



Zadání bakalářské práce

Název:	Predikce fibrilace síní prostřednictvím metod hlubokého učení
Student:	Nikola Zadorozhny
Vedoucí:	Ing. Jan Kubant
Studijní program:	Informatika
Obor / specializace:	Znalostní inženýrství
Katedra:	Katedra aplikované matematiky
Platnost zadání:	do konce letního semestru 2022/2023

Pokyny pro vypracování

Cílem práce je pomocí metod hlubokého učení navrhnout model vhodný k predikování fibrilace síní na základě záznamů poskytnutých Institutem klinické a experimentální medicíny (IKEM).

Úkoly pro řešitele jsou následující:

- Seznámení se s problematikou detekce srdečních arytmií ze snímků elektrokardiografu (EKG) pomocí metod hlubokého učení.
- Analýza dat, jejich předzpracování a tvorba datasetu.
- Navržení a implementace modelu hlubokého učení pro rozpoznávání fibrilace síní.
- Zhodnocení úspěšnosti, porovnání s existujícími přístupy zabývajícími se problematikou automatické detekce srdečních arytmií.

[1] RIBEIRO, Antônio H., et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. Nature communications, 2020, 11.1: 1-9.

[2] WEIMANN, Kuba; CONRAD, Tim OF. Transfer learning for ECG classification. Scientific reports, 2021, 11.1: 1-12.

Bakalářská práce

**PREDIKCE FIBRILACE
SÍNÍ
PROSTŘEDNICTVÍM
METOD HLUBOKÉHO
UČENÍ**

Nikola Zadorozhny

Fakulta informačních technologií
Katedra aplikované matematiky
Vedoucí: Ing. Jan Kubant
16. února 2023

České vysoké učení technické v Praze
Fakulta informačních technologií

© 2023 Nikola Zadorozhny. Všechna práva vyhrazena.

Tato práce vznikla jako školní dílo na Českém vysokém učení technickém v Praze, Fakultě informačních technologií. Práce je chráněna právními předpisy a mezinárodními úmluvami o právu autorském a právech souvisejících s právem autorským. K jejímu užití, s výjimkou bezúplatných zákonných licencí a nad rámec oprávnění uvedených v Prohlášení, je nezbytný souhlas autora.

Odkaz na tuto práci: Zadorozhny Nikola. *Predikce fibrilace síní prostřednictvím metod hlubokého učení*. Bakalářská práce. České vysoké učení technické v Praze, Fakulta informačních technologií, 2023.

Obsah

Poděkování	vii
Prohlášení	viii
Abstrakt	ix
Seznam zkratek	x
Úvod	1
Cíle práce	2
1 Medicínské zobrazovací techniky a charakteristika onemocnění	3
1.1 Srdeční činnost	3
1.2 Elektrokardiografie	4
1.2.1 Snímání EKG	4
1.2.2 Tvar fyziologického EKG	4
1.3 Srdeční arytmie	5
1.3.1 Atrioventrikulární blokáda	5
1.3.2 Síňová tachykardie	5
1.3.3 Flutter síní	5
1.3.4 Fibrilace síní	5
2 Strojové učení a neuronové sítě	7
2.1 Strojové učení	7
2.2 Neuronové sítě	8
2.2.1 Formální neuron a neuronová síť	8
2.2.2 Proces učení neuronové sítě a zpětná propagace	9
2.2.3 Aktivační funkce	10
2.2.4 Ztrátová funkce	11
2.2.5 Optimalizační metody	11
2.3 Konvoluční neuronové sítě	13
2.3.1 Vstupní vrstva	13
2.3.2 Konvoluční vrstva	13
2.3.3 Poolingová vrstva	14
2.3.4 Plně propojená vrstva	14
3 Implementace CNN a související problémy	15
3.1 Rozdělení datasetu	15
3.1.1 Křížová validace	15
3.2 Nedoučení a přeučení	16
3.3 Regularizace	17
3.3.1 Dropout	17
3.3.2 Metoda Early Stopping	17

3.3.3	BatchNorm vrstva	17
3.3.4	L1 a L2 regularizace	17
3.4	Krok učení	18
3.5	Architektury konvolučních neuronových sítí	18
3.5.1	AlexNet	18
3.5.2	VGGNet	18
3.5.3	ResNet	18
3.5.4	Architektura NN pro klasifikaci 12-svodového EKG	19
3.6	Nevyváženost datasetu	19
4	Přenesené učení a augmentace dat	21
4.1	Augmentace dat	21
4.1.1	Generativní modely	21
4.1.2	Transformace obrazu	21
4.2	Přenesené učení	24
5	Vyhodnocení úspěšnosti modelů	26
5.1	Binární klasifikace	26
6	Výzkumy v oblasti související s tématem této práce	29
6.1	Detekce arytmií pomocí extrakce rysů a základních algoritmů strojového učení	29
6.2	Detekce arytmií pomocí konvolučních neuronových sítí	30
6.3	Augmentace dat u EKG záznamů	30
6.4	Detekce arytmií pomocí přeneseného učení	31
6.5	Detekce arytmií pomocí rekurentních neuronových sítí	31
7	Vlastní řešení	32
8	Analýza dat a příprava datasetu	33
9	Základní model a jeho optimalizace	35
9.1	Architektura a optimalizace hyperparametrů	35
9.2	Augmentace dat a vybalancování datasetu	36
10	Přenesené učení	39
10.1	Volba předtrénovaného modelu a porovnání datasetů	39
10.2	Vytrénování modelu od základů	39
10.3	Přenesené učení formou extrakce rysů	40
10.4	Přenesené učení formou ladění	41
11	Srovnání výsledků	43
11.1	Systém MUSE	43
11.2	Otestování na další testovací sadě	43
11.3	Srovnání výsledků	45
12	Diskuze	46
13	Závěr	48
A	Optimalizace hyperparametrů	49
B	Porovnání EKG záznamů	51
	Obsah přiloženého média	58

Seznam obrázků

1.1	Převodní srdeční systém, převzato z [2]	3
1.2	Srdeční cyklus na záznamu EKG, převzato z [7].	4
1.3	Fibrilace síní na záznamu EKG, záznam z vlastního datasetu.	6
2.1	Schéma formálního neuronu, vlastní zpracování podle [11].	8
2.2	Schéma neuronové sítě, vlastní zpracování podle [11].	9
2.3	Schéma procesu učení neuronové sítě, vlastní zpracování podle [12].	9
2.4	Aktivační funkce sigmoid.	10
2.5	Aktivační funkce ReLU.	11
2.6	Konvoluce, vlastní zpracování podle [11].	13
2.7	Pooling, vlastní zpracování podle [11].	14
3.1	Vnořená křížová validace	16
3.2	Přeučení modelu a early stopping	17
3.3	Architektura CNN z publikace [21].	19
4.1	Způsoby transformace původního signálu EKG.	23
4.2	Volba strategie přeneseného učení v závislosti na podobnosti a velikosti datasetu.	25
5.1	Matice záměn	27
5.2	PRC křivka	28
8.1	Zastoupení jednotlivých typů srdečního rytmu v datasetu.	33
9.1	Výsledná architektura modelu.	35
9.2	Ukázka průběhu křivky učení při nevhodné volbě hodnot transformací (červeně) v porovnání s původním datasetem (modře) a rozšířeným datasetem (zeleně). Plná čára značí trénovací přesnost (resp. chybu), přerušovaná čára značí validační přesnost (resp. chybu).	36
9.3	Ukázka náhodné kombinace transformací signálu.	37
9.4	Zobrazení křivky učení (trénovací a validační chyby). Čára reprezentuje průměr všech tří běhů, obarvená oblast zobrazuje hodnotu směrodatné odchylky v daném bodě.	38
10.1	Porovnání průběhu učení v případě při požití předtrénovaných vah oproti náhodné inicializaci.	40
11.1	Porovnání matice záměn u finálního modelu (vlevo) a systému MUSE (vpravo) na testovací sadě č.2.	44
A.1	Hodnoty AUPRC v závislosti na nastavení hyperparametrů.	50
B.1	Příklad sinusového rytmu z testovacího datasetu publikace [21].	52
B.2	Příklad sinusového rytmu z vlastního datasetu.	52

B.3	Příklad fibrilace síní z testovacího datasetu publikace [21].	53
B.4	Příklad fibrilace síní z vlastního datasetu.	53

Seznam tabulek

9.1	Porovnání přístupů vybalancování datasetu.	38
10.1	Porovnání úspěšnosti modelu na testovací sadě z publikace oproti vlastním testovacím sadám.	39
10.2	Porovnání úspěšnosti přeneseného učení formou extrakce rysů.	40
10.3	Porovnání úspěšnosti přeneseného učení formou ladění.	41
10.4	Porovnání úspěšnosti přeneseného učení formou ladění se sekundárním přetrénování celého modelu.	42
11.1	Porovnání výsledků klasifikace FS na testovací sade č.2.	44
11.2	Porovnání výsledků klasifikace srdečních rytmů, především pak FS.	45

Chtěla bych poděkovat především vedoucímu své bakalářské práce Ing. Janu Kubantovi za vedení, projevenou ochotu a čas strávený konzultacemi práce. Dále děkuji MUDr. Markovi Šramkovi, Ph.D. FESC a Ditě Kvasničkové za poskytnutí dat.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracovala samostatně a že jsem uvedla veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisů, zejména skutečnost, že České vysoké učení technické v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 citovaného zákona.

V Praze dne 16. února 2023

.....

Abstrakt

Tato práce se zabývá klasifikací nejčastějšího typu srdečních arytmií, a to fibrilace síní pomocí konvolučních neuronových sítí (CNN). Činí tak skrze interpretaci 12-svodových záznamů EKG, které byly poskytnuty ve spolupráci s Institutem klinické a experimentální medicíny (IKEM). V teoretické části práce jsou představeny klíčové koncepty z oblasti neuronových sítí, arytmiologie, přeneseného učení a augmentace dat. V praktické části práce je navržena a optimalizována CNN klasifikující fibrilaci síní, problém malého množství dat tato práce řeší skrze implementaci technik přeneseného učení a augmentace dat. Optimalizovaný model navržený v této práci dosahuje v klíčových metrikách hodnot 90 % senzitivity, 94 % specificity a 0,89 F1-skóre. Oproti řešení, které se používá v klinické praxi, model dosahoval o 10 % vyšší hodnoty senzitivity, má tedy vyšší schopnost přesně detekovat pozitivní případy výskytu fibrilace síní.

Klíčová slova konvoluční neuronové sítě, fibrilace síní, EKG, přenesené učení, augmentace dat, hluboké učení

Abstract

This work focuses on the classification of the most common type of cardiac arrhythmias, namely atrial fibrillation, using convolutional neural networks (CNN). It does so through the interpretation of 12-lead ECG recordings, which were provided in cooperation with the Institute of Clinical and Experimental Medicine (IKEM). In the theoretical part of the thesis, key concepts from the field of neural networks, arrhythmology, transfer learning and data augmentation are presented. In the practical part of the work, a CNN classifying atrial fibrillation is designed and optimized, while the problem of having a small amount of data is solved by this work through the implementation of transfer learning and data augmentation techniques. The optimized model proposed in this work achieves values of 90 % sensitivity, 94 % specificity and 0.89 F1-score in key metrics. Compared to the solution that is used in clinical practice, the model achieved a 10 % higher sensitivity value. Therefore, it has a higher ability to accurately detect positive cases of atrial fibrillation.

Keywords convolutional neural networks, atrial fibrillation, ECG, transfer learning, data augmentation, deep learning

Seznam zkratek

AdaGrad	Adaptive gradient descent
Adam	Adaptive moment estimation
AF	Atrial fibrillation (fibrilace síní)
AUPRC	Area under the precision-recall curve (oblast pod PR-křivkou)
AV	atrioventrikulární
CNN	Convolutional neural network (konvoluční neuronová síť)
EKG	elektrokardiogram
FC layer	Fully connected layer (plně propojená vrstva)
FN	False negative (falešně negativní)
FP	False positive (falešně pozitivní)
FS	fibrilace síní
GAN	Generative adversarial network
GD	Gradient descent (metoda gradientního sestupu)
IKEM	Institut klinické a experimentální medicíny
LSTM	Long short-term memory
MUSE	Marquette universal system for electrocardiography
NN	Neural network (neuronová síť)
PRC	Precision-recall curve
ReLU	Rectified linear unit
RGB	Red-green-blue (barevný model červená-zelená-modrá)
RNN	Recurrent neural network (Rekurentní neuronové sítě)
SA	sinoatriální
SGD	Stochastic gradient descent (metoda stochastického gradientního sestupu)
SMOTE	Synthetic minority oversampling technique
SR	sinusový rytmus
SVM	Support vector machines (metoda podpůrných vektorů)
TL	Transfer learning (přenesené učení)
TN	True negative (skutečně negativní)
TP	True positive (skutečně pozitivní)

Úvod

Neuronové sítě dosahují v oblasti klasifikace obrazu velmi slibných výsledků. Lze tedy očekávat, že aplikace této technologie může významně zvýšit úroveň v klinické praxi, obzvláště pak v oblasti diagnostiky v kardiologii, radiologii, onkologii nebo sonografii. Neuronové sítě disponují schopností identifikovat ve velkém množství dat vzory a korelace, které nemusí být člověkem snadno rozpoznatelné. V určitých oblastech mohou neuronové sítě dokonce předčít lidského odborníka.

Pro úspěšnou aplikaci neuronových sítí ve zdravotnictví je nicméně překážkou nedostatek velkého množství anotovaných dat s vyváženým rozložením a vysokou diverzitou. Zdravotnická data jsou navíc značně komplexní, heterogenní a vícedimenzionální.

Jedním z přístupů, jakým se dá pracovat i s malým množstvím dat, je metoda přeneseného učení. Podstatou přeneseného učení je předtrénování neuronové sítě na podobném typu úlohy s velkým množstvím dat a následné přetrénování neuronové sítě na cílovém typu úlohy s menším množstvím dat.

Kardiovaskulární onemocnění jsou celosvětově hlavní příčinou úmrtí a jejich léčba představuje významnou zátěž pro zdravotní systémy a veřejné finance. Fibrilace síní patří mezi nejčastější typy srdečních arytmií. V důsledku tohoto onemocnění mohou u pacientů nastat závažné komplikace jako cévní mozková příhoda nebo srdeční selhání. Někteří pacienti nemusí pociťovat žádné příznaky, fibrilace síní tak může být diagnostikována až při některé z vážných komplikací. Tento typ arytmiie se podobně jako i jiné typy arytmií diagnostikuje především pomocí elektrokardiografie. Elektrokardiografie je neinvazivní metoda umožňující detekování srdeční činnosti, jejíž interpretace nicméně vyžaduje značné množství odborné znalosti a praxe.

Využití neuronových sítí v predikci fibrilace síní umožňuje zacílit možnosti a potenciál neuronových sítí na jedno z nejvíce rozšířených celosvětových onemocnění. Automatická interpretace EKG záznamu může zkrátit dobu potřebnou ke stanovení diagnózy a zároveň může zvýšit pravděpodobnost zachycení méně nápadných poruch srdečního rytmu. Dalším důležitým benefitem automatizace interpretace EKG záznamu je možnost rozšíření diagnostiky bez nutnosti časové zátěže lékařských odborníků potřebných k interpretaci.

Tato práce si klade za cíl navrhnout model neuronových sítí pro automatickou detekci fibrilace síní s ohledem na problematiku nedostatečného množství dat a nevyváženosti klasifikačních tříd. Data, která tato práce využívá byla poskytnuta Institutem klinické a experimentální medicíny (IKEM).

Cíle práce

Cílem této práce bylo prostřednictvím metod hlubokého učení navrhnout model vhodný ke klasifikaci fibrilace síní na základě záznamů EKG poskytnutých IKEM.

Tato práce v kapitole 1 popisuje charakteristiku srdečních arytmií, včetně fibrilace síní a shrnuje teoretické základy elektrokardiografie. Součástí kapitoly 2 je teoretický úvod do neuronových sítí (NN), konkrétně pak konvolučních NN. 3. kapitola se věnuje implementaci konvolučních NN a souvisejícím oblastem, mezi něž patří regularizace nebo metoda early stopping. 4. kapitola obsahuje seznámení se s problematikou nevyváženého datasetu, augmentace dat a přeneseného učení. Kapitola 5 shrnuje výpočet metrik pro vyhodnocení úspěšnosti modelu. Kapitola s číslem 6 shrnuje některé z dosavadních výzkumů v oblasti související s tématem práce.

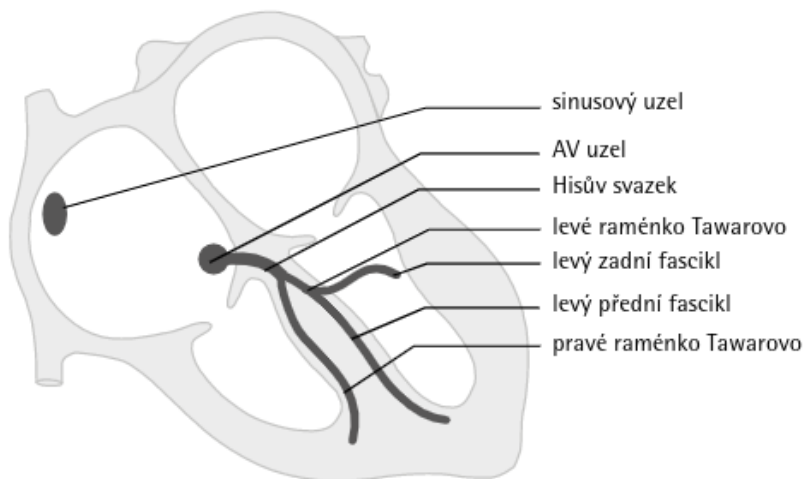
Kapitoly 7 a 8 se věnují úvodu do praktické části práce a přípravě datasetu. Kapitola 9 obsahuje návrh a optimalizaci základní konvoluční NN, implementaci vybraných metod transformace signálu EKG a porovnání několika přístupů vybalancování tříd a rozšíření datasetu. Jedním z vytyčených cílů práce bylo adresovat problém s menší velikostí datasetu, kapitola 10 implementuje a porovnává různé přístupy přeneseného učení. Na závěr pak kapitola 11 obsahuje shrnutí výsledků, otestování navržených modelů na jiné testovací sadě a vyhodnocení úspěšnosti interpretace srdečního rytmu v současné době používaným nemocničním systémem.

..... Kapitola 1

Medicínské zobrazovací techniky a charakteristika onemocnění

1.1 Srdeční činnost

Srdeční sval se skládá ze čtyř dutin, a to pravé komory, pravé síně, levé komory a levé síně. Srdeční cyklus začíná elektrickým impulsem z oblasti pravé síně nazývané sinoatriální (SA) uzel. Vzruch (vlna depolarizace) se následně šíří skrze svalovinu síně do oblasti zvané atrioventrikulární (AV) uzel, kde se vzruch zpomaluje. Zpomalení zabraňuje předčasnému šíření do komor v případě patologického rytmu, např. fibrilace síní [1]. Vzruch se dále šíří skrze tzv. Hisův svazek, který se větví na pravé a levé Tawarovo raménko. Levé raménko se dále větví do dvou svazků. Skrze Tawarova raménka a Purkyňova vlákna se vzruch šíří do pravé a levé komory [1–3]. Celý tento proces se nazývá převodní srdeční systém (viz obrázek 1.1).



■ **Obrázek 1.1** Převodní srdeční systém, převzato z [2]

V SA uzlu probíhá depolarizace rychleji než v AV uzlu nebo komorách, proto vzruch vzniká za normálních podmínek tam. SA uzel se označuje také jako tzv. primární pacemaker (udavatel rytmu), určuje frekvenci srdečního rytmu [1]. Normální srdeční rytmus, vycházející z SA uzlu, se nazývá sinusový rytmus (SR). Pokud je funkce SA uzlu narušena, vzruch může vzniknout

i v jiných místech srdce, např. v AV uzlu (takový rytmus se označuje jako junkční rytmus) nebo ve svalovině komor. Vzhledem k tomu, že v AV uzlu i komorách probíhá depolarizace pomaleji, junkční i komorový rytmus obvykle mají nižší frekvenci než rytmus sinusový [4]. Jednotlivým srdečním rytmům se blíže věnuje kapitola 1.3.

1.2 Elektrokardiografie

Srdeční činnost lze snímat různými způsoby, elektrokardiografie patří mezi standardní vyšetřovací metody, hraje klíčovou roli v diagnostice a léčbě poruch srdečního rytmu.

1.2.1 Snímání EKG

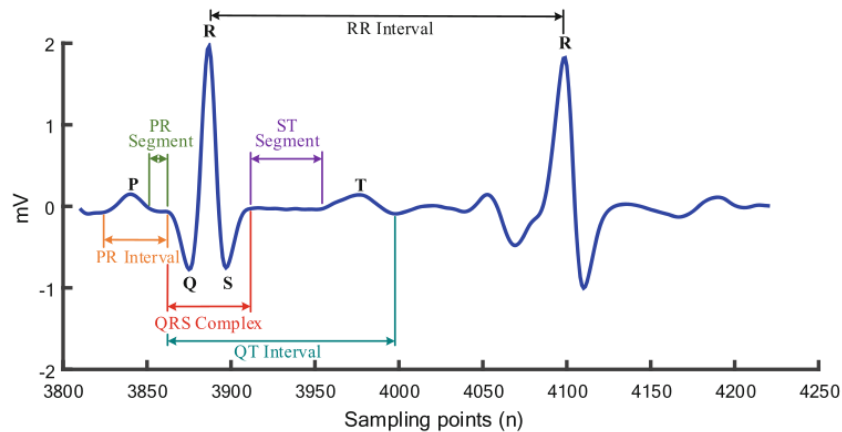
Elektrokardiograf průběh srdeční činnosti snímá prostřednictvím elektrod rozmístěných na různých částech těla. Grafický záznam průběhu měření se nazývá elektrokardiogram (dále jen EKG). Elektrody poskytují elektrický obraz srdce, jinak také nazývaný svod. Každý svod zobrazuje jiný pohled na srdeční aktivitu, tedy i jiný EKG obraz.

Při standardním vyšetření se využívá 12-svodové EKG, které se snímá prostřednictvím deseti elektrod [2], ekg-strucne. Celkem šest svodů (označují se jako I, II, III, aVR, aVL, aVF) zaznamenávají elektrody umístěné na končetinách. Končetinové svody poskytují zobrazení srdeční aktivity do frontální projekce. Svody V_1 až V_6 se připevňují k hrudníku a snímají srdce v horizontální rovině. Při diagnostice poruch srdečního rytmu je účelné pozorovat především svody s výraznou vlnou P, tedy svody II, aVR a aVF.

Při monitorování pacientů se využívá také EKG s menším počtem svodů, například 3-svodové nebo jednosvodové [5, 6].

1.2.2 Tvar fyziologického EKG

Depolarizaci a repolarizaci síní i komor lze na EKG pozorovat prostřednictvím vln (zaoblené výchylky) a kmitů (ostré výchylky). Obrázek 1.2 znázorňuje průběh srdečního cyklu na záznamu EKG.



■ **Obrázek 1.2** Srdeční cyklus na záznamu EKG, převzato z [7].

Depolarizaci pravé a levé síně odpovídá vlna P. Interval PQ (případně PR, chybí-li kmit Q) se měří od počátku vlny P k počátku kmitu Q (případně R), odpovídá depolarizaci síní, šíření vzruchu skrze AV uzlu a končí počátkem depolarizace komor. Komplex QRS reprezentuje depolarizaci obou komor, dochází zde i k repolarizaci síní, ta však není na EKG pozorovatelná.

ST segment je izoelektrickou rovinou, značí dobu mezi depolarizací a repolarizací komor, kterou zobrazuje vlna T.

1.3 Srdeční arytmie

Poruchy srdečního rytmu, jinak také srdeční arytmie, vznikají v důsledku poruchy tvorby nebo šíření vzruchu. Vychází-li vzruch z SA uzlu, hovoříme o sinusovém rytmu. Frekvence sinusového rytmu se může lišit, pomalý sinusový rytmus (pod 60-50 pulsů za minutu) se také označuje jako sinusová bradykardie. Naopak zrychlený sinusový rytmus (nad 100 pulsů za minutu) se označuje jako sinusová tachykardie.

Abnormální srdeční rytmy mohou vzniknout ve třech různých oblastech, a to v myokardu síní (síňový rytmus), v oblasti AV uzlu (junkční rytmus) nebo myokardu komor (komorový rytmus). Abnormální rytmy mohou být pomalé (bradykardie), rychlé (tachykardie) nebo výrazně nepravidelné (fibrilace). Následující odstavce se věnují charakteristice nejčastějších arytmií [4, 8, 9], zejména pak fibrilaci síní, na kterou je praktická část práce zaměřena.

1.3.1 Atrioventrikulární blokáda

Atrioventrikulární blokáda (dále jen AV blokáda) je typ arytmie způsobené poruchou převodu vzruchu skrze AV uzel. AV blokáda se rozlišuje do tří stupňů. U prvních dvou stupňů lze na EKG pozorovat prodloužení, případně nepravidelnost PR intervalů. Třetí stupeň se označuje také jako kompletní AV blokáda, v tomto případě se sinusový rytmus nahradí komorovým.

1.3.2 Síňová tachykardie

V případě síňové tachykardie dochází k depolarizaci síní abnormálně rychle. Na EKG se síňová tachykardie vyznačuje vlnami P umístěnými na vlnách T [3].

1.3.3 Flutter síní

Při překročení frekvence 250/min již není na EKG patrná izoelektrická linie mezi vlnami P. Při takto vysoké frekvenci nelze skrze AV uzel vzruch dále šířit do komor. Na záznamu EKG může být v důsledku AV blokády pozorovatelný větší počet vln P vůči jednomu komplexu QRS. Tento typ arytmie se nazývá flutter síní. V důsledku vysoké frekvence může dojít k jeho záměně za fibrilaci síní.

1.3.4 Fibrilace síní

Fibrilace síní (FS) je nejčastějším typem srdečních arytmií a ačkoliv život člověka bezprostředně neohrožuje, může vést k závažným komplikacím jako je srdeční selhání nebo cévní mozková příhoda [9]. Mezi klinické projevy fibrilace síní patří palpitace, dušnost nebo nízká výkonnost při tělesné zátěži. Někteří pacienti mohou mít asymptomatický průběh a FS se u nich prokáže až při rozvinutí komplikací.

Fibrilace síní se může u pacientů vyskytovat samostatně nebo v kombinaci s jiným typem arytmií. Existuje několik různých přístupů léčby, vždy je třeba individuálně posoudit stav pacienta. Léčba může zahrnovat užívání léků tzv. antiarytmik, kardioverzi nebo katetrizační ablaci [9].

Fibrilaci síní lze klasifikovat na základě různých kritérií, nejčastěji dle časového období výskytu na:

- první záchyt arytmie
- paroxysmální FS (FS spontánně odezní)

- perzistující FS (arytmie přetrvává po dobu déle než 7 dní nebo vyžaduje kardioverzi)
- dlouhodobá perzistující FS (arytmie trvá déle než 1 rok)
- permanentní FS (léčba arytmie není účinná)



■ **Obrázek 1.3** Fibrilace síní na záznamu EKG, záznam z vlastního datasetu.

Fibrilace síní je podobně jako flutter síní charakteristická zvýšenou frekvencí. Na rozdíl od flutteru síní je však srdeční činnost značně nepravidelná. Na záznamu EKG nejsou výrazně patrné vlny P, místo nich lze pozorovat pouze drobné nepravidelné vlny o vysoké frekvenci, případně vlnění izoelektrické linie. Především u neléčených pacientů je akce komor značně nepravidelná, to lze na EKG pozorovat jako nepravidelné rozdělení intervalů RR (intervaly mezi dílčími kmity R) viz obrázek 1.3.

..... Kapitola 2

Strojové učení a neuronové sítě

2.1 Strojové učení

Strojové učení patří mezi odvětví umělé inteligence. Algoritmy strojového učení získávají znalosti přímo z dat (tento proces označujeme jako učení algoritmu), čímž napodobují průběh lidského učení z okolního prostředí. V procesu učení se automaticky optimalizují vnitřní hodnoty algoritmu, postupně se tak zlepšuje jeho přesnost v rámci dané úlohy.

Strojové učení se dělí do tří kategorií:

- učení s učitelem (supervised learning)
- učení bez učitele (unsupervised learning)
- částečné učení s učitelem (semi-supervised learning)

V případě učení s učitelem algoritmus pracuje s anotovaným datasetem. Algoritmus upravuje vnitřní hodnoty, dokud míra odchýlení predikovaných hodnot od hodnot skutečných není minimální. Algoritmy učení s učitelem lze dále dělit podle typu úlohy na klasifikační a regresní. Cílem klasifikační úlohy je zařadit vstupní data do předem definovaných klasifikačních tříd. Pakliže lze data rozdělit do dvou klasifikačních tříd, mluvíme o binární klasifikaci. Příkladem binární klasifikace je klasifikace nádoru, zda je nebo není maligní [10]. Úlohou regrese je najít mapovací funkci mezi vstupní proměnnou a spojitou výstupní proměnnou.

V případě učení bez učitele algoritmus odhaluje skryté vzorce a korelace z neanotovaných dat. Výstupem algoritmu je rozřazení dat do skupin na základě naučených kritérií. Částečné učení s učitelem pracuje s rozsáhlým neanotovaným datasetem v kombinaci s menším anotovaným.

Algoritmů strojového učení existuje celá řada, mezi ně patří:

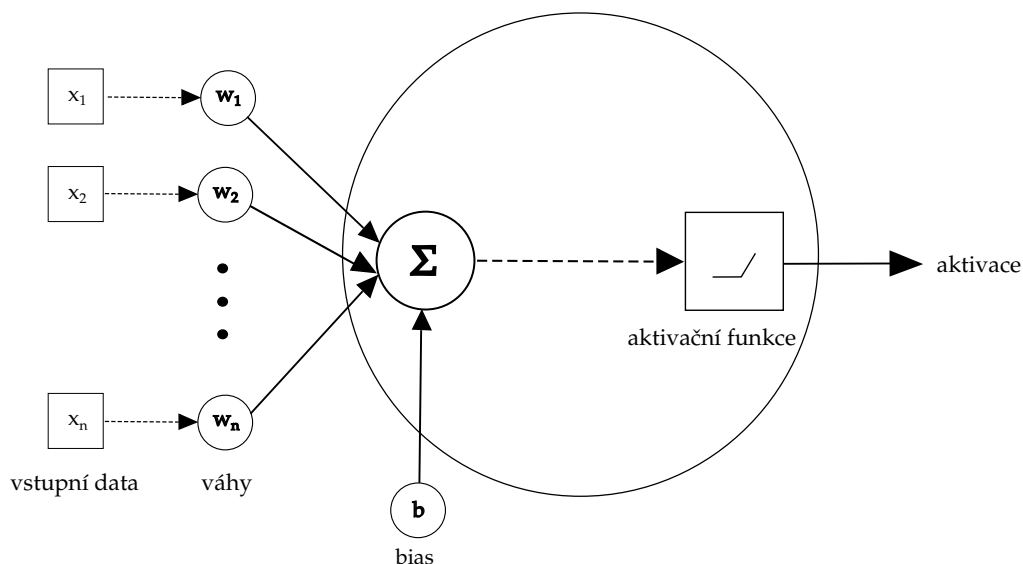
- Metoda podpůrných vektorů (SVM)
- Naivní Bayesův klasifikátor
- Rozhodovací stromy
- Algoritmus k-nejbližších sousedů
- Neuronové sítě

2.2 Neuronové sítě

Neuronové sítě jsou výpočetním modelem strojového učení inspirovaným biologickou strukturou nervových sítí živých organismů. Prostřednictvím neuronových sítí lze řešit i značně komplikované problémy, nevýhodou jejich využití je výpočetní náročnost a především nezbytnost velkého množství dat.

2.2.1 Formální neuron a neuronová síť

Základním stavebním prvkem neuronové sítě je tzv. formální neuron, který skrze matematickou funkci transformuje informace ze vstupů do jednoho výstupu. Každému vstupu neuronu je přiřazen číselný koeficient tzv. váha reprezentující jeho význam. Čím větší je hodnota váhy, tím důležitější je připojený vstup. Ze vstupních hodnot se vypočítá vážený součet, ke kterému se přičte konstanta vychýlení (angl. bias). Bias umožňuje posunutí průběhu aktivační funkce po ose x . Celkový součet se následně použije k výpočtu výstupu – aktivace neuronu skrze tzv. aktivační funkci.



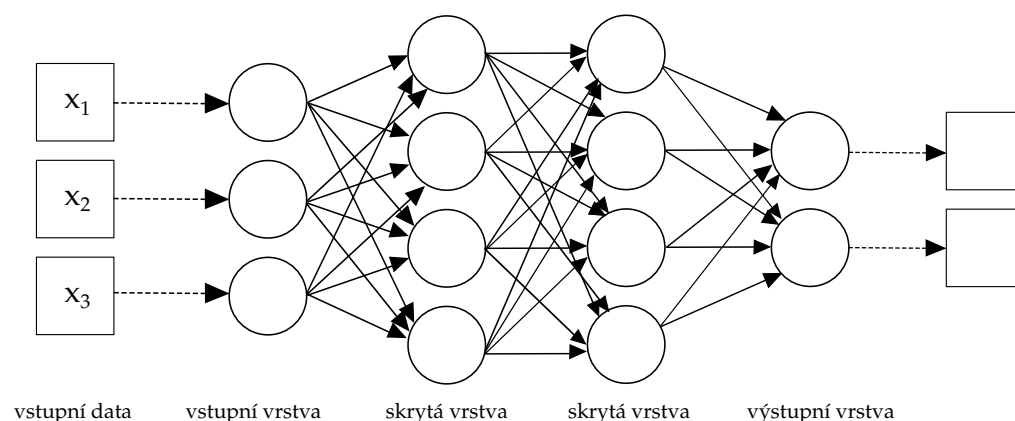
■ **Obrázek 2.1** Schéma formálního neuronu, vlastní zpracování podle [11].

Vzorec pro výpočet aktivace neuronu je následující:

$$y = f\left(b + \sum_{i=1}^n w_i a_i\right), \quad (2.1)$$

kde f je aktivační funkce, b je bias, n je počet vstupů a w_i $i \in \{1, 2, \dots, n\}$ jsou váhy přiřazené vstupům a_i pro $i \in \{1, 2, \dots, n\}$.

Jednotlivé neurony jsou vzájemně propojeny do vrstev tak, že výstup neuronu slouží jako vstup neuronu následující vrstvy (vyjma neuronu z poslední vrstvy sítě). Neuronovou síť lze rozdělit na vstupní vrstvu, skryté vrstvy a vrstvu výstupní (viz obrázek 2.2).

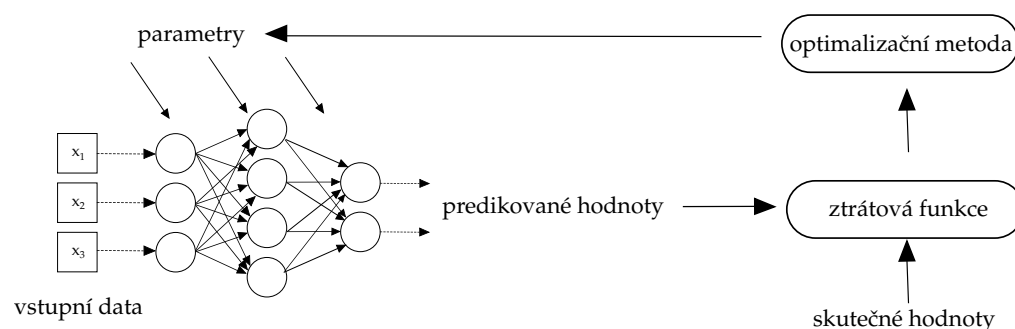


■ **Obrázek 2.2** Schéma neuronové sítě, vlastní zpracování podle [11].

Neurony první skryté vrstvy zpracovávají data ze vstupní vrstvy, která transformují pomocí aktivační funkce. Každá z následujících skrytých vrstev obdrží na vstupu aktivace neuronů její předchozí vrstvy. Počet skrytých vrstev závisí na zvolené architektuře. Hodnoty vah a vychýlení (parametry sítě) ve skrytých vrstvách reprezentují naučené vzorce extrahované ze vstupních dat. Výstupní vrstva transformuje aktivace z poslední skryté vrstvy pomocí specifické aktivační funkce, typicky se jedná o funkci sigmoid nebo softmax [11], které jsou blíže popsány v podkapitole 2.2.3. Tento postup se také označuje jako dopředný chod sítě (angl. forward pass).

2.2.2 Proces učení neuronové sítě a zpětná propagace

Naučené vzorce se uchovávají prostřednictvím parametrů jednotlivých vrstev. V procesu učení se parametry upravují tak, aby výsledná neuronová síť predikovala hodnoty s co nejmenší mírou odchýlení od hodnot skutečných. Úspěšnost trénování lze pozorovat prostřednictvím tzv. ztrátové funkce. Prostřednictvím algoritmu zpětné propagace (angl. back propagation) se vypočítá gradient ztrátové funkce v závislosti na jednotlivých parametrech. Na základě hodnoty gradientu optimalizační algoritmy modifikují parametry tak, aby ztrátovou funkci minimalizovaly. Díky zpětnému učení dochází k upřednostňování klíčových vzorů, volba ztrátové funkce a optimalizační metody je proto podstatná k úspěšnému natrénování sítě.



■ **Obrázek 2.3** Schéma procesu učení neuronové sítě, vlastní zpracování podle [12].

Proces učení neuronové sítě sestává z několika epoch, kdy jedna epocha odpovídá právě jednomu průchodu celé trénovací množiny dat.

Kromě parametrů neuronové sítě, které se v průběhu učení automaticky upravují, definujeme také hyperparametry sítě. Hyperparametry jsou hodnoty, které se nastavují manuálně tak, aby

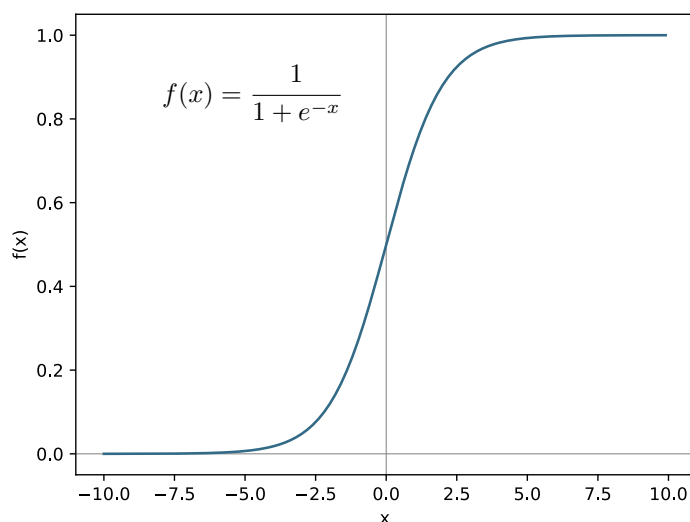
učení sítě probíhalo efektivněji. Mezi hyperparametry sítě patří výběr optimalizační metody, nastavení hodnoty kroku učení a další, popsané v následujících kapitolách.

2.2.3 Aktivační funkce

Aktivační funkce je matematická funkce transformující hodnoty ze vstupu neuronu. Výběr aktivační funkce závisí na typu vrstvy sítě a typu úlohy (klasifikace nebo regrese). V případě klasifikace je podstatný také počet klasifikačních tříd a formát výstupu. Volba aktivační funkce jednotlivých vrstev sítě je důležitou součástí procesu optimalizace hyperparametrů neuronové sítě. Tato kapitola popisuje některé z často používaných aktivačních funkcí.

Sigmoid je spojitá funkce s oborem hodnot v rozmezí intervalu (0, 1). Vzhledem ke svým vlastnostem je jednou z nejčastěji používaných aktivačních funkcí výstupní vrstvy při klasifikaci, kdy vrací hodnotu pravděpodobnosti příslušející dané klasifikační třídě. Sigmoid se typicky používá při binární klasifikaci (klasifikaci do dvou tříd). Lze ho však aplikovat i při klasifikaci do více tříd, kdy vrací pravděpodobnost příslušnosti třídám nezávisle na sobě, tedy vstup může patřit do jedné nebo více tříd, ale také do žádné.

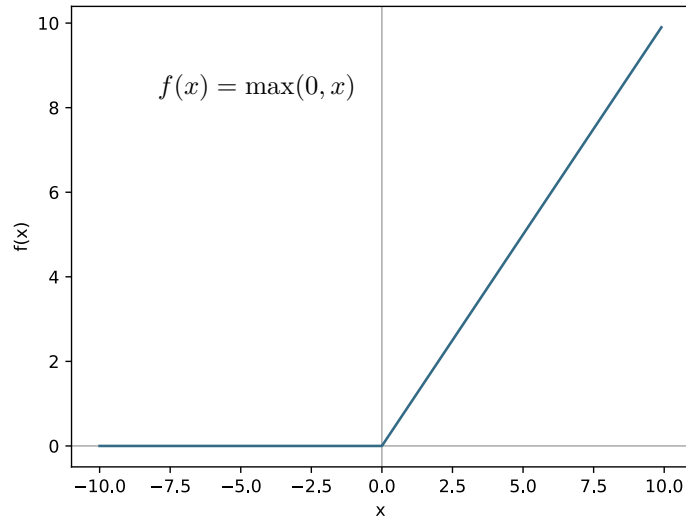
Při klasifikaci do více tříd se také často používá funkce softmax. Aktivační funkce softmax má rovněž obor hodnot v rozmezí intervalu (0, 1). Na rozdíl od funkce sigmoid při klasifikaci do více tříd každý výstup náleží právě jedné třídě, a to té s největší pravděpodobností. Součet pravděpodobností všech klasifikačních tříd se rovná hodnotě 1.



■ **Obrázek 2.4** Aktivační funkce sigmoid.

Podstatnou nevýhodou použití aktivační funkce sigmoid je problém mizejícího gradientu (angl. vanishing gradient). Pakliže se hodnota gradientu blíží k 0, nedochází v průběhu zpětné propagace k modifikaci parametrů sítě. Čímž dochází k výraznému narušení učení neuronové sítě. Z tohoto důvodu se ve skrytých vrstvách sítě využívají jiné aktivační funkce, mezi které patří aktivační funkce ReLU (rectified linear unit) [13].

Funkční hodnota aktivační funkce ReLU se při záporném vstupu rovná 0. V případě, že je na vstupu kladná hodnota, funkce se chová lineárně a výstupní hodnoty se rovnají vstupním. U této funkce nedochází k problému s mizejícím gradientem, nicméně její nevýhodou je vynulování všech záporných hodnot, kdy dochází k jevu „The Dying ReLU problem“. Pro vstupní hodnoty $x < 0$ se gradient funkce rovná 0, v důsledku čehož se příslušné parametry při zpětné propagaci neupravují. Některé neurony sítě se tak nikdy neaktivují [14].



■ **Obrázek 2.5** Aktivační funkce ReLU.

2.2.4 Ztrátová funkce

Ke kontrole úspěšnosti trénování neuronové sítě slouží ztrátová funkce, která určuje míru odchýlení predikovaných hodnot od skutečných. Volba ztrátové funkce se odvíjí od typu úlohy, při klasifikaci se často volí křížová entropie, v případě binární klasifikace konkrétně binární křížová entropie.

Výpočet binární křížové entropie je následující:

$$L(y, \hat{y}) = -\frac{1}{n} \sum_{i=1}^n (y \log(\hat{y}_i) + (1 - y) \log(1 - \hat{y}_i)), \quad (2.2)$$

kde y je skutečná hodnota, \hat{y} je predikovaná hodnota, n je počet testovacích dat.

V případě nevyváženého poměru tříd vstupních dat je vhodné křížovou entropii upravit přidáním konkrétních vstupních vah, jejichž hodnotu lze zvolit na základě procentuálního zastoupení jednotlivých tříd v datasetu. Takto modifikovanou ztrátovou funkci označujeme jako váhovou křížovou entropii.

2.2.5 Optimalizační metody

Úlohou optimalizačních metod je v průběhu trénování neuronové sítě minimalizovat ztrátovou funkci prostřednictvím úprav parametrů sítě.

Základní optimalizační metodou je metoda gradientního sestupu (angl. gradient descent, dále jen GD). Principem metody je postupné posouvání hodnot parametrů proti směru gradientu o konstantu tzv. krok učení. Tímto způsobem metoda postupuje z počátečního nastavení parametrů směrem k hodnotám s minimální ztrátovou funkcí, kdy se gradient rovná nule.

Stochastický gradientní sestup

Stochastický gradientní sestup (angl. stochastic gradient descent, dále jen SGD) vychází z metody gradientního sestupu. Zatímco u GD dochází k úpravě parametrů na základě celé trénovací sady, u SGD se parametry aktualizují na základě náhodně zvolené podmnožiny trénovací sady. Vypočítaný gradient je v tomto případě pouze odhadem skutečného gradientu, z toho důvodu

SGD konverguje oproti GD rychleji. Hyperparametr určující velikost podmnožiny pro výpočet gradientu se nazývá batch size.

Metody odvozené od GD jsou náchylné k uvíznutí v místě lokálního optima. Tomuto problému částečně předchází zavedení parametru hybnosti (angl. momentum). Principem zavedení hybnosti je zahrnutí předešlých hodnot gradientů do aktuálního výpočtu. V každé iteraci probíhá výpočet nových hodnot parametrů podle

$$v_t = \beta v_{t-1} - \alpha \frac{\delta L}{\delta \theta}, \quad (2.3)$$

$$\theta = \theta + v_t, \quad (2.4)$$

kde $\frac{\delta L}{\delta \theta}$ je gradient ztrátové funkce v závislosti na parametru θ a v_{t-1} odpovídá rychlosti (resp. gradientům z předchozích iterací vynásobených krokem učení). $\beta \in [0, 1)$ je parametr hybnosti, čím větší je tato hodnota v porovnání s hodnotou kroku učení α , tím více předchází gradienty ovlivňují aktuální výpočet nových parametrů θ [15, 16].

AdaGrad

Optimalizační metoda AdaGrad (angl. adaptive gradient descent) na rozdíl od předchozích metod neudrží krok učení konstantní, ale v průběhu učení ho upravuje. Parametry mají přiřazenou hodnotu kroku učení individuálně. Pro parametry s velkým gradientem se hodnota kroku učení radikálně zmenšuje, naopak v případě malého gradientu se hodnota kroku učení snižuje méně výrazně. Tento trend je realizován vydělením hodnoty kroku učení druhou odmocninou součtu čtverců předchozích gradientů.

Nevýhodou použití této optimalizační metody je riziko předčasného ukončení učení vlivem příliš nízké hodnoty kroku učení [15].

RMSprop

RMSprop vznikl upravením optimalizační metody AdaGrad. RMSprop k úpravě parametrů využívá exponenciální vážený klouzavý průměr čtverců předchozích gradientů. Rozdílem oproti optimalizační metodě AdaGrad je, že algoritmus RMSprop upřednostňuje hodnoty gradientů z nedávné historie [15, 16].

Adam

Algoritmus adam (angl. adaptive moment estimation) jednoznačně patří mezi nejpoužívanější optimalizační metody. Tato metoda podobně jako algoritmus RMSprop ke škálování hodnoty kroku učení využívá exponenciální vážený klouzavý průměr čtverců předchozích gradientů (second-order moment v). Algoritmus je navíc rozšířen o hybnost, která zahrnuje exponenciální vážený klouzavý průměr předchozích gradientů (first-order moment m) [15, 16]. Úprava parametrů probíhá podle rovnic

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \frac{\delta L}{\delta \theta}, \quad (2.5)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \left(\frac{\delta L}{\delta \theta} \right)^2, \quad (2.6)$$

kde $\frac{\delta L}{\delta \theta}$ je gradient ztrátové funkce v závislosti na parametru θ , β_1 a β_2 jsou hyperparametry hybnosti řídící exponenciální pokles m a v . Následně jsou hybnosti korigované podle rovnic

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad (2.7)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}, \quad (2.8)$$

Výslednou úpravu parametrů reprezentuje rovnice 2.9

$$\theta = \theta - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}, \quad (2.9)$$

kde ϵ je malá konstanta zamezující dělení nulou.

2.3 Konvoluční neuronové sítě

Konvoluční neuronové sítě (angl. convolution neural networks, dále jen CNN) jsou typem neuronových sítí, jež se především využívají pro zpracování obrazu. CNN se zpravidla skládá z vrstev konvolučních, poolingových a plně propojených. Typicky po vstupní vrstvě následuje několik konvolučních a poolingových vrstev, které plní funkci extrakce rysů. Tyto extrahované rysy následně vstupují do plně propojených vrstev, kde se přiřadí k patřičným klasifikačním třídám. Jednotlivé vrstvy jsou detailněji popsány v následujících podkapitolách.

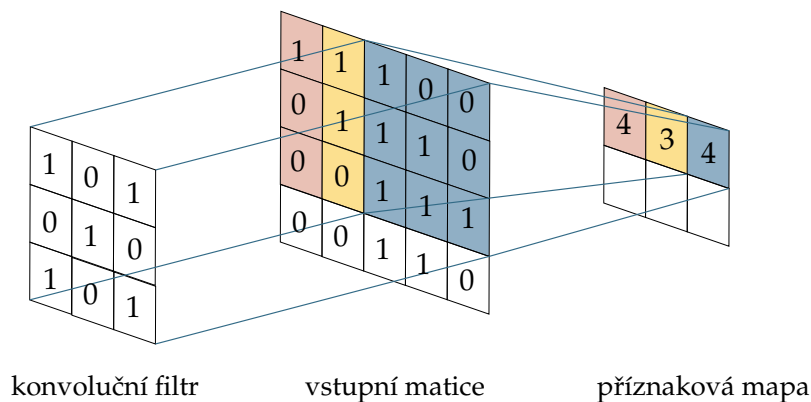
2.3.1 Vstupní vrstva

Vstupní vrstva CNN zpracovává trojrozměrný vstup, kde první dva rozměry udávají šířku a výšku obrazu, třetí rozměr reprezentuje hloubku. V případě zpracování obrazu se hloubka určuje podle počtu barevných kanálů (např. RGB má hloubku rovnou třem).

2.3.2 Konvoluční vrstva

Konvoluční vrstva je základním stavebním prvkem konvolučních neuronových sítí. CNN sestává z většího počtu konvolučních vrstev, kdy počáteční konvoluční vrstvy zachycují nízkouúrovňové prvky obrazu (např. čáry) a pozdější vrstvy extrahují prvky vysoké úrovně (např. specifické objekty). Principem konvoluční vrstvy je proces konvoluce.

Konvoluční vrstvu tvoří sada tzv. konvolučních filtrů, čtvercových matic stejného řádu. Konvoluce spočívá ve výpočtu skalárního součinu hodnot filtru s hodnotami překrývající se části vstupní matice. Konvoluční filtr se postupně posouvá po vstupní matici, výsledek konvoluce se zapisuje do tzv. příznakové mapy, která je výstupem konvoluční sítě. Velikost příznakové mapy se odvíjí od velikosti vstupní matice a konvolučního filtru. Konvoluce je znázorněna na obrázku 2.6.)



■ **Obrázek 2.6** Konvoluce, vlastní zpracování podle [11].

Počet a velikost konvolučních filtrů patří mezi hyperparametry sítě, které se v průběhu učení již nemění. Dílčí hodnoty filtrů, v obecném kontextu neuronových sítí se jedná o parametry, se v průběhu učení modifikují. Výstup konvoluční vrstvy se následně za účelem zajištění nelinearity upravuje prostřednictvím aktivační funkce, např. funkce ReLU.

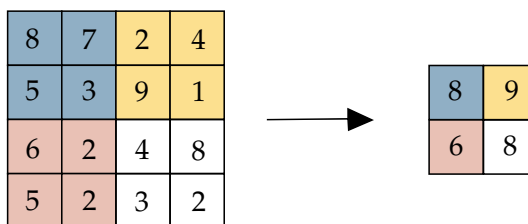
Mezi další hyperparametry konvoluční vrstvy patří tzv. krok konvoluce (angl. stride) a výplň (angl. padding). Krok konvoluce udává, o kolik vstupních prvků se konvoluční filtr bude posouvat. Padding se používá v případě, že je naším zájmem zachovat na výstupu rozměry vstupní matice. V takovém případě se kolem vstupní matice přidá určitý počet prvků nejčastěji s hodnotou 0.

V praktické části práce se využívá jednodimenzionální konvoluční síť. Vstupní vrstva 1D CNN očekává dvoudimenzionální vstup. V případě zpracování záznamu EKG první rozměr reprezentuje počet vzorků záznamu a druhý rozměr počet svodů. Rozdíl oproti dvoudimenzionální CNN je, že konvoluční filtry se pohybují pouze v jednom směru.

2.3.3 Poolingová vrstva

Mezi konvoluční vrstvy se vkládají vrstvy poolingové, jejichž účelem je zmenšení velikosti (podvzorkování) příznakových map. Poolingová vrstva redukuje velikost vstupu prostřednictvím filtrů. Zmenšení velikosti lze zajistit několika způsoby, nejčastěji výběrem maxima z části filtrem překryté vstupní matice (max-pooling, viz obrázek 2.7). Mezi jiné přístupy podvzorkování příznakových map patří výpočet průměru překryté části (average-pooling). Poolingová vrstva neobsahuje parametry, po dobu učení se proto nemění, nicméně obsahuje hyperparametry, jako je velikost filtru.

I zde se v případě 1D CNN filtry pohybují pouze v jednom směru.



■ **Obrázek 2.7** Pooling, vlastní zpracování podle [11].

2.3.4 Plně propojená vrstva

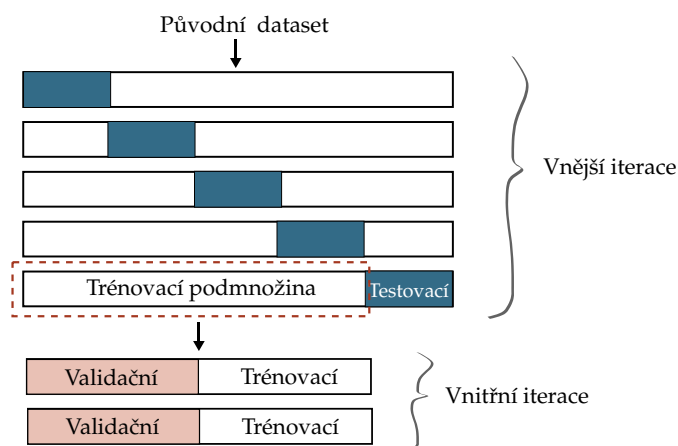
Součástí konvolučních sítí je jedna nebo více plně propojených vrstev (angl. fully connected layer, dále jen FC vrstva). Úlohou FC vrstvy je správné zařazení vstupních dat do příslušné klasifikační třídy. FC vrstva svou architekturou odpovídá obecnému schématu neuronových sítí (obrázek 2.2). Vstupem FC vrstvy je zploštělý výstup z poslední poolingové vrstvy. V případě klasifikace se v poslední vrstvě nejčastěji využívá aktivační funkce sigmoid nebo softmax.

dat a zbývající (k -tou) testovací podmnožinou. Výsledná úspěšnost modelu se označuje jako průměr výsledků napříč všemi testovacími podmnožinami.

Stratifikovaná k -násobná křížová validace je modifikací k -násobné křížové validace, kdy dataset je rozdělený do k podmnožin tak, že každá podmnožina obsahuje stejné procentuální zastoupení klasifikačních tříd. Tato modifikace křížové validace je obzvláště užitečná v případě nevyváženého datasetu.

Vnořená křížová validace je metoda užitečná v případě, že kromě samotného vyhodnocení modelu je třeba zároveň optimalizovat hyperparametry. Vnořená křížová validace probíhá ve dvou iteracích, vnitřní a vnější (viz obrázek 3.1). Vnější iterace probíhá totožně, jako je tomu v případě k -násobné křížové validace. Dataset se v každém kroku vnější iterace rozdělí na trénovací a testovací podmnožinu. V každém kroku vnější iterace se trénovací podmnožina dále rozdělí na vnitřní trénovací podmnožinu a validační podmnožinu. Účelem vnitřní iterace je identifikovat hyperparametry, které dosahují nejlepších výsledků na validační podmnožině. Model s optimalizovanými hyperparametry je natrénován na vnitřní trénovací a validační podmnožině a následně vyhodnocen na testovací podmnožině. Výsledná úspěšnost modelu se označuje jako průměr výsledků napříč všemi testovacími podmnožinami. Vnořená křížová validace může být výpočetně značně náročnou metodou, je třeba proto vhodně zvolit vnější i vnitřní hodnotu k . Pro vnitřní hodnotu k se zpravidla volí konstanta menší, např. 2 nebo 3, pro vnější hodnotu se často volí konstanta 5 nebo 10.

Při optimalizaci hyperparametrů se v kombinaci s křížovou validací často implementuje metoda Grid Search. Grid Search je algoritmus, který pro každou kombinaci hyperparametrů natrénuje a vyhodnotí model a následně na základě vybrané metriky vrátí kombinaci hyperparametrů s nejlepším výsledkem [17].



■ **Obrázek 3.1** Vnořená křížová validace

3.2 Nedoučení a přeučení

K častým komplikacím procesu učení patří nedoučení (angl. underfitting) a přeučení (angl. overfitting). Nedoučení nastává pokud model není schopen zachytit klíčové vzorce z dat a predikuje hodnoty odlišné od skutečných. Model v takovém případě není schopen dosáhnout dostatečně malé chyby na trénovací ani testovací množině (resp. validační množině) [17]. Příčinou nedoučení může být nedostatek trénovacích dat nebo příliš jednoduchá architektura modelu. Ke snížení nedoučení je vhodné při procesu učení nastavit větší množství epoch nebo zvýšit komplexitu modelu přidáním dalších vrstev [12].

K přeučení dochází jestliže se chyba na trénovací množině dat postupně stále snižuje, zatímco chyba na testovací množině (resp. validační) narůstá. To je způsobeno tím, že se model učí vzory

specifické pouze pro trénovací množinu, čímž ztrácí schopnost správně klasifikovat testovací data. Možností jak omezit přeučení modelu je více, patří mezi ně: zjednodušení modelu, augmentace dat nebo využití různých regularizačních technik, kterým se věnuje následující kapitola [12].

3.3 Regularizace

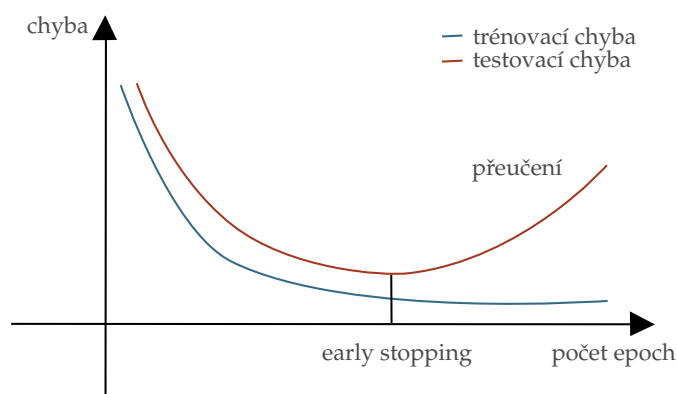
Regularizace je metoda, která se používá při trénování neuronových sítí k redukcí přeučení. Mezi regularizační techniky patří metody L1 (angl. lasso regression), L2 (angl. ridge regression), metoda dropout nebo metoda early stopping.

3.3.1 Dropout

Dropout je regularizační metoda, která částečně předchází přeučení neuronové sítě [12]. Základním principem této metody je deaktivace určitého počtu náhodně vybraných neuronů. Počet deaktivovaných neuronů se nastavuje skrze hodnotu hyperparametru sítě, zn. p . Každý neuron má potom pravděpodobnost p , že bude deaktivován. Obvykle se hodnota hyperparametru pohybuje v rozmezí 0,5 až 0,8, optimální hodnota však závisí na konkrétním případě. Dropout lze přidat ke konvolučním i plně propojeným vrstvám.

3.3.2 Metoda Early Stopping

V průběhu učení trénovací i testovací chyba v ideálním případě klesá. Při určitém počtu epoch však dospěje model do bodu, od kterého začne testovací chyba stoupat (model se přeučuje). Pomocí techniky časného zastavení (angl. early stopping), lze trénování modelu zastavit v bodě, kdy se testovací chyba již dále nesnižuje [12].



■ Obrázek 3.2 Přeučení modelu a early stopping

3.3.3 BatchNorm vrstva

BatchNorm vrstva standardizuje výstup jednotlivých konvolučních vrstev. Zařazení této vrstvy zrychluje konvergenci sítě a umožňuje nastavit vyšší hodnotu kroku učení. Batch normalizace také působí jako forma regularizace, čímž může částečně nahradit jiné formy regularizace, jako například vrstvy dropout [17, 18].

3.3.4 L1 a L2 regularizace

Principem regularizace koeficienty L1 a L2 je přidání penalizace do ztrátové funkce, která zapříčiní snížení vah. V případě L1 regularizace se penalizace rovná absolutní hodnotě vah, v případě L2 součtu druhých mocnin vah (viz 3.2 a 3.1).

$$L_R(y, \hat{y}) = L(y, \hat{y}) + \lambda \sum_{i=1}^n |w_i|, \quad (3.1)$$

$$L_R(y, \hat{y}) = L(y, \hat{y}) + \lambda \sum_{i=1}^n w_i^2, \quad (3.2)$$

kde $L(y, \hat{y})$ je ztrátová funkce, $L_R(y, \hat{y})$ je ztrátová funkce po regularizaci, λ je hyperparametr, w je parametr a n je celkový počet parametrů.

V praxi se častěji využívá L2 regularizace z toho důvodu, že v porovnání s L1 regularizací více penalizuje vysoké hodnoty vah, ale zároveň nízké hodnoty vah neomezuje na nulu [11].

3.4 Krok učení

Hodnota parametrů pro následující epochu se odvozuje od původních hodnot parametrů, gradientu vypočítaném při zpětné propagaci a kroku učení.

Krok učení patří mezi hyperparametry, výrazně ovlivňuje míru, o kterou se v průběhu optimalizace upravují parametry. Pokud je hodnota kroku učení příliš vysoká, hodnoty parametrů se skokově mění, což může zapříčinit přeskočení hledaného minima. Naopak pokud se nastaví hodnota kroku učení jako příliš nízká, parametry se příliš nemění a doba optimalizace se výrazně prodlužuje. Obecně se tak doporučuje nastavit hodnotu kroku učení dostatečně malou, aby model konvergoval k optimálnímu řešení a dostatečně velkou, z důvodu rychlosti učení sítě a problému možného uvíznutí v lokálním minimu [11].

3.5 Architektury konvolučních neuronových sítí

Tato kapitola popisuje architektury vybraných konvolučních neuronových sítí. Vývoj konvolučních neuronových sítí je spojován se soutěží ILSVRC konanou mezi lety 2010 a 2017. Předmětem soutěže byla klasifikace a detekce objektů obrazu z datasetu ImageNet čítajícím přes milion různých obrazů z tisíce klasifikačních tříd. Popsané architektury se mohou využívat i jako architektury výchozí při konstrukci nových.

3.5.1 AlexNet

AlexNet je architektura hlubokého učení z roku 2012, která zpopularizovala využití konvolučních neuronových sítí ve zpracování obrazu. Skládá se z 5 konvolučních vrstev, poolingových vrstev, 3 plně propojených vrstev a 2 dropout vrstev. Aktivační funkcí ve všech vrstvách je funkce ReLU (vyjma poslední FC vrstvy, kde je funkce softmax). Celkový počet parametrů se pohybuje kolem 60 milionů [19].

3.5.2 VGGNet

Architektura VGGNet vychází z architektury AlexNet. Sestává z 16 konvolučních vrstev, poolingových vrstev a 3 plně propojených vrstev s celkovým počtem 138 milionů parametrů [20]. VGGNet se často využívá jako předtrénovaný model k přenesenému učení formou extrakce rysů (více v kapitole 4.2).

3.5.3 ResNet

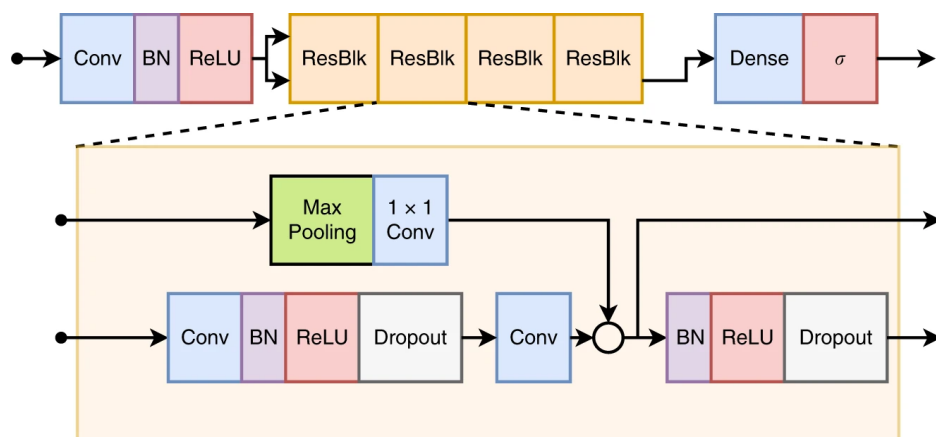
Architektura ResNet oproti předchozím CNN strukturám obsahuje BatchNorm vrstvy a tzv. reziduální bloky [18]. Se zvyšováním počtu vrstev CNN dochází v určitém okamžiku

k výraznému zhoršení přesnosti. Tento problém lze řešit právě pomocí reziduálních bloků zahrnujících tzv. „skip connections“. Skip connections umožňují v rámci reziduálního bloku přeskočit několik vrstev architektury tak, že přeměrují výstup jedné vrstvy na vstup některé vrstvy umístěné v architektuře hlouběji. Architektura ResNet dosáhla v soutěži velmi dobrých výsledků, přesností se vyrovnala člověku [18].

3.5.4 Architektura NN pro klasifikaci 12-svodového EKG

Tato podkapitola se věnuje architektuře konvoluční neuronové sítě z publikace [21]. Model navržený v této publikaci je v praktické části práce využit k přenesenému učení.

Architektura modelu z publikace [21] (viz obrázek 3.3) je odvozena od architektury sítě ResNet, ačkoliv je upravena pro jednodimenzionální vstup. První konvoluční vrstvu následují čtyři reziduální bloky, kdy každý z nich obsahuje dvě konvoluční vrstvy. Velikost konvolučních filtrů byla odvozena na základě optimalizace hyperparametrů na hodnotu 16. Počet konvolučních filtrů je v počátečních vrstvách sítě nastaven na 64, přičemž postupně se navyšuje. Výstup každé konvoluční vrstvy je normalizován prostřednictvím BatchNorm vrstvy. Jako aktivační funkce byla zvolena nelineární aktivační funkce ReLU, po níž následuje vrstva dropout. Skip connections zahrnují poolingovou a konvoluční vrstvu. Na závěr plně propojená vrstva s aktivační funkcí sigmoid klasifikuje záznamy EKG do šesti typů srdečních arytmí.



■ **Obrázek 3.3** Architektura CNN z publikace [21].

3.6 Nevyváženost datasetu

Častým problémem spojeným se zpracováním zdravotních záznamů je nevyváženost rozložení jednotlivých tříd, kdy některé z tříd jsou výrazně méně zastoupeny než jiné. V tomto případě se často jedná právě o pozitivní třídu pacientů diagnostikovaných s daným onemocněním. V důsledku tohoto problému se model může přiklonit k třídě s větším zastoupením, což negativně ovlivní výsledek klasifikace u méně početné třídy [22]. K řešení problému nevyváženého datasetu lze použít různé techniky, jako je oversampling, undersampling nebo přidání vah pro jednotlivé třídy.

Metoda oversampling náhodně vybírá záznamy z minoritní třídy, které duplikuje a zahrnuje do finálního datasetu. Tím zajišťuje, že minoritní třída je zastoupena ve stejném poměru jako třídy ostatní. Nicméně nevýhodou využití této metody je tendence k přeučení.

Dalším přístupem vybalancování tříd prostřednictvím rozšíření datasetu je metoda SMOTE (angl. synthetic minority oversampling technique). Tato metoda je blíže popsána v podkapitole 4.1.1.

Metoda undersampling náhodně vybírá záznamy majoritní třídy, které následně z finálního datasetu vyřazuje tak, aby zastoupení tříd bylo vyrovnané. Hlavní nevýhodou tohoto přístupu je ztráta hodnotných dat, což je obzvláště citelné v případě menšího datasetu.

Jiným přístupem vybalancování datasetu je přidání vah pro jednotlivé třídy. Tato metoda přiřadí dílčím třídám různé váhy, čímž snižuje dopad nevyváženosti datasetu na výsledek klasifikace. Toho lze docílit pomocí modifikace ztrátové funkce, kdy minoritní třídě je přiřazena váha vyšší. To může vést ke zlepšení výsledku klasifikace minoritní třídy, protože model je více penalizován za nesprávnou klasifikaci dat z této třídy.

Přenesené učení a augmentace dat

4.1 Augmentace dat

Učení hlubokých neuronových sítí vyžaduje velké množství dat, nicméně v praxi je počet dostupných dat často omezený. Vlivem nedostatku dat často dochází k přeučení modelu. Augmentace dat (jinak také rozšíření dat) je technika, která se využívá k navýšení velikosti a různorodosti datasetu. Toho lze docílit například kombinováním různých datasetů nebo vygenerováním nových dat odvozených od dat skutečných [23]. Existuje několik různých přístupů k rozšíření dat, následující podkapitoly se věnují augmentaci dat v oblasti zpracování obrazu.

4.1.1 Generativní modely

Technika SMOTE problém nevyváženého datasetu řeší vygenerováním syntetických vzorků méně zastoupené třídy. Algoritmus SMOTE nejdříve určí k -nejbližších sousedů vybraného obrázku. Poté mezi určitým počtem z nich a vybraným obrázkem vygeneruje prostřednictvím lineární interpolace nový obrázek [24].

Dalsí technikou využívanou ke generování syntetických dat je model strojového učení GAN (angl. generative adversarial network). Model GAN se skládá ze dvou neuronových sítí, a to generátoru a diskriminátoru. Obě sítě se trénují společně, zatímco generátor je zodpovědný za generování nových, pokud možno co nejvíce realistických dat. Úlohou diskriminátoru je správně určit, která data jsou skutečná a která jsou vygenerovaná [25].

4.1.2 Transformace obrazu

Jiným přístupem rozšíření dat je aplikace jednoduchých transformací na skutečný obrázek, jako je otočení, změna velikosti obrázku, přidání šumu nebo převrácení. Transformace dat může být účinným nástrojem ke zvýšení rozmanitosti datové sady a odolnosti modelu vůči rozdílům v rozložení dat, je však třeba dbát opatrnosti při volbě typu a míry transformace. V případě, že je transformace původních dat nevýrazná, transformované obrázky se příliš podobají původním a účinnost augmentace při učení modelu nebude signifikantní. Naopak příliš výrazná transformace obrázku může vyústit k vygenerování nereálných případů. Ačkoliv transformace obrazu patří mezi velmi užitečné přístupy rozšíření datasetu při klasifikaci obrázků, v případě klasifikace jednodimenzionálních fyziologických signálů, jako je signál EKG, jsou možnosti transformace značně omezené [23]. Následující podkapitola popisuje příklady transformací signálu EKG inspirovaných klasickými transformacemi obrazu [23, 26, 27].

4.1.2.1 Transformace signálu EKG

Gaussovský šum

Častou technikou augmentace dat je přidání Gaussovského šumu, jež může simulovat skutečný šum nefiltrovaného signálu [27]. Přidání Gaussovského šumu lze realizovat vygenerováním vektoru náhodných čísel z Gaussovského rozdělení s definovanou střední hodnotou μ a hodnotou směrodatné odchylky σ , která reprezentuje množství přidaného šumu.

Škálování

Škálování je běžnou metodou geometrické transformace při zpracování obrazu. Měřítko signálu EKG může být ve skutečnosti ovlivněno vodivostí senzoru, škálování proto patří mezi validní transformace signálu a lze jej aplikovat při augmentaci dat [27]. Škálování EKG lze realizovat vynásobením původního signálu škálovacím faktorem.

Drift

Kolísání nulové izolinie (drift) patří mezi časté artefakty při měření EKG způsobené dýcháním, pohybem nebo špatným kontaktem elektrod s kůží. K potlačení driftu se používají různé filtry, jako například pásmový filtr (angl. bandpass filter) [26]. Drift lze simulovat několika způsoby, v praktické části práce se drift implementuje podle [26] tak, že se k původnímu signálu přičte funkce kosinus s náhodně zvolenou amplitudou, frekvencí a horizontálním posunem.

Dropout

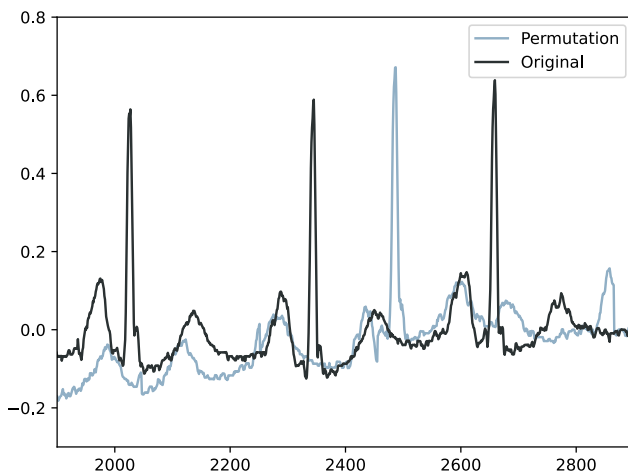
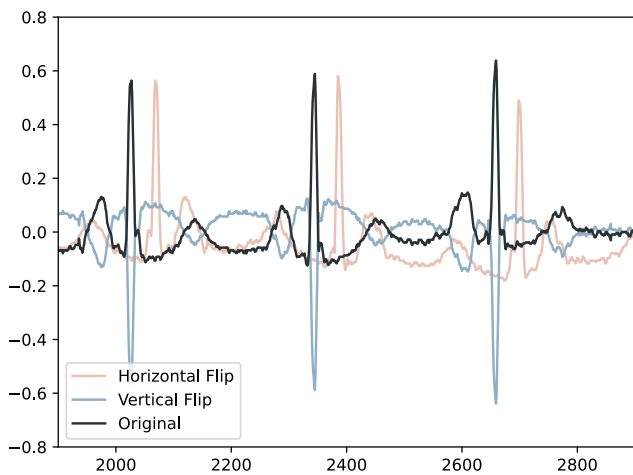
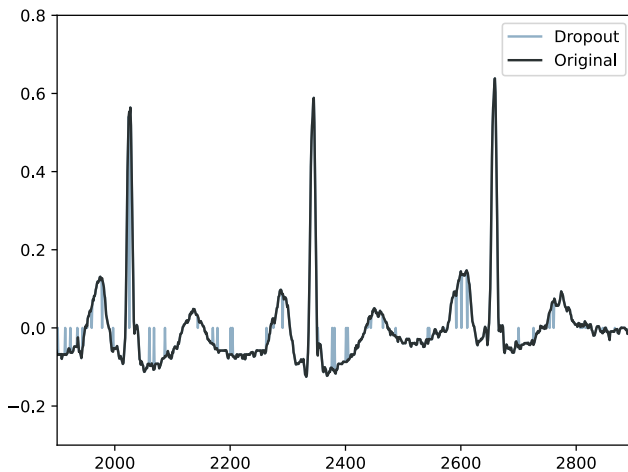
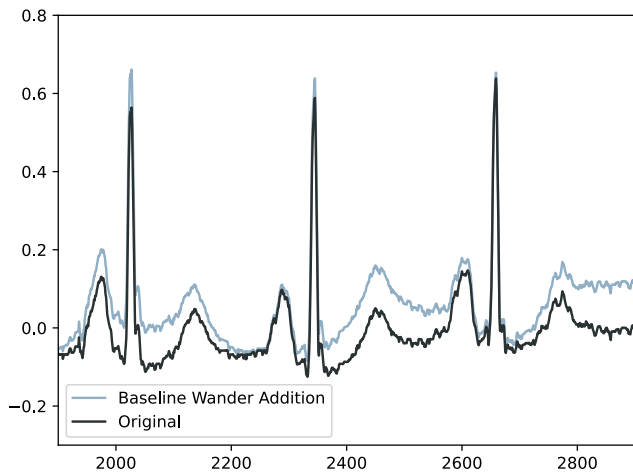
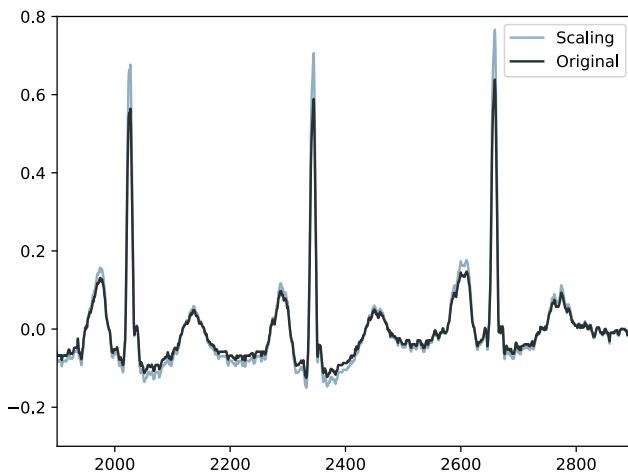
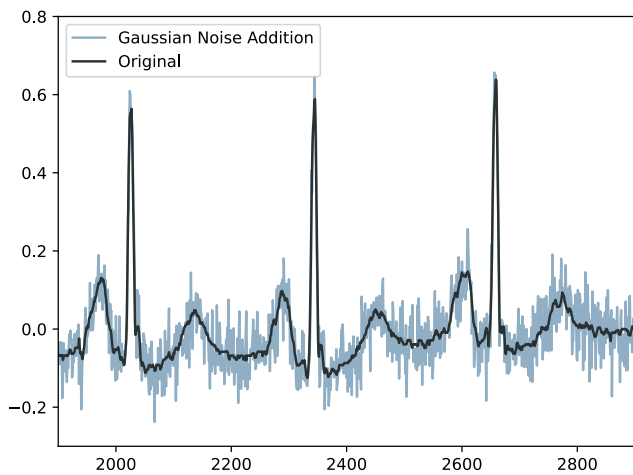
Tato metoda transformace signálu je taktéž inspirovaná transformací obrazu, kdy se náhodně zvolený počet pixelů obrazu nastavuje na hodnotu 0. V případě transformace signálu je princip této metody obdobný, tedy určité procento vstupních vzorků se přepisuje na nulu. V důsledku toho se síť učí být robustnější vůči malým změnám vstupu, což může zlepšit její schopnost zobecnění [26]. Název této metody je odvozený od regularizační techniky dropout (viz 3.3.1)

Zrcadlení

Zrcadlení je běžnou technikou augmentace dat především u symetrických obrázků. Obraz lze zrcadlit podle vodorovné osy (vertikální zrcadlení) nebo svislé osy (horizontální zrcadlení). Zrcadlení výrazně mění průběh signálu, model naučený na datasetu zahrnujícím takto upravené obrázky dosahoval dokonce horších výsledků než dataset původní [26, 27].

Permutace

Také permutace patří mezi transformace obrazu, které výrazně mění průběh signálu EKG. Principem této transformace je rozdělení signálu do m podřetězců x_1, x_2, \dots, x_m , jejich náhodné promíchání a spojení. Volba hodnoty m je zásadní, pokud je hodnota příliš malá, transformace nemusí být dostačující, pokud je naopak příliš velká, signál se podstatně deformuje a model nedokáže z dat rozpoznávat požadované vzorce [23].



■ **Obrázek 4.1** Způsoby transformace původního signálu EKG.

4.2 Přenesené učení

Přenesené učení (angl. transfer learning, TL) je v praxi často používanou metodou, kdy se model předtrénovaný pro jednu úlohu upraví a aplikuje na jinou úlohu podobného typu. Tento přístup může být užitečný obzvláště v případě, kdy je k dispozici pouze omezené množství dat. Model se v první fázi učí rozpoznávat nízkourovňové prvky na obecnějším, rozsáhlém datasetu, například datasetu ImageNet. Také je možné využít existující model s patřičnými hodnotami parametrů, jako např. VGGNet [20] nebo ResNet [18]. Ve druhé fázi se model upraví a dotrénuje na cílovém datasetu podle některého z následujících přístupů.

Trénování pouze plně propojených vrstev

Vrstvy předtrénovaného modelu jsou tzv. „zmrazené“ (angl. frozen), to znamená, že původní parametry zůstávají nezměněné. Plně propojené vrstvy předtrénovaného modelu jsou nahrazené novými. Následně je model přetrénován na cílovém datasetu. Vzhledem k tomu, že se mění parametry pouze u plně propojených vrstev, trénování probíhá relativně rychle, nicméně je náchylné na přeučení. Tento přístup se označuje také jako feature extraction.

Přetrénování části konvolučních vrstev

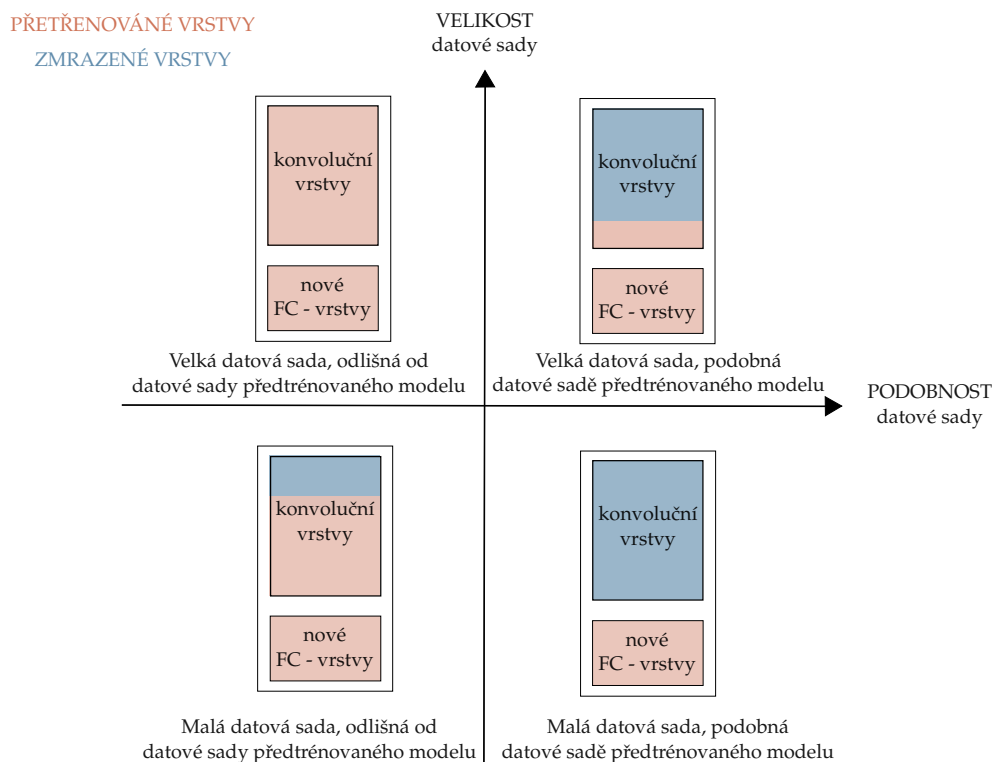
Tento přístup opět zahrnuje nahrazení plně propojených vrstev předtrénovaného modelu, rozdíl oproti předchozímu přístupu je, že se zmrazí pouze některé vrstvy modelu. Zpravidla je vhodné zmrazit počáteční vrstvy, které se učí nízkourovňové prvky a několik posledních konvolučních a poolingových vrstev přetrénovat společně s novou klasifikační částí. Tento přístup se označuje také jako fine-tuning.

Přetrénování celého modelu

Tento přístup je vhodný v případě, že se cílový dataset podstatně liší od původního. Parametry předtrénovaného modelu mohou sloužit jako iniciační při přetrénování. I v tomto případě se nahrazuje klasifikační část sítě, zároveň je vhodné zvolit nižší hodnotu kroku učení, aby nedošlo ke ztrátě předchozích informací. Nevýhodou tohoto přístupu je značná výpočetní náročnost.

Přetrénování s následným doladěním

Tato strategie se skládá ze dvou fází, v první fázi se nahradí plně propojené vrstvy sítě, ostatní vrstvy předtrénovaného modelu se zamrazí a model se natrénuje na cílovém datasetu s dostatečně velkým počtem epoch a standardní hodnotou kroku učení. Ve druhé fázi procesu se celý model „rozmrází“ a přetrénuje v rámci malého počtu epoch s řádově menší hodnotou kroku učení. Výhodami tohoto dvoufázového přístupu je, že v první fázi díky zmrazení vrstev nedochází k přepisování původních parametrů vlivem zpětného šíření chyby. Ve druhé fázi se následně parametry doladí, což může zlepšit celkovou úspěšnost modelu.



■ **Obrázek 4.2** Volba strategie přeneseného učení v závislosti na podobnosti a velikosti datasetu.

Přenesené učení je široce využívanou technikou v různých oblastech aplikací strojového učení, včetně klasifikace obrazu, zpracování přirozeného jazyka nebo rozpoznávání řeči. Aplikací přeneseného učení ve zdravotnictví, především ve zpracování signálu EKG, se věnuje následující kapitola.

Vyhodnocení úspěšnosti modelu

Výslednou úspěšnost modelu lze měřit na nezávislém souboru dat prostřednictvím různých evaluačních metrik. Tato kapitola se zaměřuje na nejčastěji používané metriky v klasifikačních úlohách.

5.1 Binární klasifikace

Výstup klasifikace může být pozitivní v případě, že vstupní prvek patří do klasifikační třídy nebo negativní v případě, že tam nepatří. Na základě výstupu algoritmu - predikovaných hodnot a hodnot skutečných, se rozlišují následující případy:

- Skutečně pozitivní (angl. true positive, dále jen TP) – algoritmus predikoval vstupní hodnotu jako pozitivní a vstupní hodnota do třídy ve skutečnosti patří.
- Skutečně negativní (angl. true negative, dále jen TN) – algoritmus predikoval vstupní hodnotu jako negativní a vstupní hodnota do třídy ve skutečnosti nepatří.
- Falešně negativní (angl. false negative, dále jen FN) – algoritmus predikoval vstupní hodnotu jako negativní, ale vstupní hodnota do třídy ve skutečnosti patří.
- Falešně pozitivní (angl. false positive, dále jen FP) – algoritmus predikoval vstupní hodnotu jako pozitivní, ale vstupní hodnota do třídy ve skutečnosti nepatří.

Matice záměn

Matice záměn zobrazuje ve sloupcích počet predikovaných hodnot označených jako pozitivní (resp. negativní) a v řádcích počet hodnot skutečně pozitivních (resp. negativních).

		predikované hodnoty	
		pozitivní	negativní
skutečné hodnoty	pozitivní	TP	FN
	negativní	FP	TN

■ **Obrázek 5.1** Matice záměn

Přesnost a preciznost

Přesnost (angl. accuracy) definuje podíl všech správně klasifikovaných hodnot oproti celkovému počtu vstupních dat.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.1)$$

Preciznost (angl. precision) stanovuje podíl skutečně pozitivních hodnot oproti všem hodnotám klasifikovaným jako pozitivní.

$$P = \frac{TP}{TP + FP} \quad (5.2)$$

Senzitivita a specificita

Senzitivita (angl. sensitivity, recall) určuje podíl hodnot, které algoritmus správně označil jako pozitivní oproti všem vstupním hodnotám, které jsou pozitivní ve skutečnosti.

$$R = \frac{TP}{TP + FN} \quad (5.3)$$

Specificita (angl. specificity) určuje podíl hodnot, které algoritmus správně označil jako negativní oproti všem vstupním hodnotám, které jsou negativní ve skutečnosti.

$$S = \frac{TN}{TN + FP} \quad (5.4)$$

F1-skóre

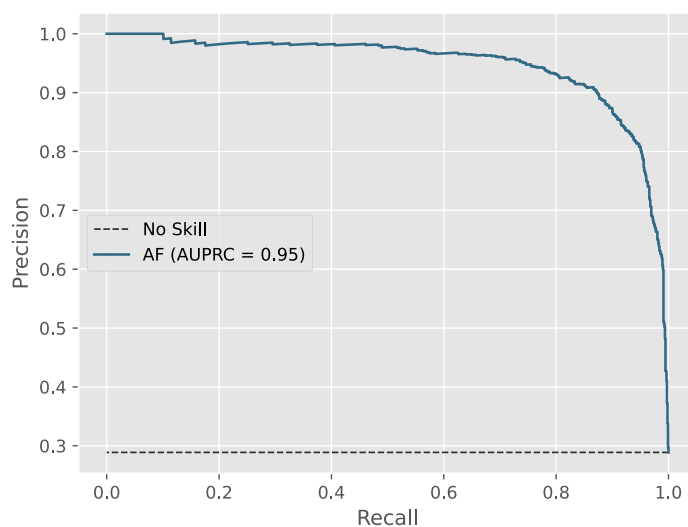
F1-skóre (angl. F1-score) definuje harmonický průměr preciznosti a senzitivity, v ideálním případě by se hodnota F1-skóre měla blížit hodnotě 1. F1-skóre má vypovídající hodnotu i v případě nerovnoměrného rozdělení tříd. Je tak vhodnou volbou metriky v případě nevyváženého datasetu.

$$F = 2 \times \frac{P \times R}{P + R} \quad (5.5)$$

PRC křivka a AUPRC

PRC křivka (angl. precision-recall curve) popisuje závislost preciznosti na senzitivě a je obzvláště užitečná v případě nevyváženého datasetu. Prostřednictvím PRC lze optimalizovat threshold tj. prahová hodnota, pro kterou platí, že když je predikovaná pravděpodobnost vstupního vzorku vyšší než threshold, vzorek se označí jako pozitivní.

AUPRC (angl. area under the precision-recall curve) je hodnota reprezentující úspěšnost klasifikace na základě výpočtu oblasti pod PRC křivkou. V ideálním případě by se její hodnota měla blížit hodnotě 1.



■ Obrázek 5.2 PRC křivka

..... Kapitola 6

Výzkumy v oblasti související s tématem této práce

Vzhledem k asymptomatickému průběhu fibrilace síní a riziku rozvoje závažných komplikací je včasná diagnostika pro pacienty klíčová. Využití algoritmů strojového učení umožňuje detekovat srdeční arytmie v téměř reálném čase, čímž může nemoc včas odhalit a zároveň asistovat lékařům při diagnostice. Signály EKG jsou navíc zatíženy množstvím artefaktů, šumem a výraznou variabilitou, proto může být užitečné v této oblasti aplikovat algoritmy strojového učení, které jsou vhodné k identifikaci vzorů z komplexních dat. V některých aspektech mohou dokonce dosahovat vyšší úspěšnosti při klasifikaci arytmií v porovnání s lékaři [21]. Tato kapitola se věnuje dosavadnímu výzkumu v oblasti detekování srdečních arytmií pomocí metod strojového učení.

6.1 Detekce arytmií pomocí extrakce rysů a základních algoritmů strojového učení

V souvislosti s diagnostikou fibrilace síní se lze setkat s tzv. Poincarého mapou (angl. Poincaré plot). Poincarého mapa je snadno interpretovatelné grafické zobrazení, které reprezentuje závislost délky intervalu RR na délce předchozího intervalu RR. Zatímco v případě fyziologického signálu se jednotlivé body shlukují do jednoho místa, pro záznamy fibrilací síní je typické nerovnoměrné rozložení s větším počtem shluků.

Metoda podpůrných vektorů (angl. support vector machine, dále jen SVM) patří mezi algoritmy strojového učení. Principem SVM je sestavit nadrovinu v n -dimenzionálním prostoru, která efektivně oddělí data náležící odlišným třídám tak, aby vzdálenost nadroviny a krajních bodů každé z tříd byla maximální. V případě binární klasifikace může jít o lineární nebo nelineární přímku.

Klasifikace arytmií prostřednictvím SVM může dosahovat dobrých výsledků, je však třeba, aby klasifikaci předcházela extrakce rysů. V publikaci [28] SVM klasifikuje srdeční arytmie (včetně fibrilace síní) z předem extrahovaných rysů, v tomto případě RR intervalů. Hlavní výhodou tohoto přístupu je robustnost algoritmu vůči zpracování signálu s větším šumem, jež obzvláště patří mezi komplikace při klasifikaci záznamů z nositelných zařízení. Jako nevýhodu lze považovat nutnost samostatné extrakce rysů a obecnost aplikování této metody na jiné typy arytmií, které se nevyznačují abnormalitami v oblasti RR intervalů.

Při klasifikaci arytmií lze SVM použít i v kombinaci s neuronovými sítěmi. V publikaci [29] se klasifikační část konvoluční neuronové sítě AlexNet zamění za SVM. Takto upravená síť na vstupu nepracovává 1D signály EKG, nýbrž spektrogramy zachycující frekvenční složku signálu.

6.2 Detekce arytmií pomocí konvolučních neuronových sítí

Podstatnou nevýhodou aplikování doposud zmíněných algoritmů je nezbytnost „manuální“ extrakce rysů, tedy i větší expertizy v dané lékařské problematice. Oproti tomu v případě neuronových sítí je extrakce rysů realizována automaticky v průběhu učení algoritmu. Konvoluční neuronové sítě patří k nejpoužívanějším algoritmům strojového učení v doméně klasifikace srdečních rytmů.

V publikaci [30] se porovnává úspěšnost klasifikace jednodimenzionální CNN a dvoudimenzionální CNN. Obě sítě vychází z architektury AlexNet, skládají se z pěti konvolučních vrstev, tří poolingových a tří plně propojených vrstev doplněných vrstvami dropout. Výsledkem práce je mimo jiné porovnání různých aktivačních funkcí, kdy překvapivě největší úspěšnosti dosáhla síť s aktivační funkcí Swish, přestože se v podobné problematice dává přednost aktivační funkci ReLu. Z výsledků práce je také zřejmé, že 1D síť si v porovnání s dvoudimenzionální vedla při klasifikaci arytmií na 1-svodovém EKG lépe.

V publikaci [43] se jednosvodové EKG záznamy z veřejně dostupného datasetu dělí do 4 klasifikačních tříd, a to FS, SR, nekvalitní záznamy a jiné rytmy. Architektura neuronové sítě byla v tomto případě navržena „od základu“, skládá se z několika střídajících se konvolučních a poolingových vrstev, BarchNorm vrstvy a dropoutu. Model navržený v této publikaci nedosahuje tak vysokých výsledků jako řešení v jiných publikacích, na rozdíl od nich je však výrazně méně komplexní.

Velmi dobrých výsledků klasifikace FS (99,7 % F1-skóre) dosahuje CNN AFibNet navržená v publikaci [44]. Model navržený v této publikaci je vytrénovaný na datasetu sjednoceném z veřejných datasetů s jednosvodovými záznamy. Data byla před zpracováním neuronovou sítí segmentována do více kratších úseků. Z pohledu architektury se CNN skládá pouze z konvolučních, poolingových a plně propojených vrstev.

Hlavním přínosem práce [21] je návrh architektury CNN, která dosahuje velmi dobrých výsledků i při zpracování 12-svodových EKG záznamů. Neuronová síť je založena na architektuře ResNet, je ovšem upravená pro jednodimenzionální signál a trénovaná na 2 milionech anotovaných dat. Model z této publikace dosahuje při rozpoznávání šesti typů arytmií F1-skóre nad 80 % a specificty vyšší než 99 %.

6.3 Augmentace dat u EKG záznamů

K úskalím spojeným s využitím konvolučních neuronových sítí patří vysoká výpočetní náročnost a riziko přeučení sítě. To lze částečně řešit různými přístupy regularizace nebo rozšířením datasetu.

V publikaci [26] se pomocí neuronové sítě ResNet porovnává vliv augmentace prostřednictvím 15 různých transformací obrazu a 6 filtrů na klasifikaci srdečních arytmií. Výsledky této práce popisují signifikantní nárůst úspěšnosti klasifikace v porovnání s neupraveným datasetem. K nejvíce efektivním transformacím podle [26] patří kolísání nulové izolinie, dropout a škálování. V publikaci byla také navržena modifikovaná verze metody RandAugment, která náhodně kombinuje několik transformací dohromady.

Autoři v knize [23] shrnují několik přístupů augmentace dat a vyhodnocují jejich vliv na úspěšnost klasifikace FS na datasetu PhysioNet prostřednictvím LSTM sítě (angl. long short-term memory, více o LSTM v podkapitole 6.5). Na základě porovnání autorů dosahuje nejlepších výsledků síť natrénovaná na datasetu rozšířeném metodou permutace nebo rozšířením metodou opakovaného spojování subsekvencí s charakteristickými body [23, 31]. Obě dvě tyto metody nicméně narušují pravidelnost intervalů RR, která je z pohledu lékařů při diagnostice klíčová.

Mezi časté techniky rozšíření dat patří metoda SMOTE. V publikaci [32] autoři pracují s databází MIT-BIH, kterou, vzhledem k nerovnoměrnému rozložení daných typů srdečních rytmů, vyvažují prostřednictvím metody SMOTE. Ke klasifikaci autoři zvolili CNN s 11 vrstvami, jež na

datasetu rozšířeném metodou SMOTE dosahuje přesnosti 98,3 %, v porovnání s původním datasetem, kde CNN dosahuje přesnosti 96,0 %. Nutno ovšem podotknout, že autoři nezpracovávají celé záznamy najednou, ale při předzpracování je segmentují.

6.4 Detekce arytmií pomocí přeneseného učení

V případě přeneseného učení je žádoucí využít již předtrénovaný model na rozsáhlejším datasetu, v praxi se využívají např. modely předtrénované na datasetu ImageNet. To ovšem v tomto případě není optimální, vzhledem k tomu, že se datové sady výrazně liší. I přesto některé publikace poukazují na zlepšení úspěšnosti klasifikace EKG při využití modelu předtrénovaného na datasetu ImageNet.

V publikaci [33] se k přenesenému učení formou extrakce rysů používá neuronová síť DenseNet předtrénovaná na datasetu ImageNet. Signály EKG jsou převedeny na spektrogramy, prostřednictvím DenseNet se extrahují vzorce, které se na závěr klasifikují skrze SVM do čtyř tříd srdečních rytmů.

V publikaci [30] se k přenesenému učení formou ladění používá dvoudimenzionální neuronová síť AlexNet předtrénovaná na datasetu ImageNet. Data zahrnutá v této publikaci pochází z veřejně dostupné databáze MIT-BIH.

V publikaci [34] se v prvním kroku neuronová síť ResNet předtrénovává na veřejně dostupném datasetu Icentia11k s 1-svodovými záznamy EKG. V dalším kroku se neuronová síť upraví a přetrénuje na dostupném datasetu s 12-svodovými záznamy ke klasifikaci fibrilace síní. Výsledkem této práce je zvýšení F1-skóre o necelých 7 % v důsledku přeneseného učení.

Mezi jiné přístupy klasifikace s nedostatečným množstvím dat patří aktivní učení nebo semi-supervizované učení [35, 36].

6.5 Detekce arytmií pomocí rekurentních neuronových sítí

Rekurentní neuronové sítě (RNN) dosahují výborných výsledků při zpracovávání sekvenčních dat, například v oblastech jako je zpracování přirozeného jazyka. RNN se vyznačují cyklickým propojením vrstev, které neberou v úvahu pouze aktuální vstup, ale také vstup předchozí. Mezi typy rekurentních neuronových sítí patří síť LSTM. Struktura LSTM umožňuje dlouhodobé uchování informací, LSTM sítě proto mohou být obzvláště užitečné v případě zpracování dlouhých časových řad, mezi něž lze řadit i signál EKG. Ke klasifikaci arytmií prostřednictvím RNN, nejčastěji pak pomocí LSTM sítí, lze přistupovat různými způsoby. Častým přístupem je rozdělení celého záznamu EKG do menších celků. Jednotlivé úseky se následně přímo zpracovávají skrze LSTM [37, 38] nebo slouží k další extrakci rysů, například délky QRS komplexu nebo amplitudy vrcholů Q a R. Takto extrahované vzory poté slouží jako vstup pro následnou klasifikaci prostřednictvím různých typů RNN, případně jejich seskupení [38, 39].

Jiným přístupem detekování arytmií může být kombinace konvolučních neuronových sítí s LSTM [40, 41]. V tomto případě je prvotní extrakce rysů realizována prostřednictvím konvolučních a poolingových vrstev modelu, výstup následně prochází vrstvou LSTM, která uchovává dlouhodobé závislosti mezi extrahovanými daty a plně propojenou vrstvou, která srdeční rytmy klasifikuje do příslušných tříd.

Kapitola 7

Vlastní řešení

Následující část práce se věnuje vlastnímu řešení využití CNN ke klasifikaci srdečních arytmií, konkrétně detekování té nejčastější z nich, fibrilaci síní. Cílem této práce bylo navrhnout model strojového učení pro automatickou detekci fibrilace síní, i při omezeném množství dat. Lékaři anotovaná data použitá v této práci byla poskytnuta Institutem klinické a experimentální medicíny (IKEM). Jedná se o 12-svodové EKG záznamy. 12-svodové EKG záznamy byly vybrány z důvodu, že poskytují komplexnější pohled na srdeční arytmiie a jsou tak primárním nástrojem pro diagnostiku těchto onemocnění v klinické praxi.

Jak již bylo zmíněno v teoretické části práce, k vytrénování spolehlivé neuronové sítě je třeba pracovat s velkým množstvím dat. Nicméně pro účely práce bylo k dispozici pouze omezené množství anotovaných záznamů, a to v řádech tisíců. Nabízí se několik možných alternativ jak k tomuto problému přistupovat. V této práci je problém omezeného počtu dat adresován použitím přeneseného učení.

Vzhledem k tomu, že se dataset ImageNet podstatně liší od datové sady čítající signály EKG, není optimální použít k přetrénování model naučený na takovém datasetu (např. VGGNet nebo ResNet). V publikaci [34] se k předtrénování využívá veřejně dostupný dataset, díky čemuž se klasifikace fibrilace síní zlepšuje o jednotky procent. Ovšem nevýhodou využití veřejně dostupných datasetů je, že se ve většině případů jedná o záznamy s menším počtem svodů.

Velmi dobrých výsledků v oblasti klasifikace srdečních arytmií dosáhl model navržený v publikaci [21]. Architektura tohoto modelu vychází z architektury ResNet, při trénování autoři pracovali s anotovaným datasetem čítajícím více než 2 miliony 12-svodových EKG záznamů. Vzhledem k velikosti, podobnosti datasetu, úspěšnosti modelu a jeho dostupnosti byl tento model v rámci mé práce zvolen jako výchozí k následnému přetrénování.

V průběhu práce bylo třeba čelit několika problémům, tím největším byl nedostatek a nevyváženost dat. Práce proto implementuje i některé techniky augmentace dat, konkrétně transformace signálu EKG. Výběr konkrétních transformací byl založen na výsledcích publikací týkajících se podobné problematiky.

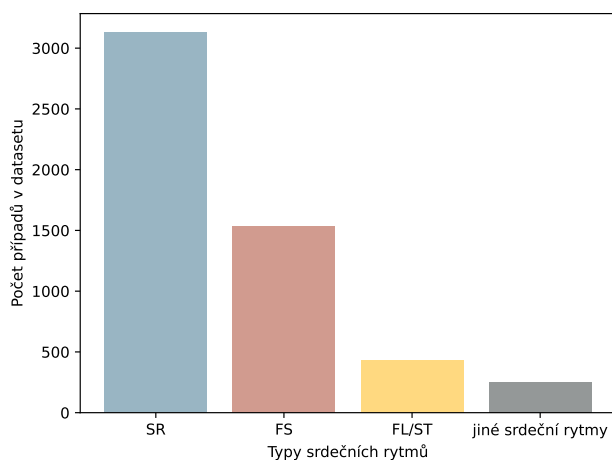
Díličí experimenty byly realizovány na notebooku Lenovo Ideapad 5 Pro s procesorem AMD Ryzen 7 5800H s 8 jádry a frekvencí 3,2 GHz, integrovanou grafickou kartou AMD Radeon™ Graphics a 16 GB RAM.

Analýza dat a příprava datasetu

Data použitá v této práci byla poskytnutá Institutem klinické a experimentální medicíny (IKEM). Jedná se o záznam 12-svodového EKG o vzorkovací frekvenci 500 Hz. Záznamy pochází z časového období listopad 2019 – duben 2020. V IKEM se neanotované záznamy EKG uchovávají v digitální formě prostřednictvím celonemocničního systému MUSE společnosti GE Healthcare. Anotace k EKG záznamům se uchovávají samostatně v csv formátu.

Pro vytvoření finálního datasetu bylo třeba přiřadit poskytnuté anotace k exportovaným EKG záznamům. Vzhledem k tomu, že jednotlivé anotace nezahrnovaly identifikátor záznamu, přiřazení nebylo vždy jednoznačné. Přiřazení anotací k příslušným záznamům bylo z toho důvodu provedeno na základě čísla pacienta, data a času záznamu. V některých případech anotacím odpovídalo více EKG záznamů, v takovém případě bylo nutné manuálně zkontrolovat, který ze záznamů k anotaci patří. Nejčastěji se tak stalo v důsledku opakovaného měření, kdy prvotní měření nebyla dostatečně kvalitní. Ve sporných případech, kdy nebylo možné příslušný záznam přiřadit, anotace nebyla zahrnuta do finálního datasetu.

Výsledný dataset obsahuje celkem 5345 záznamů, z toho 58,5 % tvoří případy s normálním sinusovým rytmem. Z arytmií patří mezi nejpočetnější fibrilace síní 28,7 % a flutter síní příp. sinusová tachykardie 8,1 %. Ačkoliv anotace dat rozlišují více typů arytmií, v této práci je stěžejní pouze rozlišení fibrilace síní od ostatních rytmů.



■ Obrázek 8.1 Zastoupení jednotlivých typů srdečního rytmu v datasetu.

Exportované signály EKG jsou již částečně předzpracované pomocí filtrů a korekce nulové izolinie. Všechny záznamy byly podvzorkovány na vzorkovací frekvenci 400 Hz, a to ze dvou důvodů, prvním je zmenšení velikosti vstupních dat a druhým důvodem je nastavení stejného rozměru vstupních dat jako je očekávaný vstup předtrénovaného modelu v kapitole 10.

Data byla rozdělena na dvě klasifikační třídy, a to třídu pozitivní (fibrilace síní) a třídu negativní (ostatní srdeční rytmy, tedy SR a jiné typy arytmií). Celkově dataset tvoří z 28,7 % třída pozitivní a 71,3 % třída negativní. Dataset je tedy poměrně nevyvážený, při rozdělení dat na trénovací, validační a testovací podmnožinu bylo třeba dbát na dodržení poměrového zastoupení tříd.

V průběhu psaní práce byl poskytnut navíc z IKEM soubor dat obsahující 1793 záznamů. Protože byl tento dataset dodán později, sloužil pouze k sekundární validaci výsledku. Dataset pochází z pozdějšího časového období, a to duben 2022 – říjen 2022, navíc od primárního datasetu se liší v rozložení jednotlivých tříd.

Základní model a jeho optimalizace

Tato kapitola se věnuje návrhu a optimalizaci vlastního, základního modelu. Vzhledem k velikosti datasetu nelze očekávat při klasifikaci příliš vysokou úspěšnost, nicméně tento model byl dále využit k širšímu otestování metod vybalancování datasetu a augmentace dat.

9.1 Architektura a optimalizace hyperparametrů

Ke klasifikaci byla navržena konvoluční síť o 16 vrstvách (dropout se v tomto případě označuje také jako vrstva). Finální konfigurace byla vybrána na základě rozsáhlé optimalizace hyperparametrů, implementované prostřednictvím k -násobné křížové validace, kde $k = 3$ a softwarového nástroje wandb. Počet epoch byl po zvážení průběhu prvního běhu nastaven na 20. Jednotlivé konfigurace byly porovnány na základě validační chyby a průměrné hodnoty AUPRC. V případě, že se při trénování validační chyba po dobu pěti epoch nesnížila, hodnota kroku učení se zmenšila desetkrát.

Předmětem optimalizace byla velikost filtrů konvolučních vrstev {16, 32, 64}, hodnota batch size {16, 32, 64}, počáteční hodnota kroku učení {0,01, 0,001, 0,0001}, volba optimalizační metody {adam, SGD} a hodnota dropout vrstvy {0, 0,25, 0,5, 0,8}. V rámci optimalizace byly také porovnány různé regularizační techniky, kromě dropout vrstev také BatchNorm vrstvy nebo L2 regularizace. Graf v příloze A shrnuje některé vybrané výsledky pro dílčí konfigurace.

Z porovnání optimalizace měla největší vliv na úspěšnost modelu velikost konvolučního filtru a hodnota dropout vrstev. Finální konfigurace (viz obrázek 9.1) byla v další kapitole využita při vyhodnocení různých přístupů augmentace dat a vybalancování tříd.



■ Obrázek 9.1 Výsledná architektura modelu.

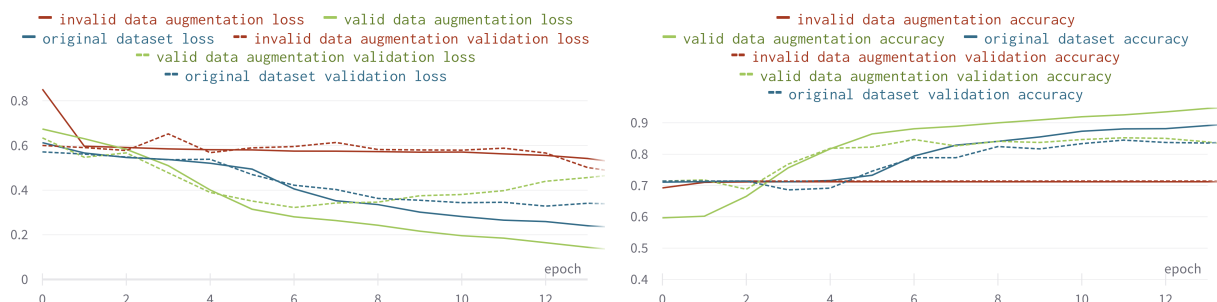
9.2 Augmentace dat a vybalancování datasetu

Vzhledem k tomu, že použitý dataset je menšího rozsahu a navíc třída negativních případů poměrně přesahuje třídu pozitivních případů, stálo za zvážení pokusit se data rozšířit pomocí techniky augmentace dat a zastoupení jednotlivých případů vyrovnat pomocí některé z metod vybalancování dat, jako je resampling nebo přidání vah. Univerzální přístup, jakým toho docílit není, záleží na volbě klasifikátoru i vlastnostech dat. Proto jedním z cílů práce bylo jednotlivé přístupy porovnat.

Augmentace dat

K transformování původních EKG signálů byly implementovány tři techniky augmentace dat, a to škálování, Gaussovský šum a drift viz obrázky 4.1. Škálování bylo implementováno pro kompresi (případně amplifikaci) signálu náhodně zvoleným škálovacím faktorem z intervalu $[0,1, 1,5)$. Gaussovský šum byl implementován za účelem přidání šumu s náhodně zvolenou střední hodnotou σ v intervalu $[0,025, 0,1)$. Přidání driftu bylo realizováno prostřednictvím kosinové funkce s náhodně zvolenou amplitudou (z intervalu hodnot $[0,02, 0,1)$), délkou intervalu kosinové funkce $\{5\pi, 10\pi, 20\pi\}$ a horizontálním posunem (z intervalu $[0, 2\pi)$).

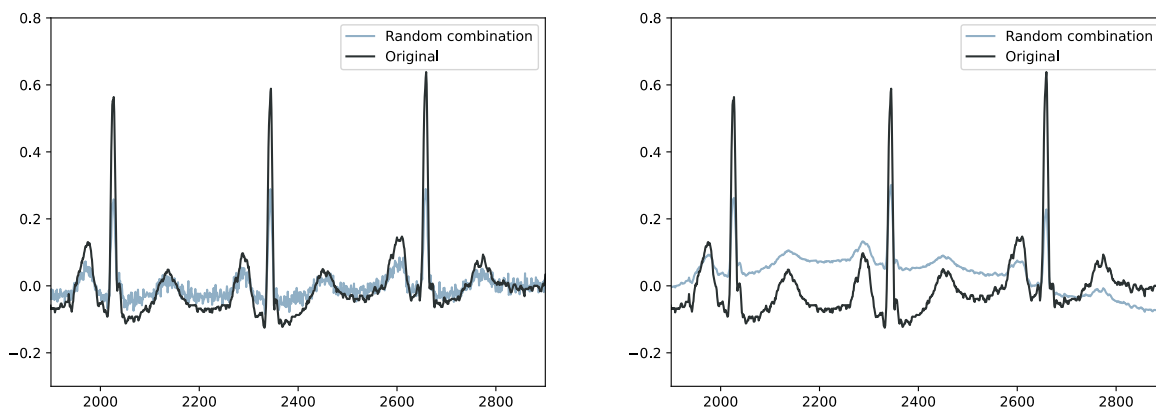
Zmíněné intervaly částečně vycházejí z publikací, i přesto bylo třeba všechny hodnoty ověřit a případně upravit v závislosti na vlastním datasetu. V první fázi byla transformace s dílčími hodnotami aplikována na trénovací podmnožinu datasetu, po vytrénování modelu byl efekt transformace ověřen na validační sadě, a to v rámci několika nezávislých běhů.



■ **Obrázek 9.2** Ukázka průběhu křivky učení při nevhodné volbě hodnot transformací (červeně) v porovnání s původním datasetem (modře) a rozšířeným datasetem (zeleně). Plná čára značí trénovací přesnost (resp. chybu), přerušovaná čára značí validační přesnost (resp. chybu).

Závěrem lze říci, že vybrané transformace s intervaly zásadně nenarušují trénování modelu. V další části práce se již pracuje s náhodným výběrem z těchto transformací (případně jejich kombinací) s náhodnou volbou patřičných hodnot, jako je škálovací faktor nebo amplituda kosinové funkce.

Obrázek 9.3 zobrazuje dva příklady náhodné kombinace transformace signálu. Na obrázku vlevo se k původnímu signálu v první fázi přidal Gaussovský šum, kde $\sigma = 0,03$. Následně byl signál škálovaný faktorem 0,44. Na obrázku vpravo byl původní signál v první fázi škálovaný faktorem 0,39. Následně byl k takto modifikovanému signálu přidán drift s délkou intervalu 20π , amplitudou 0,02 a horizontálním posunem $1,25\pi$



■ **Obrázek 9.3** Ukázka náhodné kombinace transformací signálu.

Vybalancování datasetu

Původní množina dat (vyjma testovací podmnožiny) obsahovala necelých 30 % případů fibrilace síní. Zbytek případů tvořily sinusový rytmus a jiné typy arytmií shrnuté do jedné třídy.

Vybalancování tříd bylo realizováno prostřednictvím různých přístupů, mezi ně patří metody oversampling, undersampling a metoda SMOTE. Tyto přístupy byly dále porovnány s částečným vybalancováním pomocí augmentace dat. V tomto případě byla všechna data třídy skutečně pozitivní (záznamy s FS) použita k augmentaci, zatímco z třídy negativní (záznamy SR a jiné typy arytmií) bylo náhodně vybráno a transformováno pouze 20 % záznamů. Po této úpravě datasety jednotlivých běhů obsahovaly přes 40 % pozitivních případů. Důvodem, proč byla alespoň částečně rozšířena i třída negativní, je zohlednění rizika zkreslení reprezentace menšinové třídy. Při rozšíření pouze menšinové třídy se model může přizpůsobovat vzorům, které se v reálném datasetu nevyskytují.

Zmíněné metody řeší problém nevyváženého datasetu pomocí vygenerování nových záznamů. Zcela odlišným přístupem, jak tento problém řešit, je modifikace ztrátové funkce prostřednictvím přiřazení vah, které reprezentují zastoupení tříd v datasetu. Tento přístup byl rovněž implementován a je zahrnutý do porovnání v následující podkapitole.

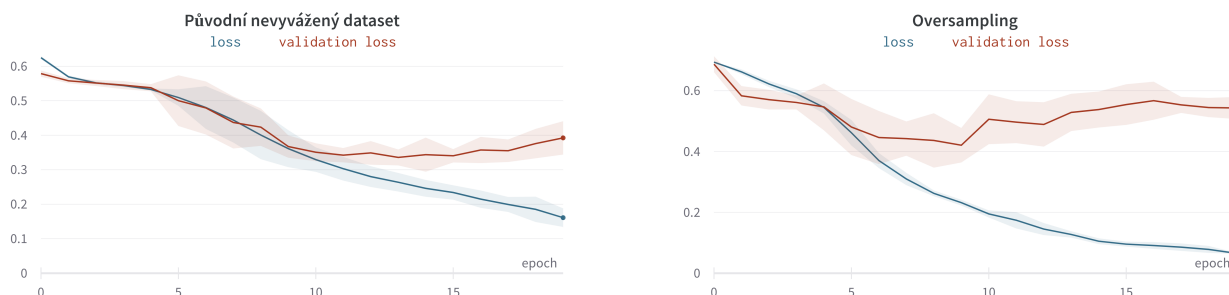
Výsledné porovnání přístupů

Jednotlivé přístupy rozšíření datasetu byly vyhodnoceny na základě tří nezávislých běhů s různou inicializací a různým rozdělením na trénovací a validační podmnožinu (všechna rozdělení obsahovala stejné původní procentuální zastoupení tříd, tedy necelých 30 % případů s FS). Díky tomuto postupu jsou výsledky reprezentativnější. Z každého běhu se zaznamenával model s nejlepším výsledkem (resp. nejnižší validační chybou), jednotlivé běhy byly omezené prostřednictvím techniky EarlyStopping a postupného snižování kroku učení. Rozšíření dat bylo v každé skupině aplikováno pouze na trénovací podmnožinu s náhodným výběrem maximálně tří transformací (mohou se opakovat) s náhodným výběrem patřičných parametrů, jako například škálovacího faktoru. Po ukončení trénování byl na validační podmnožině vyhodnocený graf závislosti F1-skóre na jednotlivých prahových hodnotách (dále jen threshold). Na základě tohoto grafu byl zvolen optimální threshold (takový který maximalizuje F1-skóre).

Každý z běhů byl na závěr vyhodnocen na stejné testovací podmnožině. Tato testovací podmnožina nebyla součástí ani předchozího trénování, ani optimalizace hyperparametrů. Navíc obsahuje stejné procentuální zastoupení třídy FS, jako trénovací a validační podmnožina. Vzhledem k nevyváženosti datasetu byly jako výsledné porovnávací metriky vybrány AUPRC, F1-

skóre, senzitivita a specificita. Finální metriky jednotlivých přístupů odpovídají průměru metrik všech tří nezávislých běhů.

Metriky, optimalizace thresholdu a křivka učení byly zaznamenány prostřednictvím softwarového nástroje wandb.



■ **Obrázek 9.4** Zobrazení křivky učení (trénovací a validáční chyby). Čára reprezentuje průměr všech tří běhů, obarvená oblast zobrazuje hodnotu směrodatné odchylky v daném bodě.

Následující tabulka shrnuje výsledky vybalancování a augmentace dat.

	AUPRC	F1-skóre	Senzitivita	Specificita
Původní dataset	0,78	0,78	89,9 %	83,9 %
Undersampling	0,54	0,61	84,6 %	61,6 %
Oversampling	0,70	0,72	83,1 %	81,2 %
SMOTE	0,75	0,75	88,4 %	81,6 %
Kombinace transformací signálu	0,74	0,77	90,1 %	82,1 %
Přidání vah	0,77	0,76	91,2 %	80,7 %

■ **Tabulka 9.1** Porovnání přístupů vybalancování datasetu.

Z tabulky 9.1 je zřejmé, že vyrovnání datasetu nemělo na úspěšnost modelu zásadní vliv, naopak obzvláště v případě oversamplingu docházelo k výraznému přeučení viz obrázek 9.4. Metoda undersampling měla výrazně negativní vliv na rychlost i průběh učení. Z metod rozšíření datasetu nejlepších výsledků dosahovala metoda SMOTE a augmentace dat prostřednictvím kombinace transformací signálu. I přesto ani jedna z těchto metod nedosahovala tak vysokých výsledků jako klasifikace na původním datasetu. V případě kombinace transformací signálu a metody SMOTE model dokázal rozpoznávat vzorce z dat rychleji, avšak podobně jako v případě metody oversampling docházelo k výraznějšímu přeučení modelu. Přidání vah nemělo z pohledu F1-skóre zásadní vliv na úspěšnost modelu. V případě kombinace transformací signálu a přidání vah lze však pozorovat menší nárůst v senzitivě modelu a pokles ve specificitě modelu.

Jak již bylo zmíněno, žádná z metod vybalancování datasetu není univerzálně optimální. Může se tedy stát, že na konkrétním příkladu žádná z těchto metod ke zlepšení klasifikace nevede. Takový výsledek může být způsoben volbou architektury modelu nebo vlastnostmi datasetu.

Možností jak by se mohlo při optimalizaci modelu postupovat dále, je kombinace více regularizačních technik nebo snížení komplexity modelu. Ani to však nemusí vést k dostatečně přijatelnému výsledku. K další optimalizaci navrženého modelu by bylo třeba získat více skutečných dat. Z tohoto důvodu byl pro další postup zvolen jiný přístup, a to metoda přeneseného učení, která umožňuje efektivně pracovat i s menší velikostí datasetu.

Kapitola 10

Přenesené učení

10.1 Volba předtrénovaného modelu a porovnání datasetů

V této práci byl k přenesenému učení jako výchozí zvolen model z publikace [21], jehož architektura je blíže popsána v podkapitole 3.5.4. Tento model byl vybrán z následujících důvodů:

- Model dosahuje velmi dobrých výsledků v oblasti klasifikace šesti typů arytmií, včetně FS (přes 80 % F1-skóre a specifická přes 99 %).
- Model je natrénovaný na nadstandardně velkém množství dat, navíc použité záznamy EKG jsou 12-svodové.
- Architektura modelu i s patřičnými váhami je plně dostupná.

Datasety ovšem nejsou plně totožné, model je zaměřený na šest různých typů arytmií.

Autoři publikace nezveřejnili trénovací podmnožinu dat, zveřejnili pouze testovací podmnožinu, ze které lze nahlédnout, jak jimi použitý dataset vypadal. Příklad porovnání záznamu z datasetu z publikace oproti záznamu použitým v této práci je uveden v příloze B. Moje práce se specializuje na detekování fibrilace síní, čemuž odpovídá i sestavený dataset, který obsahuje vyšší procentuální zastoupení FS než dataset z publikace (obsahuje pouze necelé 2 % případů FS).

Tabulka 10.1 popisuje úspěšnost modelu z publikace (bez úprav) na mé testovací podmnožině. Klasifikace dosahovala nižších výsledků než v publikaci, což může být právě v důsledku výše zmíněných rozdílů. Následující podkapitoly popisují aplikace různých forem přeneseného učení, jejichž cílem je zvýšit úspěšnost klasifikace na datasetu reálných případů pacientů IKEM.

	F1-skóre	Senzitivita	Specifická
Testovací sada č.1	0,84	86,3 %	92,6 %
Publikovaná testovací sada (vzhledem k FS)	0,87	76,9 %	100 %

■ **Tabulka 10.1** Porovnání úspěšnosti modelu na testovací sadě z publikace oproti vlastním testovacím sadám.

10.2 Vytrénování modelu od základů

Tato část popisuje trénování modelu od základu, tedy bez využití původních optimalizovaných vah. Bylo provedeno několik experimentů, nejlepších výsledků bylo dosaženo s nastavením opti-

malizační metody adam, hodnoty kroku učení 0,001 a batch size 32. Při trénování docházelo k výraznému přeučení, výsledek klasifikace nedosahoval lepších výsledků než základní model z kapitoly 9.1. Od tohoto způsobu tak bylo upuštěno, obrázek 10.1 zobrazuje rozdíl v průběhu učení bez a s využitím původních vah.



■ **Obrázek 10.1** Porovnání průběhu učení v případě při požití předtrénovaných vah oproti náhodné inicializaci.

10.3 Přenesené učení formou extrakce rysů

V případě přeneseného učení formou extrakce rysů se poslední vrstva původního modelu nahradí novou klasifikační vrstvou, původní vrstvy se zmrazí a model se vytrénuje. K optimalizaci hyperparametrů a vyhodnocení modelů byla použita vnořená stratifikovaná křížová validace. Počet skupin vnější křížové validace byl nastaven na 5, počet skupin vnitřní křížové validace na 2. Vzhledem k výpočetní náročnosti a rychlému průběhu učení byl počet epoch nastaven na nižší hodnotu, a to 20 epoch. Hyperparametry (krok učení 0,01, 0,001, 0,0001 a optimalizační metoda adam, SGD, RMSprop) byly optimalizovány pomocí metody GridSearch, pro porovnání při optimalizaci byla zvolena metrika AUPRC. Model s nejvyšší hodnotou AUPRC byl použit k optimalizaci thresholdu na trénovací podmnožině. K porovnání přístupů přeneseného učení formou extrakce rysů byly vybrány metriky F1-skóre, senzitivita a specifita. Tabulka 10.2 porovnává tři různé případy, a to trénování na původním datasetu, na datasetu rozšířeném pomocí metody SMOTE a na datasetu rozšířeném prostřednictvím kombinace transformací obrazu (viz kapitola 9.2). Obě z metod rozšíření datasetu byly aplikovány pouze na trénovací podmnožinu.

Typ datasetu	AUPRC	F1-skóre	Senzitivita	Specifita
Původní dataset	0,90	0,83	83,2 %	92,9 %
Dataset rozšířený pomocí kombinace transformace signálu	0,91	0,83	91,8 %	87,5 %
Dataset rozšířený pomocí metody SMOTE	0,92	0,81	93,3 %	85,2 %

■ **Tabulka 10.2** Porovnání úspěšnosti přeneseného učení formou extrakce rysů.

Z tabulky 10.2 je zřejmé, že přenesené učení formou extrakce rysů z pohledu AUPRC dosahovalo ve všech případech datasetu porovnatelných výsledků. Lze však pozorovat výrazný rozdíl v poměru hodnot senzitivity a specifity. Zatímco u původního datasetu model dosahuje výsledků senzitivity 83,2 %, při rozšíření datasetu (skrze SMOTE nebo kombinaci transformací signálu) se senzitivita zlepšila o minimálně 8 %. Na druhou stranu lze při rozšíření datasetu pozorovat pokles ve specifitě.

10.4 Přenesené učení formou ladění

Další široce používanou metodou využití existujícího modelu je tzv. přenesené učení formou ladění. V tomto případě se poslední klasifikační vrstva předtrénovaného modelu nahradí za vlastní plně propojenou vrstvu a zmrazí se pouze některé vrstvy modelu (případně se přetrénuje celý původní model). V této práci bylo porovnáno několik možných přístupů zahrnujících různý počet přetrénovaných vrstev, nastavení hyperparametrů a rozšíření datasetu. K optimalizaci hyperparametrů a vyhodnocení modelů byla použita vnořená stratifikovaná křížová validace. Mezi otestované hyperparametry byly zahrnuty krok učení 0,01, 0,001, 0,0001 a optimalizační metoda adam, SGD, RMSprop, hodnota dropout 0, 0,25, 0,5, počet plně propojených vrstev 2, 3 a počet jejich výstupů 256, 512, 1024. Počet skupin vnější křížové validace byl nastaven na 5, počet skupin vnitřní křížové validace na 3. Hyperparametry byly optimalizovány pomocí metody Grid-Search, pro porovnání při optimalizaci byla zvolena metrika AUPRC. Model s nejvyšší hodnotou AUPRC byl použit k optimalizaci thresholdu na trénovací podmnožině. K porovnání přístupů přeneseného učení formou ladění byly vybrány metriky F1-skóre, senzitivita a specificita.

Napříč optimalizací hyperparametrů nejvyšších hodnot AUPRC ve většině přístupů z tabulky 10.3 dosahovala konfigurace: 2 plně propojené vrstvy s počtem výstupů 512, bez dropout vrstvy, optimalizační metoda RMSprop a hodnota kroku učení 0,01.

Typ datasetu	Počet zmrazených vrstev	AUPRC	F1-skóre	Senzitivita	Specificita
Původní dataset	49/50	0,93	0,86	86,4 %	94,5 %
Kombinace transformací signálu	49/50	0,93	0,85	90,0 %	91,1 %
SMOTE	49/50	0,92	0,84	90,9 %	89,1 %
Původní dataset	37/50	0,89	0,82	84,7 %	92,2 %

■ **Tabulka 10.3** Porovnání úspěšnosti přeneseného učení formou ladění.

Z tabulky 10.3 je zřejmé, že přenesené učení formou ladění dosahovalo celkově lepších výsledků než přenesené učení formou extrakce rysů. Co se týká porovnání počtu zmrazených vrstev, výrazně horších výsledků dosahoval model s rozmrazením části původního modelu. A to zřejmě v důsledku zásadních rozdílů parametrů v rozmrazených vrstvách oproti zmrazeným vrstvám s původními parametry. Porovnání datasetů z pohledu F1-skóre dosahovalo ve všech případech stejných výsledků, ovšem i zde, podobně jako u extrakce rysů, lze pozorovat rozdíl v poměru hodnot senzitivity a specificity. Zatímco u původního datasetu model dosahuje výsledků senzitivity 86,4 %, při rozšíření datasetu (skrže SMOTE nebo kombinaci transformací signálu) se senzitivita zlepšila o téměř 4 %. Na druhou stranu lze při rozšíření datasetu pozorovat pokles ve specificitě se stejným rozdílem.

Modely z tabulky 10.3, které dosahovaly nejlepších výsledků, byly na závěr přetrénovány s rozmrazením všech vrstev modelu. Jako optimalizační metoda byla zvolena metoda RMSprop, vzhledem ke skutečnosti, že v předchozích experimentech tato optimalizační metoda dosahovala nejlepších výsledků. Hodnota počátečního kroku učení byla nastavena na nižší hodnotu, a to $1E-5$, aby nedocházelo k výraznému přepisování parametrů počátečních vrstev modelu. Výsledky jsou shrnuty v tabulce 10.4. Množina dat byla rozdělena prostřednictvím 3-násobné stratifikované křížové validace na trénovací, testovací a validační podmnožinu. Validační podmnožina byla použita z důvodu monitorování průběhu učení, postupné redukce hodnoty kroku učení a včasného zastavení z důvodu předcházení přeučení. V každém kroku křížové validace byl model vyhodnocen na testovací podmnožině s již optimalizovanou hodnotou threshold. Finální vyhodnocení modelu je průměrem takto získaných výsledků.

Typ datasetu	Počet zmrazených vrstev	AUPRC	F1-skóre	Senzitivita	Specifická
Původní dataset	0/50	0,95	0,89	89,9 %	95,5 %
Kombinace transformací signálu	0/50	0,94	0,87	90,5 %	92,7 %

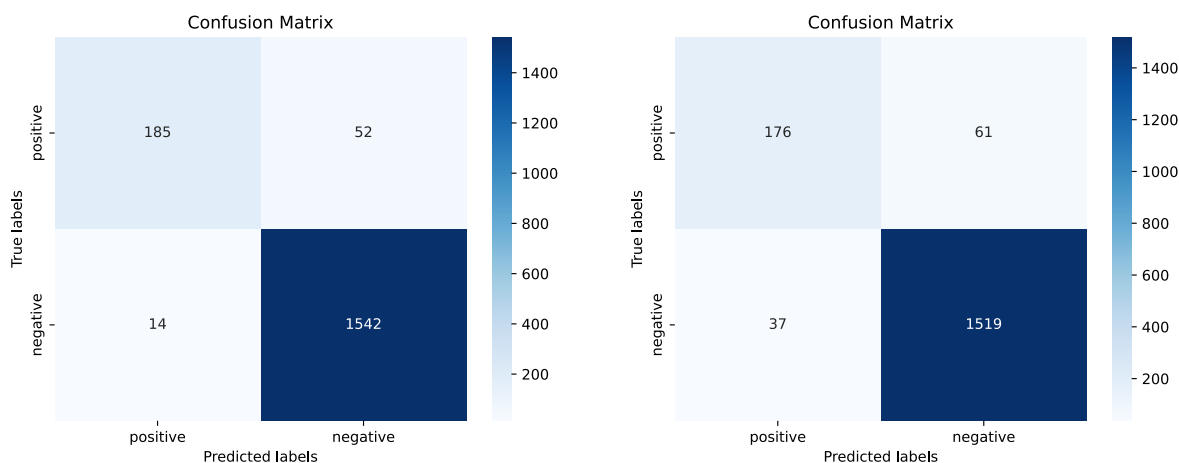
■ **Tabulka 10.4** Porovnání úspěšnosti přeneseného učení formou ladění se sekundárním přetrénování celého modelu.

Z tabulky 10.4 je zřejmé, že sekundární přetrénování celého modelu výrazně zlepšilo výsledek klasifikace fibrilace síní. Ze všech forem přeneseného učení byl tento přístup výpočetně nejvíce náročný, nicméně z porovnání vyhodnocovacích metrik si vedl nejlépe. Vlivem sekundárního přetrénování modelu s původním datasetem došlo ke zlepšení: F1-skóre o 3 %, senzitivity o 3,5 % a specifické o 1 %. U modelu s kombinací transformace signálu došlo ke zlepšení: F1-skóre o 2 %, senzitivity o 0,5 % a specifické o 1,6 %.

Typ datasetu	F1-skóre	Rozdíl v F1-skóre	Senzitivita	Rozdíl v senzitivě	Specificita	Rozdíl ve specificitě
Systém MUSE	0,78	-0,07	74,3 %	-4,7 %	97,6 %	+0,6 %
Model z publikace [21]	0,71	-0,13	62,4 %	-12,1 %	98,0 %	+5,4 %
Vlastní model na původním datasetu	0,85	-0,04	78,0 %	-11,9 %	99,0 %	+3,5 %
Vlastní model na datasetu s kombinací transformací	0,84	-0,03	79,0 %	-11,5 %	98,7 %	+6 %

■ **Tabulka 11.1** Porovnání výsledků klasifikace FS na testovací sadě č.2.

Z tabulky 11.1 je zřejmé, že navržené modely dosahují horší úspěšnosti než v předešlých experimentech na primárním datasetu. Důležité je však poznamenat, že systém MUSE i model z publikace [21] také dosahují menších výsledků na testovacím datasetu č.2 oproti primární testovací sadě. Jaký je důvod tohoto poklesu v úspěšnosti není zřejmé, do budoucna stojí za uvážení konzultovat špatně klasifikované záznamy s lékařem.



■ **Obrázek 11.1** Porovnání matice záměn u finálního modelu (vlevo) a systému MUSE (vpravo) na testovací sadě č.2.

11.3 Srovnání výsledků

Tabulka 11.2 shrnuje dosažené výsledky v porovnání s publikovanými výsledky v oblasti klasifikace fibrilace síní prostřednictvím CNN.

Typ datasetu	F1-skóre	Senzitivita	Specifická	Poznámka
Vlastní model 1	0,78	89,9 %	83,9 %	základní model 16 vrstev
Vlastní model 2	0,89	89,9 %	95,5 %	využití přeneseného učení
Systém MUSE	0,85	79,0 %	97,0 %	používá se v klinické praxi
Model z publikace [21]	0,84	86,3 %	92,6 %	upravená 1D ResNet
Model z publikace [43]	0,78	-	-	jednoduchá architektura, jednosvodové záznamy
Model z publikace [44]	0,99	99,8 %	99,8 %	jednoduchá architektura, jednosvodové záznamy, žádná regularizace

■ **Tabulka 11.2** Porovnání výsledků klasifikace srdečních rytmů, především pak FS.

Základní model s 16 vrstvami nedosahoval tak dobrých výsledků v porovnání s jinými publikacemi. Obzvláště model z publikace [44] dosahuje výrazně vyšších výsledků, i přestože není více komplexní než navržený model. Je však třeba zohlednit skutečnost, že model z publikace [44] se zaměřuje pouze na klasifikaci jednosvodových záznamů, v publikaci navíc autoři nezohledňují žádné regularizační techniky.

Druhý model navržený v této práci dosahuje slibných výsledků. V porovnání s řešením používaném v klinické praxi model dosahuje o téměř 11 % vyšší hodnoty senzitivity. Oproti modelu z publikace [21] (výchozí předtrénovaný model) druhý model dosahuje o 5 % vyšší hodnoty F1-skóre. Výsledky jsou detailněji prodiskutovány v kapitole 12.



Kapitola 12

Diskuze

Model navržený v počáteční fázi práce (CNN složená z 16 vrstev) ani po rozsáhlé optimalizaci hyperparametrů, kdy byly otestovány různé regularizační techniky, jako je dropout, batch normalizace a L2 regularizace, nedosahoval dostatečně optimálního výsledku. Důvodů, které mohly vést k nízké přesnosti modelu, může být více. A to nevyváženost datasetu, přeučení modelu v důsledku nedostatečného množství dat nebo přeučení vlivem velkého počtu parametrů modelu.

Problém nevyváženého datasetu mohl způsobit, že se model příliš zaměřil na třídu s větším zastoupením, což negativně ovlivnilo výsledek klasifikace pro třídu FS. Proto bylo otestováno několik způsobů vybalancování datasetu, mezi něž patří undersampling, oversampling, SMOTE a přidání vah k jednotlivým třídám. Žádný z otestovaných přístupů nicméně nezlepšil celkový výsledek klasifikace. Obzvláště metody undersampling a oversampling výrazně zhoršily úspěšnost klasifikace modelu. Zlepšení oproti původnímu datasetu nastalo pouze v hodnotě senzitivity u metody přidání vah, ale zároveň v tomto případě došlo ke snížení hodnoty specificity. Jak již bylo zmíněno, žádná z metod vybalancování datasetu není univerzálně optimální. Může se tedy stát, že na konkrétním příkladu žádná z těchto metod ke zlepšení klasifikace nevede. Takový výsledek může být způsoben volbou architektury modelu nebo vlastnostmi datasetu.

Při rozšíření datasetu metodou oversampling docházelo k přeučení modelu zřejmě vlivem toho, že se všechny záznamy FS pouze duplikují. Proto byl otestován přístup rozšíření datasetu prostřednictvím transformace signálu, konkrétně přidáním Gaussovského šumu, škálováním a zavedením driftu. Tyto tři techniky byly vybrány s ohledem na skutečnost, že signál vzniklý transformací může nastat i při skutečném měření vlivem různých okolností, jako je pohyb pacienta nebo špatný kontakt elektrod s kůží. Pro větší efekt byly tyto transformace náhodně voleny vždy v kombinaci maximálně tří z nich. Ani v tomto případě však nedošlo k celkovému zlepšení klasifikace, pouze nepatrnému zvýšení hodnoty senzitivity jako v případě přidání vah. Z křivky průběhu učení bylo patrné výrazné přeučení.

Při retrospektivním zhodnocení lze podotknout, že proti přeučení mohlo pomoci ukládání parametrů modelu z epochy s nejmenší validační chybou ještě před momentem, kdy se začne model přeučovat.

Jak již bylo zmíněno v rámci výsledků, nabízí se více možností, jak při optimalizaci modelu dále postupovat. Ani to by však nemuselo vést ke zlepšení výsledku. K další optimalizaci modelu by proto bylo třeba získat více skutečných dat. Z tohoto důvodu byl pro další postup zvolen jiný přístup, a to metoda přeneseného učení, která umožňuje efektivně pracovat i s menší velikostí datasetu.

Pro přenesené učení byla vybrána CNN z publikace [21]. Výhoda jejího využití spočívá v tom, že je natrénovaná na velkém množství 12-svodových záznamů EKG. Díky tomu by model v případě přetrénování na vlastním 12-svodovém datasetu mohl dosahovat lepších výsledků než v případě využití CNN natrénované na jednosvodovém nebo obrázkovém datasetu.

Nejprve byla úspěšnost klasifikace prostřednictvím této sítě bez jakékoliv změny ověřena samostatně na vlastním testovacím datasetu. Model dosahoval dobrých výsledků, avšak existovaly možnosti, jak ho dále vylepšit pomocí přístupů přeneseného učení.

Metoda přeneseného učení formou extrakce rysů výsledek klasifikace nezlepšila, naopak výsledek byl nepatrně horší. To ovšem může být v důsledku toho, že při extrakci rysů byl výsledek vypočítán jako průměr výsledků křížové validace, zatímco model bez úpravy byl otestován na jednorázové testovací sadě. Výsledek původního modelu tak mohl být příliš optimistický. Což se později také potvrdilo při testování na jiné testovací podmnožině. Model upravený formou extrakce rysů dosahoval výrazně vyšší senzitivity v případě natrénování na rozšířeném datasetu než na datasetu původním. Ke zvýšení senzitivity zřejmě došlo v důsledku toho, že model natrénovaný na nevyváženém datasetu větší část vstupních dat predikoval jako rytmy jiné než FS. Vyrovnáním tříd se model naučil lépe detekovat rysy specifické pro FS.

V případě metody přeneseného učení formou ladění bylo dosaženo lepších výsledků v důsledku přidání další plně propojené vrstvy. I v tomto případě nejlepších výsledků dosahoval model natrénovaný na původním datasetu nebo datasetu s kombinací transformací. A i zde byl patrný nárůst senzitivity při využití druhého zmíněného datasetu, rozdíl je však menší než při extrakci rysů. Modely s nejlepším výsledkem byly použity pro další doladění formou druhé fáze přetrénování celého modelu. Čímž se výsledek klasifikace výrazně zlepšil. To naznačuje, že datasety se přece jen lišily a bylo třeba parametry konvolučních vrstev doladit. Díky přetrénování celého modelu došlo ke zlepšení klasifikace o 5 % F1-skóre.

V závěru práce bylo provedeno doplňující testování na nezávislé testovací datové sadě. Model při klasifikaci na druhé testovací sadě nedosahoval tak dobrých výsledků jako v případě primárního datasetu. Důležité je však poznamenat, že tento pokles není specifický pouze pro navržený model, ale lze ho pozorovat i po otestování systému MUSE nebo publikovaného modelu. Proto je pravděpodobně důvodem tohoto poklesu změna vlastností datasetu, nikoliv architektura modelu. Oproti primárnímu datasetu testovací sada č.2 obsahovala výrazně menší procento FS. Jedním z faktorů poklesu výsledků mohla být horší kvalita dat. Ani to však není příliš pravděpodobné vzhledem ke skutečnosti, že záznamy EKG byly získány a zpracovány stejným způsobem jako v případě primárního datasetu. Navíc testovací množina byla dostatečně početná na to, aby se předešlo náhodné agregaci nekvalitních záznamů. Pokles by také neměl být způsoben argumentem optimistického výsledku na náhodně vybrané původní testovací sadě, protože navržené modely byly otestovány pomocí křížové validace a výsledek by tak měl být reprezentativní.

Navržený model byl mimo jiné porovnán se systémem MUSE, kdy model dosahuje minimálně o 10 % vyšší hodnoty senzitivity. Navržený model sice dosahuje nižší hodnoty specifcity (o 1,5 %), ale z pohledu F1-skóre je výsledek modelu lepší o 4 %. V případě včasného záchytu srdečních abnormalit může být právě vyšší senzitivita modelu klíčová.

Navržený model v porovnání s výsledky jiných publikací dosahuje dobrých výsledků. Do budoucna by bylo zajímavé model dále optimalizovat, k dalším experimentům je však třeba rozšířit dataset. Nabízí se několik možností, a to například rozšíření datasetu prostřednictvím GAN nebo využití neanotovaných záznamů z IKEM. Dále by bylo zajímavé upravit architekturu přidáním LSTM sítě. V praxi by pak stálo za zvážení klasifikaci rozšířit i na další typy arytmií nebo pracovat s neuronovou sítí na jiném typu úlohy, jako je dlouhodobá předpověď vývoje zdravotního stavu na základě aktuálně naměřeného záznamu EKG.



Kapitola 13

Závěr

Tato práce se zabývá problematikou klasifikace fibrilace síní ze snímků EKG prostřednictvím konvolučních neuronových sítí. Cílem práce bylo navrhnout a implementovat vlastní řešení, které pracuje s datasetem poskytnutým Institutem klinické a experimentální medicíny (IKEM).

V teoretické části práce (kapitoly 1–5) jsou uvedeny koncepty z oblasti neuronových sítí, arytmiologie, přeneseného učení a augmentace dat. Kapitola 6 se věnuje dosavadnímu výzkumu v oblasti klasifikace srdečních arytmií prostřednictvím metod strojového učení.

V praktické části práce byla implementována konvoluční neuronová síť s 16 vrstvami. I přes optimalizaci hyperparametrů a implementaci různých regularizačních technik u modelu docházelo k přeučení, a to zřejmě v důsledku menší velikosti datasetu.

Z toho důvodu bylo otestováno několik různých přístupů rozšíření datasetu včetně technik SMOTE a augmentace dat formou náhodné kombinace transformací signálu, mezi něž patří škálování, přidání Gaussovského šumu a přidání driftu. Žádná z otestovaných metod nicméně nevedla ke zlepšení výsledku klasifikace, naopak přeučení modelu trénovaném na rozšířeném datasetu bylo výraznější. Nejlepších výsledků dosahoval model natrénovaný na původním datasetu, a to 89 % senzitivity, 83 % specificity a 0,78 F1-skóre. Takový výsledek není příliš optimální, k další optimalizaci modelu by však bylo třeba získat více dat. Z toho důvodu bylo od této architektury upuštěno a v další části práce s přihlédnutím na velikost datasetu byla navržena architektura na základě přeneseného učení s využitím již předtrénovaného modelu.

Nejlepších výsledků, a to 90 % senzitivity, 94 % specificity a 0,89 F1-skóre dosahoval model přetrénovaný formou ladění ve dvou fázích, kdy se v první fázi trénují pouze nahrazené plně propojené vrstvy modelu a ve druhé fázi se přetrénuje celý model s nižší hodnotou kroku učení. V případě modelu založeného na konceptu přeneseného učení lze pozorovat výrazné zvýšení senzitivity při trénování na rozšířeném datasetu skrze augmentaci dat formou náhodné kombinace transformací signálu nebo metodou SMOTE.

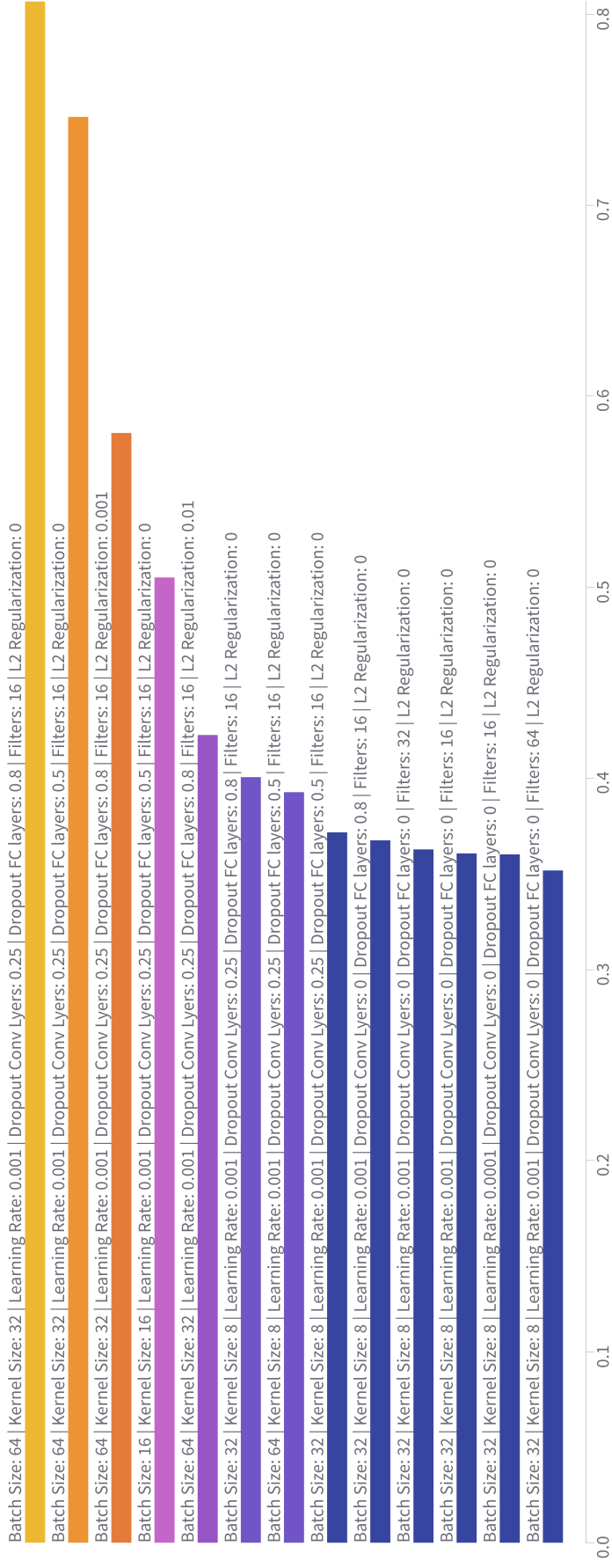
Tento model dosahuje porovnatelných výsledků s jinými publikacemi a má vyšší schopnost přesně detekovat pozitivní případy výskytu fibrilace síní v porovnání s řešením používaným v klinické praxi (systém MUSE). Model dosahuje až o 10 % vyšší hodnoty senzitivity než systém MUSE.

Možným rozšířením práce by bylo navrhnout model pro klasifikaci více typů arytmií nebo upravit architekturu modelu přidáním LSTM sítě. K dalším experimentům je však třeba dataset rozšířit. Nabízí se několik možností, a to například rozšíření datasetu prostřednictvím GAN, využití neanotovaných záznamů nebo sjednocení datasetu s veřejně dostupnými daty.

..... Příloha A

Optimalizace hyperparametrů

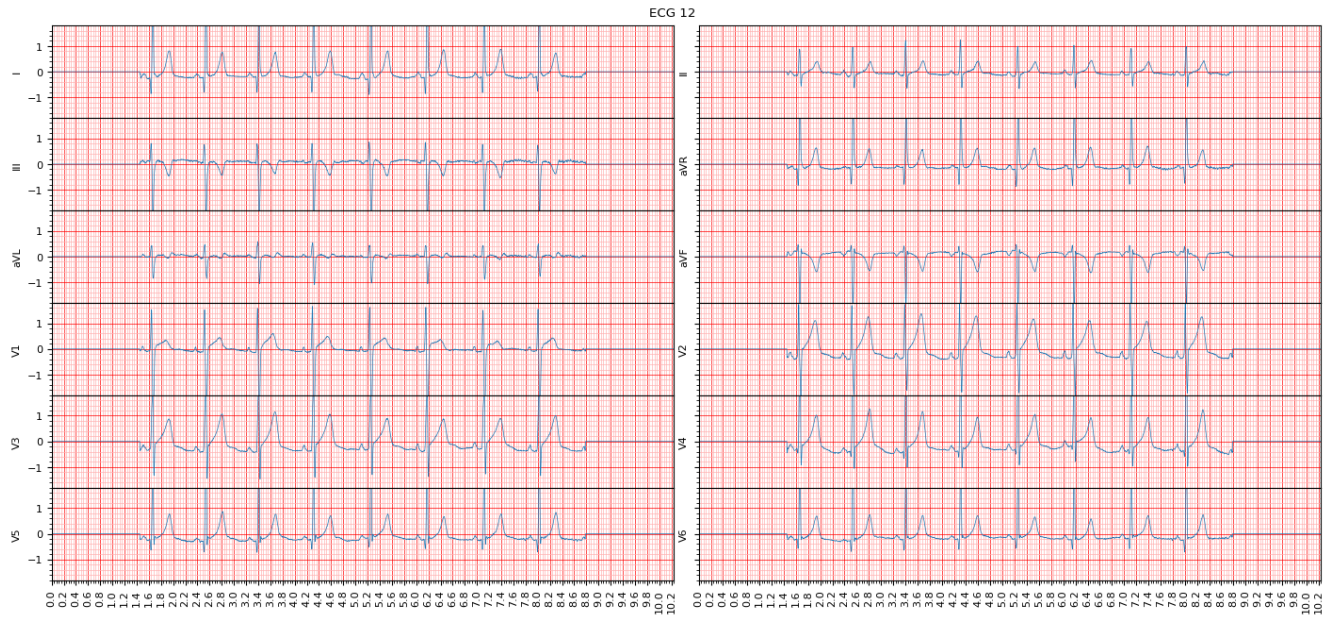
AUPRC



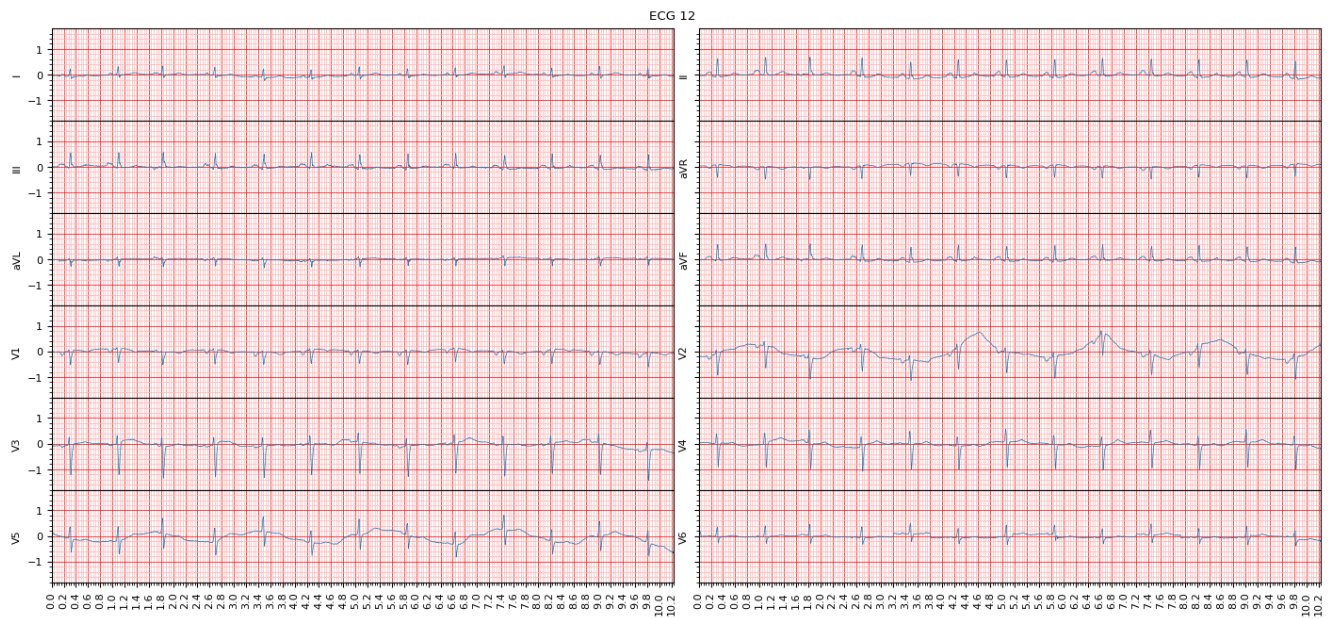
■ **Obrázek A.1** Hodnoty AUPRC v závislosti na nastavení hyperparametrů.

..... Příloha B

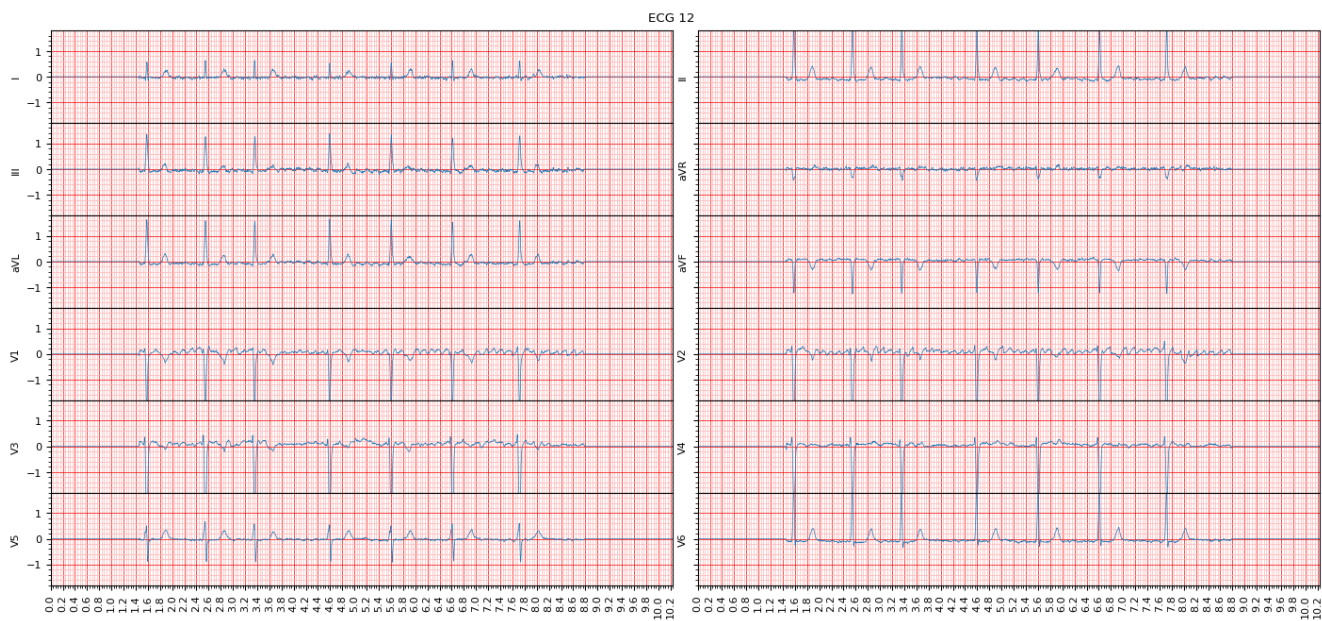
Porovnání EKG záznamů



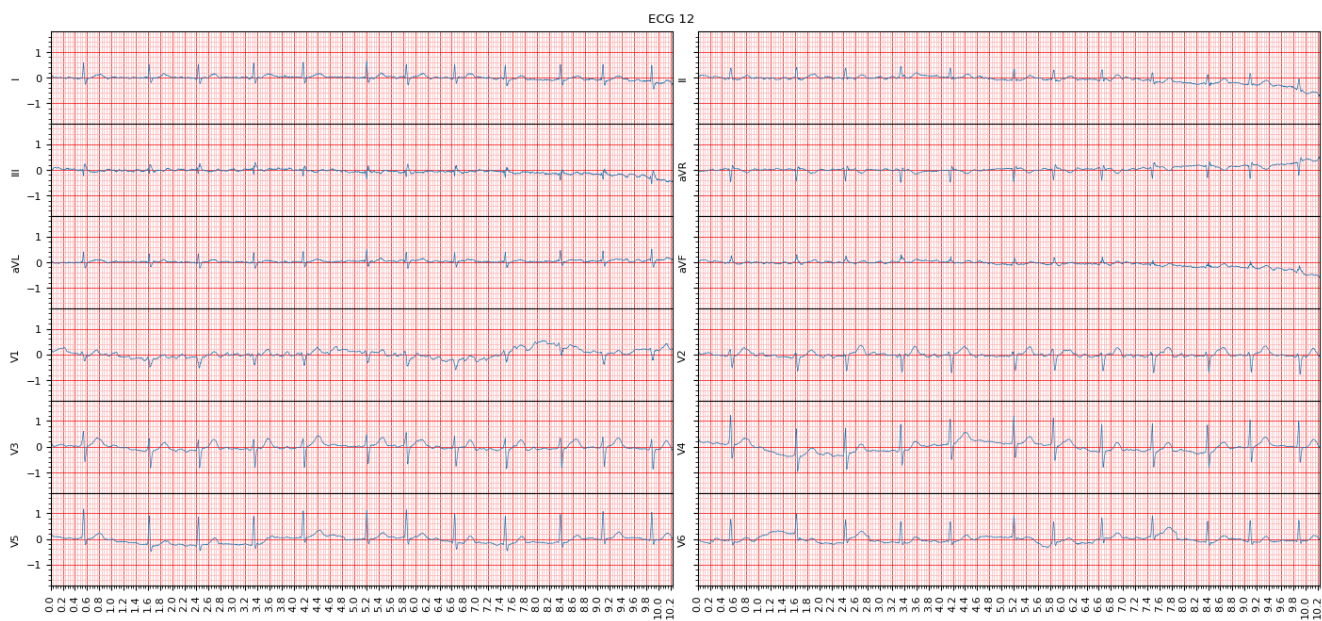
■ **Obrázek B.1** Příklad sinusového rytmu z testovacího datasetu publikace [21].



■ **Obrázek B.2** Příklad sinusového rytmu z vlastního datasetu.



■ Obrázek B.3 Příklad fibrilace síní z testovacího datasetu publikace [21].



■ Obrázek B.4 Příklad fibrilace síní z vlastního datasetu.

Bibliografie

- [1] TROJAN, Stanislav. *Lékařská fyziologie*. Vyd. 4., přeprac. a dopl. Praha: Grada, 2003. ISBN 80-247-0512-5.
- [2] HABERL, Ralph. *EKG do kapsy*. Praha: Grada, 2012. ISBN 978-80-247-4192-5. Překlad 4. vydání.
- [3] HAMPTON, John R. *EKG stručně, jasně, přehledně*. Vyd. 2., rozš. Praha: Grada, 2005, c1996. ISBN 80-247-0960-0. Překlad 6. vydání.
- [4] BENNETT, David H. *Srdeční arytmie: praktické poznámky k interpretaci a léčbě*. 8. Praha: Grada, 2014. ISBN 978-80-247-5134-4.
- [5] SU, Li, Stefan BOROVIČ a Bernhard ZRENNER. 12-lead Holter electrocardiography. *Herzschrittmachertherapie Elektrophysiologie*. 2013, 24(2), 92-96. ISSN 0938-7412. Dostupné z: doi:10.1007/s00399-013-0268-4.
- [6] XIAOLIN, Li, Barry CARDIFF a Deepu JOHN. A 1D Convolutional Neural Network for Heartbeat Classification from Single Lead ECG. 2020 *27th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. IEEE, 2020, 2020-11-23, 27, 1-2. ISBN 978-1-7281-6044-3. Dostupné z: doi:10.1109/ICECS49266.2020.9294838.
- [7] HE, Runnan, Kuanquan WANG, Qince LI, Yongfeng YUAN, Na ZHAO, Yang LIU a Henggui ZHANG. A novel method for the detection of R-peaks in ECG based on K-Nearest Neighbors and Particle Swarm Optimization. *EURASIP Journal on Advances in Signal Processing*. 2017, 2017(1). ISSN 1687-6180. Dostupné z: doi:10.1186/s13634-017-0519-3.
- [8] LAZZERINI, Pietro Enea, Franco LAGHI-PASINI, Mohamed BOUTJDIR a Pier Leopoldo CAPECCHI. Inflammatory cytokines and cardiac arrhythmias: the lesson from COVID-19. *Nature Reviews Immunology*. 2022, 22(5), 270-272. ISSN 1474-1733. Dostupné z: doi:10.1038/s41577-022-00714-3.
- [9] Fibrilace síní. *IKEM* [online]. Praha: Ikem, c2015–2022 [cit. 2022-11-15]. Dostupné z: <https://www.ikem.cz/cs/fibrilace-sini/a-436/>.
- [10] SIDEY-GIBBONS, Jenni A. M. a Chris J. SIDEY-GIBBONS. Machine learning in medicine: a practical introduction. *BMC Medical Research Methodology*. 2019, 19(1). ISSN 1471-2288. Dostupné z: doi:10.1186/s12874-019-0681-4.
- [11] PATTERSON, Josh a Adam GIBSON. *Deep learning: a practitioner's approach*. Sebastopol, CA: O'Reilly, 2017. ISBN 978-1-491-91425-0.
- [12] CHOLLET, François. *Deep learning with Python*. Second edition. Shelter Island: Manning, 2021. ISBN 978-161-7296-864.

- [13] MAAS, Andrew, Awni HANNUN a Andrew NG. Rectifier Nonlinearities Improve Neural Network Acoustic Models. *JMLR: WCP: Proceedings of the 30 th International Conference on Machine Learning*. Atlanta, Georgia, USA, 2013, 30(28).
- [14] LU, Lu. Dying ReLU and Initialization: Theory and Numerical Examples. ArXiv. Dostupné z: arxiv.org/abs/1903.06733.
- [15] SOYDANER, Derya. A Comparison of Optimization Algorithms for Deep Learning. *International Journal of Pattern Recognition and Artificial Intelligence*. 2020, 34(13). ISSN 0218-0014. Dostupné z: [doi:10.1142/S0218001420520138](https://doi.org/10.1142/S0218001420520138).
- [16] RUDER, Sebastian. An overview of gradient descent optimization algorithms. ArXiv. Dostupné z: arxiv.org/pdf/1609.04747.
- [17] GOODFELLOW, Ian, Yoshua BENGIO a Aaron COURVILLE. *Deep learning*. Cambridge, MA: MIT press, 2016. Adaptive computation and machine learning series. ISBN 978-026-2035-613.
- [18] HE, Kaiming, Xiangyu ZHANG, Shaoqing REN a Jian SUN. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2016, 2016, 29, 770-778. ISBN 978-1-4673-8851-1. Dostupné z: [doi:10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- [19] KRIZHEVSKY, Alex, Ilya SUTSKEVER a Geoffrey E. HINTON. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*. 2017, 60(6), 84-90. ISSN 0001-0782. Dostupné z: [doi:10.1145/3065386](https://doi.org/10.1145/3065386).
- [20] SIMONYAN, Karen. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2014.
- [21] RIBEIRO, Antônio H., Manoel Horta RIBEIRO, Gabriela M. M. PAIXÃO, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nature Communications*. 2020, 11(1), 1-9. ISSN 2041-1723. Dostupné z: [doi:10.1038/s41467-020-15432-4](https://doi.org/10.1038/s41467-020-15432-4).
- [22] XIONG, Yingnan, Lin WANG, Qingnan WANG, Shan LIU a Bo KOU. Improved convolutional neural network with feature selection for imbalanced ECG Multi-Factor classification. *Measurement*. 2022, 189. ISSN 02632241. Dostupné z: [doi:10.1016/j.measurement.2021.110471](https://doi.org/10.1016/j.measurement.2021.110471).
- [23] CHENGYU, Liu a Li JIANQING. *Feature Engineering and Computational Intelligence in ECG Monitoring*. Singapore: Springer, 2020. ISBN 978-981-15-3824-7.
- [24] CHAWLA, N. V., K. W. BOWYER, L. O. HALL a W. P. KEGELMEYER. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*. 2002, 16, 321-357. ISSN 1076-9757. Dostupné z: [doi:10.1613/jair.953](https://doi.org/10.1613/jair.953).
- [25] PIACENTINO, Esteban, Alvaro GUARNER a Cecilio ANGULO. Generating Synthetic ECGs Using GANs for Anonymizing Healthcare Data. *Electronics*. 2021, 10(4). ISSN 2079-9292. Dostupné z: [doi:10.3390/electronics10040389](https://doi.org/10.3390/electronics10040389).
- [26] AN, Junmo, Richard E. GREGG a Soheil BORHANI. Effective Data Augmentation, Filters, and Automation Techniques for Automatic 12-Lead ECG Classification Using Deep Residual Neural Networks. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2022, 44, 1283-1287. ISBN 978-1-7281-2782-8. Dostupné z: [doi:10.1109/EMBC48229.2022.9871654](https://doi.org/10.1109/EMBC48229.2022.9871654).
- [27] SOLTANIEH, Sahar, Ali ETEMAD a Javad HASHEMI. Analysis of Augmentations for Contrastive ECG Representation Learning. *International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2022, 1-10. ISBN 978-1-7281-8671-9. Dostupné z: [doi:10.1109/IJCNN55064.2022.9892600](https://doi.org/10.1109/IJCNN55064.2022.9892600).
- [28] ASL, Babak Mohammadzadeh, Seyed Kamaledin SETAREHDAN a Maryam MOHEBBI. Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. *Artificial Intelligence in Medicine*. 2008, 44(1), 51-64. ISSN 09333657. Dostupné z: [doi:10.1016/j.artmed.2008.04.007](https://doi.org/10.1016/j.artmed.2008.04.007).

- [29] ÇINAR, Ahmet a Seda Arslan TUNCER. Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks. *Computer Methods in Biomechanics and Biomedical Engineering*. 2021, 24(2), 203-214. ISSN 1025-5842. Dostupné z: doi:10.1080/10255842.2020.1821192.
- [30] U, Yunan. *A Comparison of 1-D and 2-D Deep Convolutional Neural Networks in ECG Classification*. 2018.
- [31] CAO, Ping, Xinyi LI, Kedong MAO, Fei LU, Gangmin NING, Luping FANG a Qing PAN. A novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation. *Biomedical Signal Processing and Control*. 2020, 56. ISSN 17468094. Dostupné z: doi:10.1016/j.bspc.2019.101675.
- [32] PANDEY, Saroj Kumar a Rekh Ram JANGHEL. Automatic detection of arrhythmia from imbalanced ECG database using CNN model with SMOTE. *Australas Phys Eng Sci Med*. 2019, 42(4), 1129-1139. ISSN 0158-9938. Dostupné z: doi:10.1007/s13246-019-00815-9.
- [33] SALEM, Milad, Shayan TAHERI a Jiann-Shiun YUAN. ECG Arrhythmia Classification Using Transfer Learning from 2- Dimensional Deep CNN Features. *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*. IEEE, 2018, 2018, 1-4. ISBN 978-1-5386-3603-9. Dostupné z: doi:10.1109/BIOCAS.2018.8584808.
- [34] WEIMANN, Kuba a Tim O. F. CONRAD. Transfer learning for ECG classification. *Scientific Reports*. 2021, 11(1), 1-12. ISSN 2045-2322. Dostupné z: doi:10.1038/s41598-021-84374-8.
- [35] XIA, Yufa, Huailing ZHANG, Lin XU, Zhifan GAO, Heye ZHANG, Huafeng LIU a Shuo LI. An Automatic Cardiac Arrhythmia Classification System With Wearable Electrocardiogram. *IEEE Access*. 2018, 6, 16529-16538. ISSN 2169-3536. Dostupné z: doi:10.1109/ACCESS.2018.2807700.
- [36] ZHAI, Xiaolong, Zhanhong ZHOU a Chung TIN. Semi-supervised learning for ECG classification without patient-specific labeled data. *Expert Systems with Applications*. 2020, 158. ISSN 09574174. Dostupné z: doi:10.1016/j.eswa.2020.113411.
- [37] SINGH, Shraddha, Saroj Kumar PANDEY, Urja PAWAR a Rekh Ram JANGHEL. Classification of ECG Arrhythmia using Recurrent Neural Networks. *Procedia Computer Science*. 2018, 132, 1290-1297. ISSN 18770509. Dostupné z: doi:10.1016/j.procs.2018.05.045.
- [38] PEREIRA, Rafael a Rodrigo Varejão ANDREÃO. Inter-patient detection of atrial fibrillation in short ECG segments based on LSTM network with multiple input layers. *Research on Biomedical Engineering*. 2022, 38(2), 465-476. ISSN 2446-4740. Dostupné z: doi:10.1007/s42600-022-00198-z.
- [39] SCHWAB, Patrick, Gaetano Claudio SCEBBA, Jia ZHANG, Marco DELAI a Walter KARLEN. Beat by Beat: Classifying Cardiac Arrhythmias with Recurrent Neural Networks. *Computing in Cardiology (CinC)*. IEEE, 2017, 2017, 1-4. Dostupné z: doi:10.22489/CinC.2017.363-223.
- [40] PETMEZAS, Georgios, Kostas HARIS, Leandros STEFANOPOULOS, Vassilis KILINTZIS, Andreas TZAVELIS, John A ROGERS, Aggelos K KATSAGGELOS a Nicos MAGLAVERAS. Automated Atrial Fibrillation Detection using a Hybrid CNN-LSTM Network on Imbalanced ECG Datasets. *Biomedical Signal Processing and Control*. 2021, 63. ISSN 17468094. Dostupné z: doi:10.1016/j.bspc.2020.102194.
- [41] CHEN, Chen, Zhengchun HUA, Ruiqi ZHANG, Guangyuan LIU a Wanhui WEN. Automated arrhythmia classification based on a combination network of CNN and LSTM. *Biomedical Signal Processing and Control*. 2020, 57. ISSN 17468094. Dostupné z: doi:10.1016/j.bspc.2019.101819.
- [42] LIU, Hongwei, Reid COLLINS, Robert J. H. MILLER, et al. Use of a Clinical Electrocardiographic Database to Enhance Atrial Fibrillation/Atrial Flutter Identification Algorithms Based on Administrative Data. *Journal of the American Heart Association*. 2021, 10(7). ISSN 2047-9980. Dostupné z: doi:10.1161/JAHA.120.018511.

- [43] HSIEH, Chaur-Heh, Yan-Shuo LI, Bor-Jiunn HWANG a Ching-Hua HSIAO. Detection of Atrial Fibrillation Using 1D Convolutional Neural Network. *Sensors*. 2020, 20(7). ISSN 1424-8220. Dostupné z: doi:10.3390/s20072136.
- [44] TUTUKO, Bambang, Siti NURMAINI, Alexander Edo TONDAS, Muhammad Naufal RA-CHMATULLAH, Annisa DARMAWAHYUNI, Ria ESAFRI, Firdaus FIRDAUS a Ade Iri-ani SAPITRI. AFibNet: an implementation of atrial fibrillation detection with convolutive neural network. *BMC Medical Informatics and Decision Making*. 2021, 21(1). ISSN 1472-6947. Dostupné z: doi:10.1186/s12911-021-01571-1.

Obsah přiloženého média

	README.txt	stručný popis obsahu média
	src	zdrojové kódy implementace
	thesis		
		tex zdrojová forma práce ve formátu L ^A T _E X
		thesis.pdf text práce ve formátu PDF