] 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.
- [51] 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.
- [52] 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/jmihi.2019.2815.
- [53] 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.
- [54] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Philip 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.
- [55] 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.
- [56] Xinggang 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.
- [57] 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.
- [58] 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.
- [59] Lucie Maršánová, Andrea Němcová, Radovan Smíšek, Tomáš Goldmann, Martin Vitek, 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.

- [60] 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.
- [61] 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.

## A Curriculum vitae

---

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

---

### Vzdělání

---

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

---

### Pracovní praxe

---

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

---

## Lektorská činnost

---

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

---

## Projekty

---

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

---

## Kurzy a další vzdělání

---

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

---

## Soutěže a ocenění

---

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

## **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.