

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2020

Bc. Markéta Smrčková



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

BIOMETRIE S VYUŽITÍM SNÍMKŮ SÍTNICE S NÍZKÝM ROZLIŠENÍM

RETINAL BIOMETRY WITH LOW RESOLUTION IMAGES

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Markéta Smrčková

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. Radim Kolář, Ph.D.

BRNO 2020

Diplomová práce

magisterský navazující studijní obor **Biomedicínské a ekologické inženýrství**

Ústav biomedicínského inženýrství

Studentka: Bc. Markéta Smrčková

ID: 186689

Ročník: 2

Akademický rok: 2019/20

NÁZEV TÉMATU:

Biometrie s využitím snímků sítnice s nízkým rozlišením

POKYNY PRO VYPRACOVÁNÍ:

1) Seznamte se s principem snímání sítnice pomocí experimentálního video-ofthalmoskopu. 2) Provedte rešerši aktuálních přístupů z oblasti biometrie, především pak biometrie s využitím sítnice. 3) Navrhněte, popište a proveďte implementaci pokročilé metody pro biometrické rozpoznávání osob na základě snímků sítnice. 4) Z dodaných a veřejně dostupných snímků sítnice si připravte sadu dat pro testování navržených postupů a proveďte testování. 5) Podle výsledků se pokuste metodu upravit za účelem co největší přesnosti metody. 6) Provedte komplexní testy, výsledky dostatečně popište a diskutujte.

DOPORUČENÁ LITERATURA:

[1] Dražanský M. et al., „Biometrie“, Brno Computer Press, 2011

[2] Nixon M., a Aguado A., „Feature Extraction and Image Processing“, Elsevier, 2005

Termín zadání: 3.2.2020

Termín odevzdání: 29.5.2020

Vedoucí práce: doc. Ing. Radim Kolář, Ph.D.

prof. Ing. Ivo Provazník, Ph.D.
předseda oborové rady

UPOZORNĚNÍ:

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

ABSTRAKT

Tato práce se zabývá návrhem metody pro biometrickou identifikaci ze snímků oční sítnice. První část práce je zaměřena především na principy a pojmy, využívané v biometrii, anatomii oka a metody pro biometrii sítnice. Je zde popsána podstata neuronových sítí a hlubokého učení, které budou v praktické části využity. V poslední části se práce věnuje popisu zvoleného identifikačního algoritmu, jeho implementaci a zhodnocení výsledků biometrického systému.

KLÍČOVÁ SLOVA

Biometrie, identifikace sítnice, neuronové sítě, konvoluční sítě, hluboké učení, klasifikace obrazů, strojové učení, Matlab.

ABSTRACT

This thesis attempts to find an alternative method for biometric identification using retinal images. First part is focused on the introduction to biometrics, human eye anatomy and methods used for retinal biometry. The essence of neural networks and deep learning methods is described as it will be used practically. In the last part of the thesis a chosen identification algorithm and its implementation is described and the results are presented.

KEYWORDS

Biometry, retina identification, neural networks, convolutional networks, deep learning, image classification, machine learning, Matlab.

SMRČKOVÁ, Markéta. *Biometrie s využitím snímků sítnice s nízkým rozlišením*. Brno, 2020, 70 s. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce: doc. Ing. Radim Kolář, Ph.D.

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Biometrie s využitím snímků sítnice s nízkým rozlišením“ jsem vypracovala samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

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

Brno

.....

podpis autorky

PODĚKOVÁNÍ

Ráda bych poděkovala vedoucímu diplomové práce panu doc. Ing. Radimu Kolářovi, Ph.D. za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Brno

.....

podpis autorky

Obsah

Úvod	10
1 Biometrie a základní pojmy	11
1.1 Biometrické znaky	11
1.2 Biometrické systémy	12
1.3 Hodnocení kvality systému	14
2 Anatomie oka a sítnice	16
2.1 Sítnice oka	16
2.2 Patologie sítnice	17
2.3 Měřicí přístroje	19
3 Biometrie sítnice	21
3.1 Akvizice snímku	21
3.2 Segmentace cévního řečiště	21
3.3 Extrakce příznaků	22
3.4 Vyhodnocení podobnosti	23
3.5 Support Vector Machine (SVM)	24
4 Umělé neuronové sítě (ANN, NN)	27
4.1 Hluboké učení (deep learning)	28
4.2 Konvoluční neuronové sítě (CNN)	29
4.3 Transfer Learning	31
5 Praktická část	36
5.1 Příprava datasetu	36
5.2 Výběr architektury CNN	39
5.3 Transfer learning na zvolených CNN	40
5.4 Modifikace sítě ResNet50	42
6 Výsledky diplomové práce	44
6.1 Výsledky Transfer learning na zvolených CNN	44
6.2 Možná rozšíření práce	48
Závěr	49
Literatura	50
Seznam symbolů, veličin a zkratk	55

Seznam příloh	56
A Přílohy	57
A.1 Příložená data a kódy	57
A.2 Obrazové přílohy - přehled trénovacích procesů jednotlivých sítí . . .	58

Seznam obrázků

1.1	Schéma verifikačního a identifikačního systému [3]	13
1.2	Pravděpodobnostní rozložení chybových funkcí FMR a FNMR [4]. . .	15
2.1	Anatomie lidského oka [7]	16
2.2	Příklad snímku zdravé sítnice s vyznačenými anatomickými body. [8]	17
3.1	Lineární SVM klasifikace [28].	25
3.2	Nelineární SVM klasifikace. Přepřacováno z [28].	26
4.1	Schéma perceptronu neuronové sítě. Přepřacováno z [30].	27
4.2	Schéma propojení vrstev klasické neuronové sítě. Přepřacováno z [31].	28
4.3	Schéma propojení vrstev konvoluční neuronové sítě. Přepřacováno z [31].	29
4.4	Příklad vzorů uložených v kernelech CNN [31].	30
4.5	Architektura konvoluční neuronové sítě AlexNet [36].	32
4.6	Architektura konvoluční neuronové sítě GoogLeNet [38].	33
4.7	Architektura konvoluční neuronové sítě VGG-19 [40].	34
4.8	Architektura konvoluční neuronové sítě typu ResNet [41].	35
5.1	Příklad dat, zleva: šedotónový obraz, RGB obraz a binární obraz segmentovaného cévního stromu.	36
5.2	Příklad snímků, které byly z databáze odstraněny.	37
5.3	RGB snímek, rozložený šedotónové reprezentace barevných kanálů. Zleva: červený, zelený a modrý kanál.	38
5.4	Zleva: původní snímek, snímek po úpravě oříznutím a změnou velikosti.	39
5.5	Průběh trénování sítě ResNet18 v režimu verifikace.	40
5.6	Průběh trénování sítě GoogLeNet v režimu identifikace.	41
5.7	Přesnost SVM klasifikace z příznaků v dané vrstvě ResNet50.	43
6.1	Výsledky testovaných sítí v režimu verifikace.	45
6.2	Výsledky testovaných sítí v režimu identifikace.	46
6.3	Příklad špatně klasifikovaného snímku sítě GoogLeNet, ResNet18 a ResNet37 v režimu verifikace.	46
6.4	Příklad špatně klasifikovaného snímku sítí ResNet18 v režimu identi- fikace.	47
6.5	Příklad snímků sítnice z databáze DRIVE [45].	48
A.1	Trénovací proces sítě AlexNet v režimu verifikace.	59
A.2	Trénovací proces sítě GoogLeNet v režimu verifikace.	60
A.3	Trénovací proces sítě VGG-19 v režimu verifikace.	61
A.4	Trénovací proces sítě ResNet18 v režimu verifikace.	62
A.5	Trénovací proces sítě ResNet37 v režimu verifikace.	63
A.6	Trénovací proces sítě ResNet50 v režimu verifikace.	64

A.7	Trénovací proces sítě AlexNet v režimu identifikace.	65
A.8	Trénovací proces sítě GoogLeNet v režimu identifikace.	66
A.9	Trénovací proces sítě VGG-19 v režimu identifikace.	67
A.10	Trénovací proces sítě ResNet18 v režimu identifikace.	68
A.11	Trénovací proces sítě ResNet37 v režimu identifikace.	69
A.12	Trénovací proces sítě ResNet50 v režimu identifikace.	70

Úvod

Biometrie je věda, věnující se snímání a vyhodnocování určitých rysů jedince, za účelem jeho následné identifikace. Snahou je především nahradit dosud využívaný koncept identifikačních čísel, průkazů, karet a přístupových údajů jako jsou hesla.

Problémem těchto forem identifikace může být jejich poměrně snadné zneužití, odcizení či ztráta karet, zapomenutí hesla apod. Jako protipól výhod jednoznačné identifikace se naopak nabízí otázka, zda by se měla takto citlivá data shromažďovat a zda se nejedná o narušení práva jedince na ochranu osobních údajů.

Nejznámější a dosud nejrozšířenější biometrická metoda je metoda snímání otisků prstů, která již dlouho nachází své využití především v kriminalistice, novodobě také v zabezpečení mobilních telefonů či střežených objektů. S rozvojem výpočetní techniky se v posledních letech rozvíjí další, velmi spolehlivé metody. Tato práce se konkrétně zaměřuje na identifikaci osob pomocí snímků sítnice.

V první části budou představeny základní pojmy z oboru biometrie, přehled biometrických metod pro identifikaci a také statistické veličiny pro hodnocení identifikačních systémů. Bude zde věnována pozornost především na teoretický podklad pro biometrii sítnice.

Další část práce se již konkrétněji zaměřuje na biometrii sítnice, objasní se zde jednotlivé kroky a metody, vedoucí k identifikaci osob. Dále se práce věnuje především neuronovým sítím, které mohou být pro biometrii v této práci využity.

Poslední část obsahuje popis postupu při implementaci zvoleného algoritmu na reálná data, prezentaci a zhodnocení výsledků klasifikace a představení případných možností rozšíření tohoto algoritmu.

1 Biometrie a základní pojmy

Biometrie je věda, využívající unikátních biologických znaků pro jednoznačnou identifikaci osob. S rozvojem nových technologií je mnohem snazší tyto znaky zaznamenávat, uchovávat, a především vyhodnocovat. Zabezpečení biometrického typu se již běžně používá ve vojenství, bankovníctví, ochraně střežených objektů a dalších aplikacích. Mnohem častěji se s biometrií však začínáme setkávat i v běžném životě, například u chytrých telefonů, které již využívají jak otisky prstů, tak rozpoznávání obličeje.

Rozvoj biometrických metod je umožněn především snižující se cenou potřebných technologií jako jsou senzory, kamery a také vyhodnocovací softwary. V dnešní době již navíc nemusí platit závislost ceny komerčně dostupných snímačů na kvalitě a přesnosti výsledků. Závislost ceny technického vybavení na spolehlivosti identifikace je důležitá především u biometrie oka, ať už se jedná o biometrii sítnice či oční duhovky, nebo otisků prstů. Kde však cena identifikačního systému příliš nerozhoduje jsou metody, které nezaručují jednoznačnost vstupních dat. Například hlasové rozpoznávání může být i přes kvalitní senzor ovlivněno šumem pozadí apod., u rozpoznávání obličeje může mít na výsledek vliv rozdílná akvizice porovnávaných snímků obličeje, odlesky a jiné obrazové artefakty [1].

1.1 Biometrické znaky

Identifikace osob je umožněna jedinečnými znaky člověka (tzv. markanty). Tyto znaky se dělí do dvou základních skupin:

Anatomické znaky

Anatomické znaky jsou konkrétní fyzické rysy lidského těla, které se v průběhu života víceméně nemění. Jejich nezpochybnitelnou výhodou je, že prakticky nejsou ovlivnitelné jak momentálním stavem člověka, tak ani podmínkami prostředí. Výjimku mohou tvořit případy některých onemocnění či závažných poranění (např. otisk prstu při popálení rukou, oční zákal, změna cév v sítnici jako následek diabetu). Metody založené na anatomických rysech jsou nejrozšířenější a zároveň nejspolehlivější. Mezi snímané znaky patří například:

- Sítnice oka
- Duhovka oka
- Otisk prstu
- Geometrie obličeje
- Geometrie ruky

- Geometrie ušních lalůček
- Dentální obraz
- DNA
- a další

Behaviorální znaky

Jak již plyne z názvu, tyto znaky jsou založeny na určitém typu chování člověka. Do jisté míry mohou i tyto znaky být podmíněné anatomicky, neboť například styl chůze ovlivňuje fyzické proporce člověka, jako je výška, váha, muskulatura apod. Přesto je důležité brát v potaz možné dynamické změny v behaviorálních znacích. Znamená to tedy, že tyto znaky nemusí být po celý život člověka stejné. Mezi vlastnosti zkoumané pro behaviorální biometrii patří například:

- Styl chůze
- Hlasový projev
- Mimika obličeje
- Pohyb rtů
- Podpis

1.2 Biometrické systémy

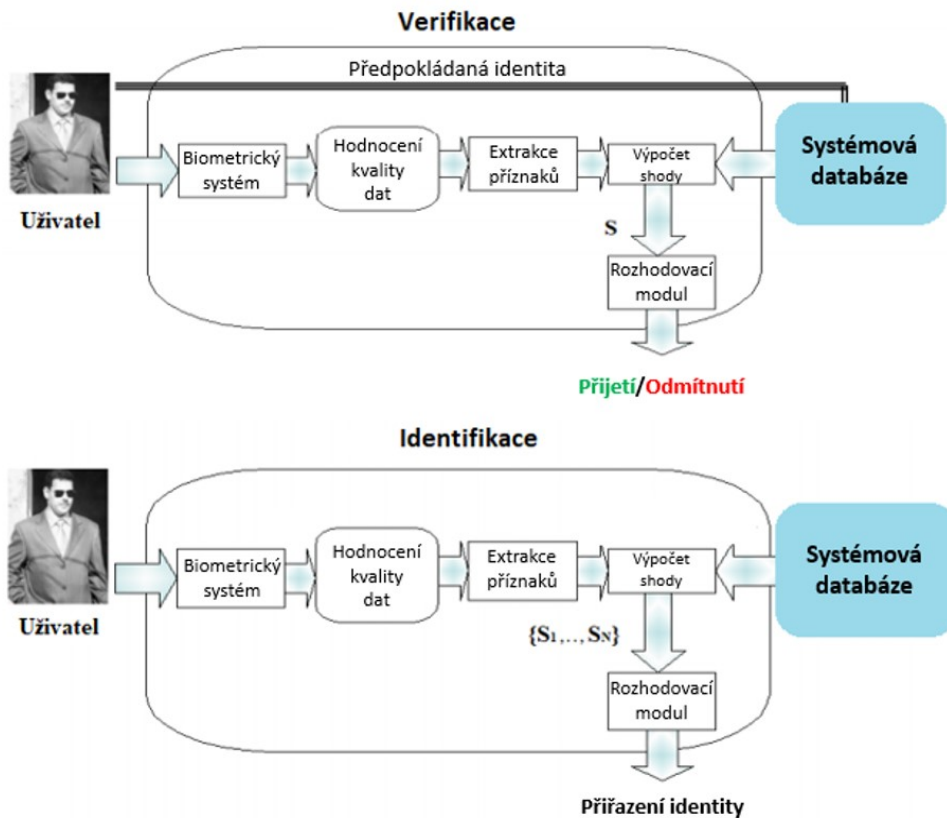
Složitost systému je dána množstvím porovnávaných znaků. Pokud systém pracuje pouze s jedním vstupním biometrickým znakem (např. otisk prstu), jedná se o *unimodální* systém. Pro pokročilejší systémy je však typická kombinace vícero biometrických markantů (např. otisk prstu + geometrie obličeje), což umožňuje vyšší stupeň zabezpečení. Tyto systémy se nazývají *multimodální*. Další variantou jsou hybridní systémy, které kombinují biometrické znaky s autentizačními (např. otisk prstu + heslo). Biometrické systémy mohou fungovat na dvou základních principech, kterými jsou verifikace a identifikace [2]. Schématický popis obou systémů je zobrazen na Obr. 1.1.

Verifikace

Verifikace je postup, při kterém se systému předloží biometrická data a předpoklad o identitě. Systém následně rozhodne, zdali tato data opravdu přísluší identitě konkrétní osoby na základě předchozích dat, uložených v databázi systému (popř. může být výstupem informace, do jaké míry data odpovídají identitě). Data se tedy porovnávají pouze s jedním uloženým vzorem. Toto prohledávání, takzvaně 1:1, přináší výhodu především díky nízké vypočetní a časové náročnosti algoritmu [2].

Identifikace

Identifikace je operace, při které se systému předloží biometrická data, avšak bez informace o příslušné identitě. Úlohou systému je tedy projít celou databází za účelem nalezení shody a následného přiřazení odpovídající identity (popř. seznam kandidátů s mírou příslušnosti dat a identity). Tento postup, takzvaně 1:N, je oproti verifikaci značně výpočetně i časově náročnější, v závislosti na velikosti prohledávané databáze [2].



Obr. 1.1: Schéma verifikačního a identifikačního systému [3]

Dalším rozlišením systému je druh požadavku odezvy. Z tohoto hlediska můžeme systémy rozlišovat na pozitivní a negativní.

Pozitivní požadavek odezvy systému

Pozitivním požadavkem rozumíme operaci, kdy systém má za úkol označit předložená biometrická data a předpokládanou identitu za odpovídající/nedopovídající, hledá však v databázi existující šablonu. Tento postup zabraňuje více osobám použít jednu identitu. Využití nachází například v přístupových zabezpečovacích systémech [3].

Negativní požadavek odezvy systému

Negativní požadavek naopak zabraňuje jedné osobě využít vícero identit. Systém prohledává databázi za účelem potvrzení, že příslušná osoba (resp. její biometrická data) se v databázi nenachází. Typickým příkladem takového požadavku je rozdělování sociálních dávek. Pokud byla dávka dané osobě již vyplacena, systém má její data uchovaná - dávka bude vyplacena pouze tomu, kdo ještě v systému není zadán [3].

1.3 Hodnocení kvality systému

Základní úloha verifikačního systému je porovnat vstupní vzorek se zapamatovanou šablonou z databáze. Odpovědí systému je tedy buďto přijetí či odmítnutí a obě tato rozhodnutí mohou být správná i nesprávná. Existují tedy čtyři možné výsledky rozhodnutí takového systému a na jejich základě byly zavedeny chybové funkce, sloužící k objektivnímu hodnocení spolehlivosti verifikačního systému:

1. **Míra chybného přijetí - False Acceptance Rate (FAR)**
2. **Míra chybného odmítnutí - False Rejection Rate (FRR)**

U identifikačních systémů většinou není výsledkem jedna konkrétní identita, avšak seznam kandidátů, přičemž každému z nich je přiřazena míra shody (s). Tato hodnota se pohybuje v rozmezí intervalu $\langle 0,1 \rangle$. Systém je většinou opatřen volbou prahu T , který určuje, v jakém rozmezí shody ještě systém označí šablonu jako potencionálního kandidáta. Volba prahu má zásadní vliv na výsledky identifikace - může být příliš benevolentní nebo příliš přísný - což se projeví při výpočtu následujících chybových funkcí [3]:

1. **Míra chybné shody - False Match Rate (FMR)**

vyjadřuje pravděpodobnost, že systém označí biometrická data jedné osoby za shodná s daty jiné osoby. Matematicky je vyjádřena takto:

$$FMR(T) = \int_T^1 p(s|H_1)ds, \quad (1.1)$$

kde T je rozhodovací práh, H_1 je výrok, že vzor a šablona nejsou jedna identita, p je pravděpodobnostní hustota, že výrok je pravdivý a s je míra shody.

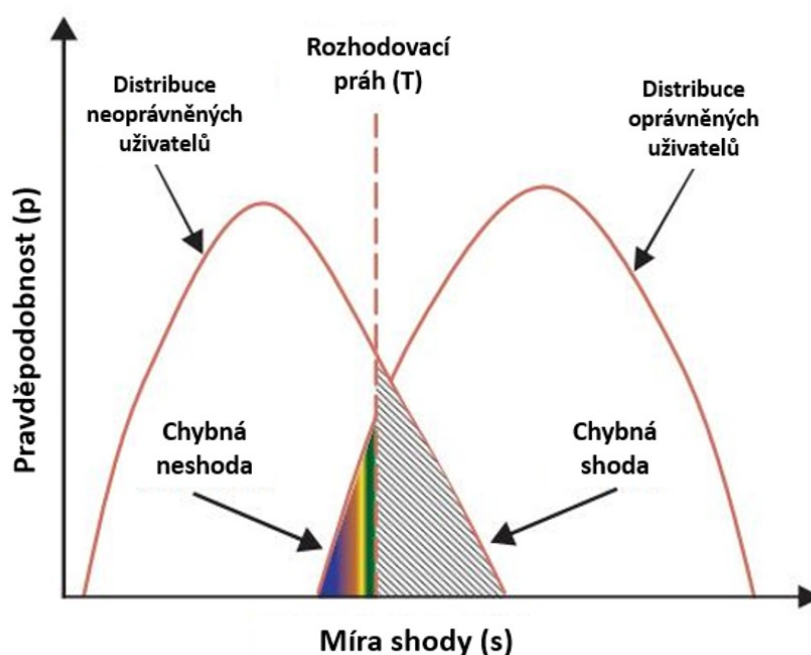
2. Míra chybné neshody - False Non-Match Rate (FNMR)

vyjadřuje pravděpodobnost, že systém neoznačí shodu dat vstupních s odpovídajícími daty v databázi. Matematicky je vyjádřena takto:

$$FNMR(T) = \int_0^T p(s|H_0)ds \quad (1.2)$$

kde T je rozhodovací práh, H_0 je výrok, že vzor a šablona jsou jedna identita, p je pravděpodobnostní hustota, že výrok je pravdivý a s je míra shody.

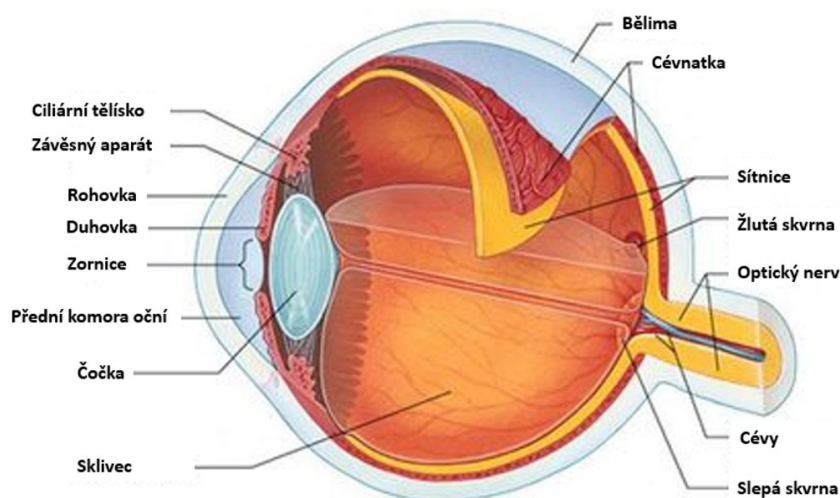
Volba prahu se často odvíjí od výsledků těchto chybových funkcí a může být nastaven na hodnotu takovou, kdy se $FMR = FNMR$. Tento bod se nazývá míra vyrovnání chyb - Equal Error Rate (EER). Pokud bude práh T nastaven na hodnotu EER, bude mít systém stejnou pravděpodobnost výskytu chybné shody i chybné neshody. Příklad možného nastavení prahu je zobrazen na Obr. 1.2. Volba prahu se také odvíjí od oblasti využití identifikačního systému. Například ve forenzní sféře se nastavuje práh pro větší hodnotu chybné shody, zatímco u zabezpečujících systémů je žádoucí spíše větší pravděpodobnost chybné neshody [3].



Obr. 1.2: Pravděpodobnostní rozložení chybových funkcí FMR a FNMR [4].

2 Anatomie oka a sítnice

Oko je párový smyslový orgán, který zprostředkovává převod světelného podnětu na elektrický, což umožňuje jeho zpracování nervovou soustavou a tedy vznik zrakového vjemu. Anatomie oka je popsána na Obr. 2.1. Lidské oko tvoří soustavu světlolomných prostředí, soustřeďujících světelné paprsky na světločivou vrstvu oka. Světelný paprsek prochází zornicí a dopadá na čočku, která je prvním důležitým světlolomným prvkem oka. Čočka umožní zaostřit paprsek tak, aby po průchodu sklivcem dopadal ideálně na jeden bod na sítnici.

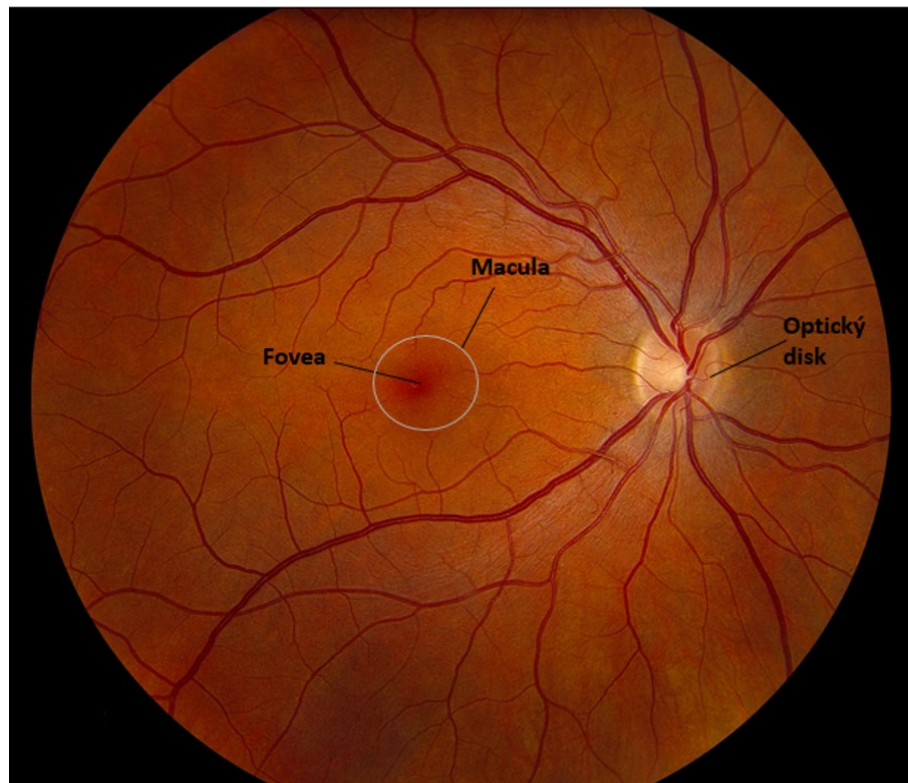


Obr. 2.1: Anatomie lidského oka [7]

2.1 Sítnice oka

Sítnice je tenká vrstva v zadní vnitřní části oka, která obsahuje zrakové receptory - tyčinky a čípky. Na tuto vrstvu je soustředěn světelný paprsek, který dráždí příslušné receptory. V receptorech dochází k elektrochemické reakci za vzniku akčního napětí, které se dále nervovými zakončeními šíří do zrakového centra v mozku. V místě, kde se čípky vyskytují s největší hustotou, se nachází žlutá skvrna (fovea), tedy bod neostřejšího vidění. Její okolí se nazývá macula. Hustota čípků směrem od tohoto bodu postupně klesá. Dalším významným bodem na sítnici je slepá skvrna, tedy místo neobsahující žádné receptory. Nachází se v bodě, kde z oka vystupuje zrakový nerv. Z pohledu snímku sítnice se tato oblast nazývá optický disk [5]. Snímek sítnice pořízený fundus kamerou je znázorněn na Obr. 2.2.

Receptory v sítnici musí být pro správnou funkci dostatečně zásobované kyslíkem, který je přiváděn dvěma cévními systémy. Centrální sítnicová tepna vstupuje do oka v oblasti optického disku a vyživuje vnitřní stranu sítnice. Větví se obvykle na dvě hlavní části - horní a dolní, které se nadále rozvětvují. Druhý cévní systém, zajišťující prokrvení vnější strany sítnice a receptorů, je tvořen hustou sítí jemných kapilár cévnatky. Strukturu cév sítnice lze zachytit pomocí oftalmologických zařízení, což poskytuje možnosti jak pro diagnostiku, tak pro biometrii [6].



Obr. 2.2: Příklad snímku zdravé sítnice s vyznačenými anatomickými body. [8]

2.2 Patologie sítnice

Sítnice je vhodným nástrojem pro biometrii především z toho důvodu, že je sítnice uvnitř oka dobře chráněná proti vnějším vlivům. Je ovšem nutné zmínit určitá onemocnění, která mají za následek změny v sítnici, což ztěžuje či dokonce znemožňuje identifikaci. Poruchy sítnice jsou ve většině případů nevratné a často mohou být příčinou slepoty. Mezi tyto poruchy patří:

Makulární degenerace

Makulární degenerace patří mezi nejčastější příčiny slepoty. Je z části podmíněná geneticky, avšak významným ovlivňujícím faktorem u tohoto onemocnění je věk a životní styl [9]. Existuje tzv. suchá a vlhká forma tohoto onemocnění. U vlhké formy dochází k nekontrolovanému růstu cév za sítnicí, což vede k prosakování krve. Tlak krve způsobí zprohýbání sítnice a odumření zrakových receptorů v okolí žluté skvrny (makuly), tedy místa nejostřejšího vidění. To má za následek snížení schopnosti zaostřit. Pacienti mají zpočátku problém s viděním v centrálním zorném poli, postupně se porucha zraku rozšiřuje i na periferní vidění [10].

Vlhká forma je částečně léčitelná, avšak postupuje velmi rychle. Na suchou formu zatím nebyla objevena léčba. Z hlediska snímků sítnice se toto onemocnění projeví shlukem skvrn v oblasti makuly, které reprezentují místa degenerovaných pigmentových buněk [10].

Diabetická retinopatie (DR)

Nedostatečná léčba cukrovky může mít za následek poškození zraku. Nadbytečné množství glukózy v krvi se ukládá v jemných cévách sítnice, což vede k jejich ucpaní či poškození. Průběh tohoto onemocnění je bezbolestný, a tak se většinou projeví až v pokročilejším stadiu. Důležité jsou proto preventivní prohlídky, zejména u diabetických pacientů. V případě odhalení diabetické retinopatie je možné onemocnění částečně léčit chirurgicky, nebo za použití laseru [9].

Odchlípení sítnice

Při odchlípení sítnice dochází k jejímu natržení, což způsobí uvonění krve z cévy sítnice do prostoru sklivce. To se nejprve projeví skvrnami v zorném poli pacienta, které způsobují přítomné červené krvinky. Při dalším krvácení do prostoru sklivce může být zrak zcela zastíněn. Při včasném odhalení problému je dnes již možné odchlípnutou sítnici chirurgicky napravit. Základem diagnostiky je viditelná trhлина, popřípadě krvácení na snímku sítnice [9].

Další onemocnění

Mezi další onemocnění patří například záněty sítnice. Dále i přímý kontakt sítnice se sklivcem může způsobovat deformace v zorném poli. Ačkoliv se nejedná o onemocnění sítnice, je pro tuto práci vhodné zmínit také oční zákal. Šedý zákal (katarakta) postihuje oční čočku a její zakalení se projeví i na snímku sítnice, což může značně zhoršit proces identifikace [9]. Zelený zákal (glaukom) je způsoben degenerací optického nervu a na snímcích sítnice jsou patrné změny v oblasti optického disku.

2.3 Měřicí přístroje

Přestože v této práci použijeme snímky sítnice pro biometrii, běžně se zařízení umožňující jejich záznam primárně používají pro klinické účely. Pořízené snímky mohou mít podstatnou úlohu při odhalení počátečních indikací některých zmíněných onemocnění. Pro tuto práci je vhodné popsat základní princip přístrojů, které se pro snímání sítnice používají.

Fundus kamera

Fundus kamery poskytují možnost zobrazit oční pozadí (fundus), a tedy cévní řečiště zásobující oko. Jedná se o speciální druh mikroskopu opatřený fotoaparátem. Zařízení je vybaveno soustavou optických filtrů a zrcadel, které fokusují prstenec infračerveného světla na oční sítnici. Světelný paprsek prochází zornicí a šíří se dále jednotlivými vrstvami oka, kde dochází ke konstantnímu útlumu jeho intenzity dle vlastností tkáně. Princip zobrazování je v měření proměnlivého útlumu, který pozorujeme v cévách sítnice [11]. Příklad snímku pořízeného fundus kamerou je na Obr. 2.2.

Množství absorbovaného světla v sítnici se mění v závislosti na aktuálním množství krve dle srdečního cyklu. Světlo, modifikované vlivem absorbcí všech tkání, se od sítnice odráží a vystupuje zornicí zpět do zařízení fundus kamery, kde je detekováno a fotoaparátem zachyceno.

Fotoaparát u tohoto zařízení může pracovat s různými velikostmi zorného pole a má tak zásadní vliv na podobu výsledného snímku. Širokoúhlé fotoaparáty mají zorný úhel od 45° do 140° a poskytují proporcionálně menší zvětšení sítnice. Ty s malým zorným úhlem mají potom méně než 20° [11].

Video-oftalmoskop

Tento přístroj poskytuje možnost sledovat vývoj změn sítnice v čase. Z tohoto hlediska jsou však na systém kladeny vysoké technické požadavky, které je zapotřebí dodržet. Hlavním z nich je nutnost synchronizace snímání obou očí, aby byla zajištěna časová přesnost pulzních změn v sítnici. Pokud by byl systém časově nestálý, mohly by se neodpovídající změny na sítnicích obou očí jevit jako patologie. Zároveň je žádoucí vysoká citlivost systému i na malé objemové změny v prokrvení. To je zajištěno vhodným výběrem světla o vlnové délce 577 nm, která je krví nejvíce pohlcována. Narozdíl od klasických Fundus kamer se zde jako zdroj světla používá LED dioda [12].

Důležitá je také fixace očí na jeden bod, aby nedocházelo ke znehodnocení záznamu pohybem očí pacienta. I přes fixaci se však pohybové artefakty v záznamu vyskytují a mohou být odstraněny fázovou korelací snímků s vybraným referenčním

snímkem (obvykle první snímek sekvence). Pro ještě preciznější odstranění pohybových artefaktů se pokračuje porovnáváním souřadnic vybraných bodů na segmentovaném cévním stromu sítnice [12].

3 Biometrie sítnice

Tato kapitola popisuje základní postup, uplatněný při biometrii sítnice. Identifikaci na základě struktury cév oční sítnice umožňuje její unikátnost pro každého jedince. Ta platí i v případě jednovaječných dvojčat [13]. Následující podkapitoly se věnují jednotlivým blokům základního biometrického algoritmu.

3.1 Akvizice snímku

Pro získání snímku sítnice lze použít některý z již zmíněných oftalmologických přístrojů. Tato fáze má zásadní vliv na konečný výsledek identifikace, neboť vstupní data mohou být při nedostatečné spolupráci měřeného pacienta nekvalitní. Pro správnou akvizici snímku sítnice je potřeba zajistit vhodné podmínky. Mezi základní z nich patří snímání ve vnitřním prostředí, při definovaném konstantním osvětlení [13]. Je také zapotřebí eliminovat pohyb očí i pacienta, odstranit předměty, které by mohly výsledný snímek ovlivnit (brýle, kontaktní čočky) a také zaručit dostatečnou relaxaci zornice pro průchod světelného paprsku [14].

Surová data následně prochází předzpracováním, během kterého je možné dodatečně upravit snímky, které budou použity pro identifikaci (odstranění šumu, artefaktů apod.). Pro biometrii sítnice může být dalším přínosným krokem segmentace.

3.2 Segmentace cévního řečiště

Segmentace je jedna z nejzákladnějších morfologických úprav, používaných pro zpracování obrazu. Na segmentovaném obrazu dochází k vymezení prostoru, který by v ideálním případě měl odpovídat ohraničené oblasti či předmětu, vyobrazeném na snímku. Ne vždy je toto možné zaručit, neboť tvar a velikost objektů na snímku nemusí vždy odpovídat realitě, například kvůli nevhodnému nasvětlení při akvizici.

Výstupem segmentace je obraz o stejné velikosti jako obraz původní [15]. Dle typu použitého algoritmu může být výsledný obraz binární, RGB nebo indexovaný, v závislosti na počtu dělených oblastí.

Pro konkrétní úlohu segmentace cévního řečiště sítnice lze použít nejznámější metody, jako je dělení a slučování oblastí, metoda narůstání oblastí, metoda rozvodí či modifikace Houghovy transformace. Velmi oblíbenou metodou je však přizpůsobená filtrace, kdy je obraz korelován s maskami, které obsahují vzor cévy (v různých šířkách a orientacích). Problémem u segmentace cévního řečiště je obvyklé narušení spojitosti cév, proto je vhodné následně tyto artefakty odstranit [15]. Pro ještě větší efektivitu lze pro segmentaci využít neuronové sítě [16].

3.3 Extrakce příznaků

Příznaky lze chápat jako číselné ohodnocení významných bodů či oblastí cévního stromu, které mohou být reprezentativní pro identifikaci. V praxi se využívají různé kombinace příznaků, z důvodu zvýšení spolehlivosti identifikačního systému. Hodnoty příznaků jsou uloženy v příznakovém vektoru, který je jednoznačně přidělen původnímu obrazu, ve formě referenční šablony a provádí se jejich srovnání s předloženými daty. Můžeme porovnávat jejich množství, pozice či jiné parametry [17]. Příznaky rozdělujeme na globální a lokální. Mezi globální příznaky se řadí například celkový počet pixelů, náležící segmentovanému cévnímu stromu a umožňuje prvotní vyloučení nevyhovující části šablon z databáze. Až poté se provádí porovnání lokálních příznaků s omezenou množinou šablon. Tento postup zvyšuje časovou efektivitu systému [17]. Mezi lokální příznaky patří například:

Bifurkace (rozvětvení) a koncový bod cévy

Lokalizace cévního větvení je jedním z nejčastějších příznaků. Pro extrakci takového příznaku se jeví jako nejvhodnější nejprve provést skeletonizaci segmentovaného cévního řečiště. Výstupem je totiž binární obraz, který obsahuje středové linie, odpovídající průběhu segmentovaných cév. Tyto linie mají tloušťku pouze jeden pixel, což značně zlehčuje možnosti detekce bifurkací [18].

Na segmentovaný obraz je aplikována maska, která je posouvána po obraze a vyhodnocuje, kolik sousedů s hodnotou 1 má aktuální pixel. Dva sousední pixely poukazují na normální průběh cévy, jeden sousední pixel odpovídá koncovému bodu a tři nebo více sousedních pixelů získáme v místě bifurkace, případně v místě křížení dvou cév [18]. Další možností, odvíjející se od nalezení bifurkací je určení úhlu větvení.

Pozice a velikost optického disku

Možností lokalizace optického disku je více. Je možné vycházet z původních snímků sítnice, nebo použít již segmentovaná data. Optický disk se na snímcích z fundus kamery či video-oftalmoskopu jeví jako světlejší kulovitý útvar. V tomto případě lze využít Houghovu transformaci pro detekci kružnic. Jejím výstupem jsou souřadnice středu nalezené kružnice a její poloměr [19].

V druhém případě, tedy u segmentovaných dat, není informace o poloze optického disku zachována v jasových hodnotách. Můžeme však využít faktu, že optický disk je centrem vstupu i výstupu cév, zásobujících oko. Oblast optického disku je tedy typická nejvyšší hustotou a zároveň šířkou těchto cév. Samotná detekce poté

probíhá tak, že po segmentovaném obraze se posouvá okno, obsahující kruhový obrazec, vyplněný hodnotou 1. Výstup korelace této masky s obrazem bude nejvyšší právě v místě, kde je na segmentovaném obraze nejvíce pixelů s hodnotou 1. Toto místo by ideálně mělo odpovídat optickému disku. Střed obrazce v poloze s nejvyšším korelačním koeficientem je následně označen za střed optického disku [19].

Pozice bifurkací vzhledem k optickému disku

Pro zvýšení robustnosti identifikačního algoritmu je možné využít kombinaci předchozích příznaků. Problém samotné lokalizace bifurkací či optického disku je ten, že při různých akvizicích se mohou jejich souřadnice měnit. Dochází k vzájemné translaci i rotaci. Nezmění se však jejich vzájemná poloha, změny budou pouze poměrové. Střed optického disku lze pro jednodušší přepočítání nastavit na počátek nového souřadnicového systému [20].

3.4 Vyhodnocení podobnosti

Příznaky na vstupu jsou porovnávány s šablonami z databáze (při procesu identifikace), popřípadě s vybranou konkrétní šablonou (v případě verifikace). Je důležité zmínit, že nejčastěji bývá výstupem identifikačního systému fuzzy množina, poskytující informace o míře podobnosti vstupních dat a dostupných šablon [21]. Pro porovnání lze použít různé druhy metrik a na základě nejvyšší shody přiřadit odpovídající identitu. Mezi nejčastěji používané rozhodovací metody patří:

Vzdálenost příznakových vektorů

Nejjednodušším a nejrychlejším způsobem porovnání dvou vektorů je stanovení jejich vzdálenosti. Tato vzdálenost může být vypočítána různými metrikami. Typickým příkladem je Euklidovská vzdálenost (vypočítaná vztahem 3.1. [22]) resp. Manhattanská vzdálenost. V obou případech je cílem identifikačního programu najít a označit v databázi vektor s nejmenší vzdáleností od vstupního vektoru. Dva totožné vektory příznaků mají vzdálenost nulovou, s rostoucím výskytem odlišností se vzdálenost zvyšuje.

$$d_E = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, \quad (3.1)$$

kde n je délka vektoru a x_i, y_i jsou souřadnice prvku vektoru na pozici i .

Korelace

Korelace je matematická operace, hodnotící lineární závislost dvou příznakových vektorů. Ve výsledku získáme hodnotu (Pearsonova) korelačního koeficientu r , která náleží intervalu $\langle -1, 1 \rangle$. Korelační koeficient, odpovídající hodnotě 1, značí přímou závislost vektorů. Hodnotu 0 dostaneme na výstupu u nezávislých a -1 v případě nepřímo závislých vektorů. [23]. Výpočet je reprezentován vzorcem 3.2 [24].

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (3.2)$$

kde r je korelační koeficient, n je délka vektoru a x_i, y_i jsou souřadnice prvku vektoru na pozici i .

Template matching

Metoda template matching (TM) spočívá v porovnání vzorů konkrétních tvarů. Tyto vzory vzniknou tak, že z původního segmentovaného obrazu se zachovají pouze vyznačené pozice významných lokalizovaných bodů. Nejčastěji jsou to již zmiňovaná větvení či zakončení cév. Linie těchto bodů vytváří určitý obrazec, který je považován za referenční vzor. Rozdíl oproti klasické korelaci spočívá v rozdělení referenčního obrazu na segmenty za vzniku menší šablony [25].

Posouváním šablony po obraze se jejich korelací hledá shoda obrazců. Lze tak kompenzovat translaci souřadnic bifurkací. Problémem však zůstává možnost natočení bifurkačních bodů, kterou tato metoda neřeší. V praxi je však translace mnohem markantnější, rotace může být správným upevněním pozice hlavy pacienta v průběhu akvizice omezena [26].

Typickým využitím TM metody je například strojové čtení, kdy za vzor je považováno konkrétní písmeno. Geometrie písmene je ojedinělá a po porovnání tohoto vzoru s celým obrazem získáme informace o výskytu daného písmene. Teorie porovnávání vzorů je také spojována s principem činnosti lidského mozku [26].

3.5 Support Vector Machine (SVM)

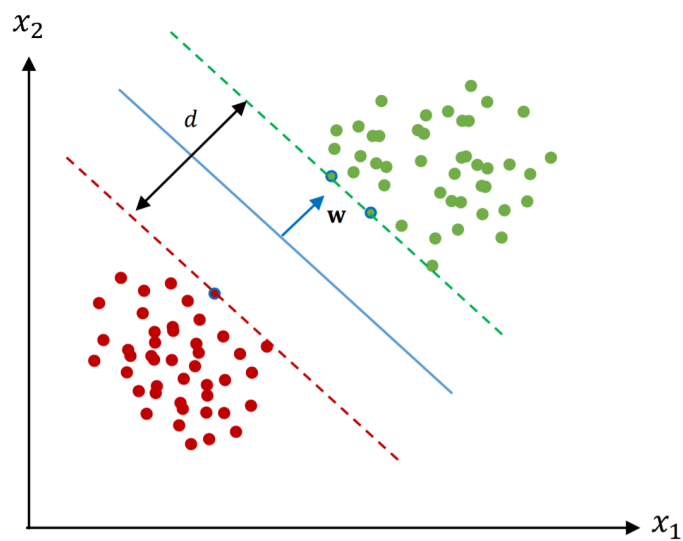
Support Vector Machine je metoda strojového učení, která slouží zejména k binární klasifikaci dat. Využívá způsob lineární či nelineární klasifikace. Na vstup SVM se předloží trénovací množina, obsahující data a informaci o odpovídající kategorii.

Data jsou reprezentována vektorem příznaků (x_1, \dots, x_n) v příznakovém prostoru, jehož rozměry závisí na jejich počtu. Pokud jsou příznaky jednotlivých kategorií dostatečně odlišné, v příznakovém prostoru by měly tvořit oddělené shluky. SVM se snaží proložit příznakový prostor dat takovou křivkou, která by tyto shluky, a tedy jednotlivé kategorie vhodně rozdělovala [27].

Lineární SVM

Pro proložení se využívají krajní příznaky oddělovaných kategorií. Těmto příznakům se říká pomocné vektory, z anglického support vectors. Počet těchto vektorů závisí na rozložení trénovacích dat. Oddělující křivka může například procházet středem mezi těmito body. Testovací data jsou poté touto křivkou proložena a je vypočítána přesnost klasifikace [27].

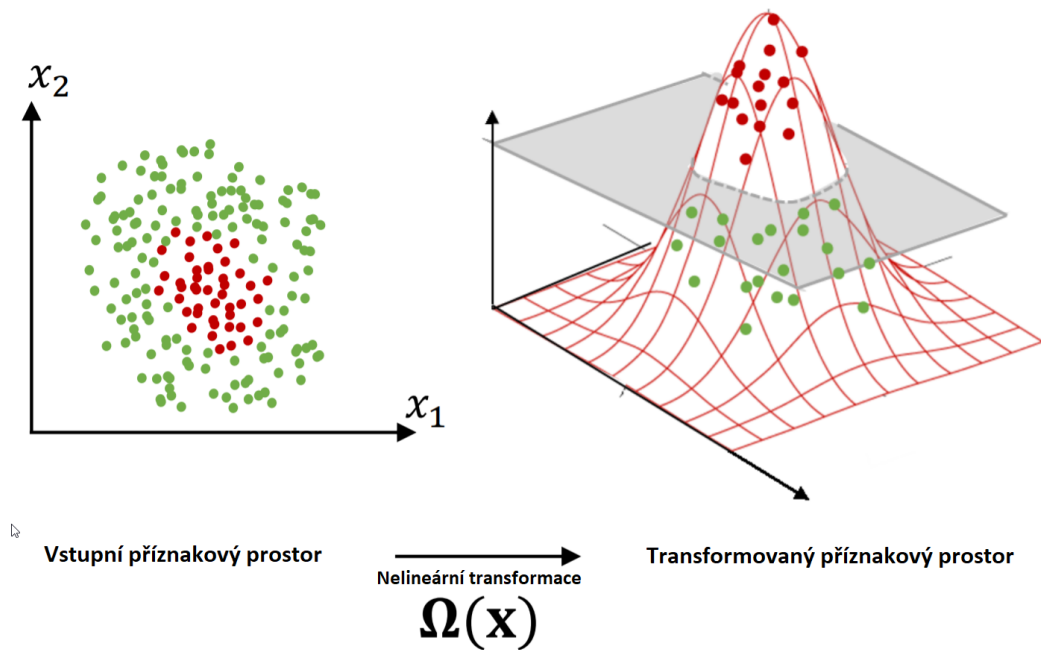
U reálných dat se často setkáváme s kategoriemi, které nejsou takto jednoduše oddělitelné, a je tedy nutné připustit určité misklasifikace. Kromě vhodného sklonu oddělující křivky, tak lze také nastavit její okraje, jako vzdálenost k okrajovým pomocným vektorům (w). Tímto způsobem lze definovat jakési oddělující pásmo (d). Cílem je nastavit okolí křivky tak, aby bylo schopné pojmout případné misklasifikace [27], [28]. Příklad rozdělení dat je znázorněn na Obr. 3.1.



Obr. 3.1: Lineární SVM klasifikace [28].

Nelineární SVM

Při složitějším rozložení dat v příznakovém prostoru lze tento prostor nelineárně transformovat. Často se provádí transformace na vyšší dimenzi. Cílem je vytvořit nový příznakový prostor, ve kterém by příznakové reprezentace dat byly lépe separovatelné. Tato transformace jde na úkor zvýšení výpočetní i paměťové náročnosti klasifikace [28]. Princip klasifikace ve vícerozměrném prostoru je patrný z Obr. 3.2.



Obr. 3.2: Nelineární SVM klasifikace. Přepřacováno z [28].

4 Umělé neuronové sítě (ANN, NN)

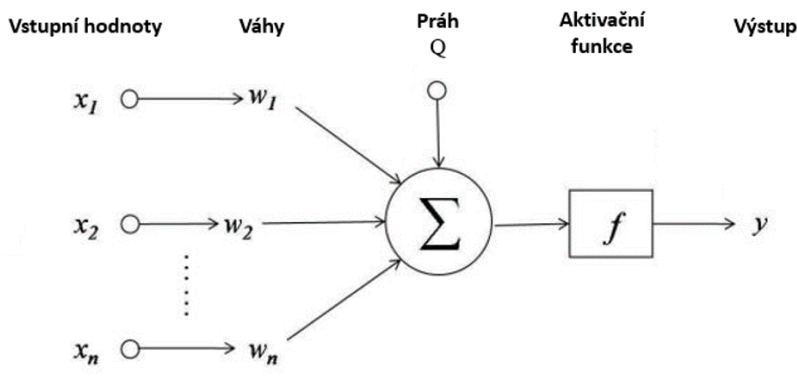
Umělé neuronové sítě (ANN, NN) jsou v poslední době nejúčinnějším prostředkem nejen pro klasifikaci dat. Princip NN spočívá v simulaci funkce skutečného biologického mozku. Síť je tvořena elementárními prvky - umělými neurony, nazývané perceptrony. Počet a uspořádání těchto elementů udává topologii (architekturu) neuronové sítě, která má zásadní vliv na výsledek klasifikace.

Perceptrony provádí nezávislé matematické operace se vstupními hodnotami, které jsou váhované. Díky váhám perceptron rozhoduje, jak moc bude daný vstup brát v úvahu při výpočtu výstupní hodnoty. Rozhodování sítě je uskutečněno pomocí prahu a aktivační funkce každého perceptronu. Výstup z aktuálního perceptronu postupuje dle topologie na vstupy další vrstvy a jeho výpočet je patrný ze vzorce 4.1 [29]. Schéma perceptronu je znázorněné na Obr. 4.1.

$$Y = \sum_0^n (w_n * x_n) + \theta, \quad (4.1)$$

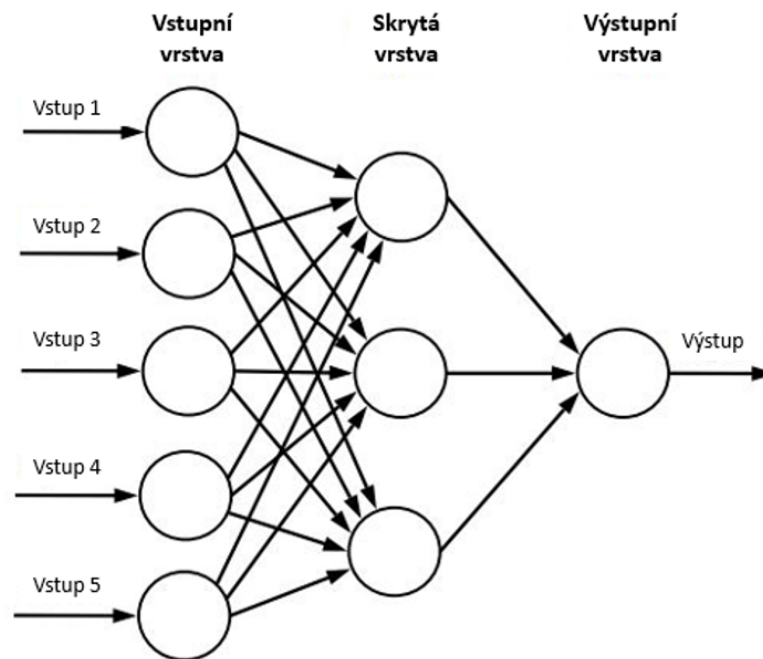
kde Y je výstup perceptronu, x_n je vstup, w_n hodnota váhy a θ je práh.

Váhy jsou inicializované na určitou hodnotu, která se v průběhu učení sítě mění. Optimalizace vah probíhá na základě různých algoritmů. Může se použít metoda zpětného šíření chyby či gradientní sestup s použitím k-násobné křížové validace. Při tomto postupu je s jistou opakovací frekvencí síti předložena validační množina (disjunktní s množinou trénovací) a na základě zjištěných odchylek predikce modelu se váhy iteračně upravují, dokud není dosaženo požadované přesnosti či počtu epoch. Architektura jednoduché neuronové sítě je znázorněna na Obr. 4.2.



Obr. 4.1: Schéma perceptronu neuronové sítě. Přepřacováno z [30].

ANN vykazují vysokou úroveň klasifikace již při topologii sítě s jednou vnořenou vrstvou [30]. Přidáním dalších vrstev se zvyšuje robustnost klasifikace. NN, obsahující více než tři skryté vrstvy, se označují jako hluboké neuronové sítě. Tento způsob učení je dnes velmi vyhledávaný v oblasti počítačového vidění a klasifikace obrazů [30].



Obr. 4.2: Schéma propojení vrstev klasické neuronové sítě. Přepřacováno z [31].

4.1 Hluboké učení (deep learning)

Moderním nástrojem v klasifikaci obrazů je metoda hlubokého učení. Při hlubokém učení se dodržuje přísná hierarchie vrstev. Postupuje se k hlubším vrstvám a každá vrstva poskytuje na výstupu odlišný druh příznakové informace o vstupních datech. Rozdíl oproti klasickému učení sítě je v extrakci příznaků. U klasického přístupu jsou na vstup NN přiváděny kompletní příznakové vektory, které byly extrahovány předchozí analýzou dat. Naproti tomu u hlubokého učení jsou na vstup přiváděna obrazová data a extrakce příznaků je součástí celého algoritmu NN. Hlavní výhodou je tedy zjednodušení a urychlení celého procesu, neboť návrh samostatného algoritmu pro extrakci příznaků bývá časově náročný [31].

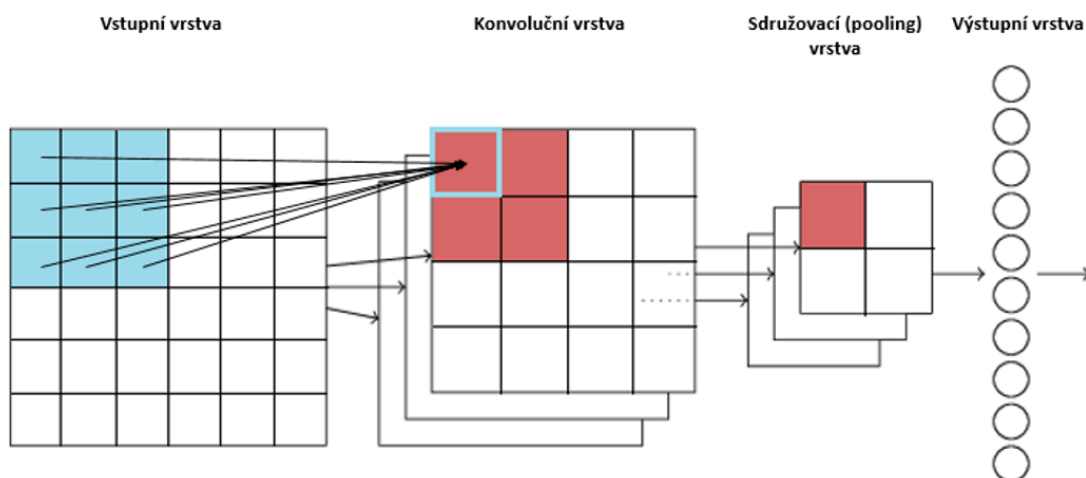
Informace z příznaků, získaných na výstupu jednotlivých vrstev, se s postupem k hlubším vrstvám v hierarchii nabalují. Čím hlubší je aktuální vrstva algoritmu,

tím blíže se výstup podobá konečnému rozhodnutí neuronové sítě. Učení sítě probíhá opět nejčastěji způsobem zpětného šíření chyby (back propagation), kdy se iteračně upravují hodnoty vah jednotlivých perceptronů [31].

Výhodou hlubokého učení je vysoká úroveň abstrakce nad množinou trénovacích dat. Hlubokého učení využívají komplexní NN, které obsahují zpravidla více než 3 vrstvy. Typickým příkladem využití jsou konvoluční neuronové sítě. V dnešní době se stále zdokonalují prostředí, ve kterých je možné hluboké učení realizovat. Tato prostředí navíc podporují širokou škálu programovacích jazyků a zároveň obsahují knihovny s přednastavenými NN různých topologií [31].

4.2 Konvoluční neuronové sítě (CNN)

Vícevrstvé neuronové sítě, obsahující alespoň jednu konvoluční vrstvu, se nazývají konvoluční neuronové sítě (CNN). Architektura CNN často obsahuje kombinaci konvolučních vrstev s vrstvami sdružovacími, popřípadě plně propojenými. Konvoluční vrstvy jsou obvykle ty skryté, na ně poté navazují sdružovací vrstvy a nakonec plně propojené vrstvy na výstupu [31]. Princip konvoluční NN je zobrazen na Obr. 4.3.



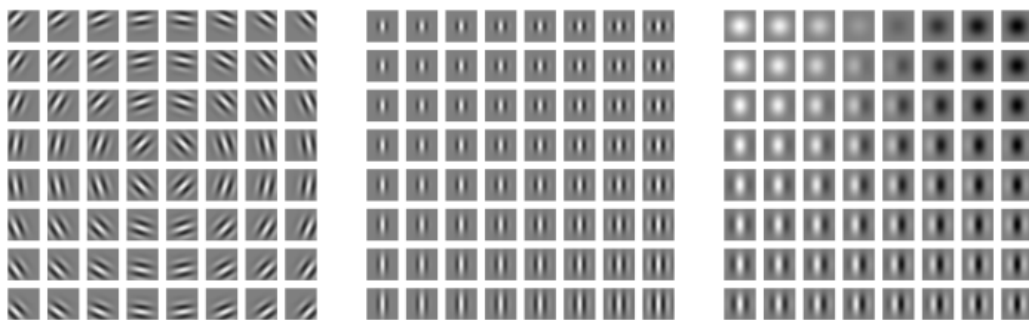
Obr. 4.3: Schéma propojení vrstev konvoluční neuronové sítě. Přepřacováno z [31].

Konvoluční vrstva

V klasických vrstvách NN jsou všechny perceptrony vzájemně propojené. U konvoluční vrstvy jsou propojení pouze částečná a formují tak určité shluky perceptronů. Tyto shluky slouží jako filtry, někdy také nazývané kernely. Jejich velikost bývá obvykle 3x3 nebo 5x5 pixelů, tedy výrazně menší, než velikost vstupních dat. Každý

tento kernel se posouvá po původním obraze a pomocí konvoluce se získá výstup na tomto konkrétním shluku perceptronů konvoluční vrstvy. Oblast obrazu, která vstupuje do jednoho perceptronu skryté vrstvy, se nazývá lokální receptorové pole. [30].

Každý perceptron konvoluční vrstvy je stejně jako u klasických NN definován prahem a váhami, které se v průběhu učení iteračně posouvají směrem k optimálním hodnotám. Díky těmto parametrům rozhoduje perceptron, jaká informace projde na výstup neuronu. Každý kernel konvoluční vrstvy umožňuje získání jiného druhu příznaku ze vstupního obrazu, neboť vyhledává v obraze jiný vzor [31][30]. Příklad vyhledávaných vzorů lze vidět na Obr. 4.4.



Obr. 4.4: Příklad vzorů uložených v kernelech CNN [31].

Díky tomuto uspořádání vykazují CNN selektivní charakter, zatímco klasické NN zpracovávají hodnoty z celého vstupu. Díky tomu je již s použitím jedné konvoluční vrstvy možné vyhledávat velké množství struktur, orientovaných hran, barevných kombinací apod. Přenos výstupu z konvoluční vrstvy je stejně jako u klasických NN podmíněn tvarem aktivační funkce. Shluky konvoluční vrstvy obsahují menší počet perceptronů a navíc jsou definovány tzv. sdílenými váhami. Toto sdílení parametrů značně sníží výpočetní náročnost [31].

Aktivační funkce

Aktivační funkce je matematická rovnice, určující výstup z každého perceptronu dané vrstvy. Výsledná hodnota po průchodu váhovaného vstupu je tak posunuta na další úroveň sítě. Aktivační funkce navíc normalizuje data, která následně putují dál do sítě. Podle jejího tvaru je daný výstup v rozmezí 0 a 1, popřípadě -1 a 1. Aktivační funkce může mít různý průběh, například [32]:

- binární schodovitá funkce,
- sigmoida,

- ReLu funkce
- softmax (nejčastěji na výstupní vrstvě).

Sdružovací vrstva

Sdružovací, neboli pooling vrstva, navazuje na vrstvu konvoluční a její podstatou je podzorkování výstupu konvoluční vrstvy. Ten obsahuje velké množství informací, vzhledem k množství získaných příznaků. Určitá oblast výstupu konvoluční vrstvy (nejčastěji 2x2 pixely) je tak na výstupu vrstvy sdružovací reprezentována jednou hodnotou. Ta může být získána jako průměr, či maximum z dané oblasti [31].

Plně propojená vrstva

Tato vrstva bývá umístěna až na závěr celé architektury konvoluční sítě. Její význam spočívá v konečném přiřazení vstupního obrazu do konkrétní kategorie. Slouží tedy jako klasifikátor. Perceptrony v této vrstvě jsou plně propojené, každý výstup předchozí vrstvy je tedy přiveden na každý vstup této vrstvy. Počet perceptronů je volen dle požadovaného počtu klasifikačních tříd [31].

4.3 Transfer Learning

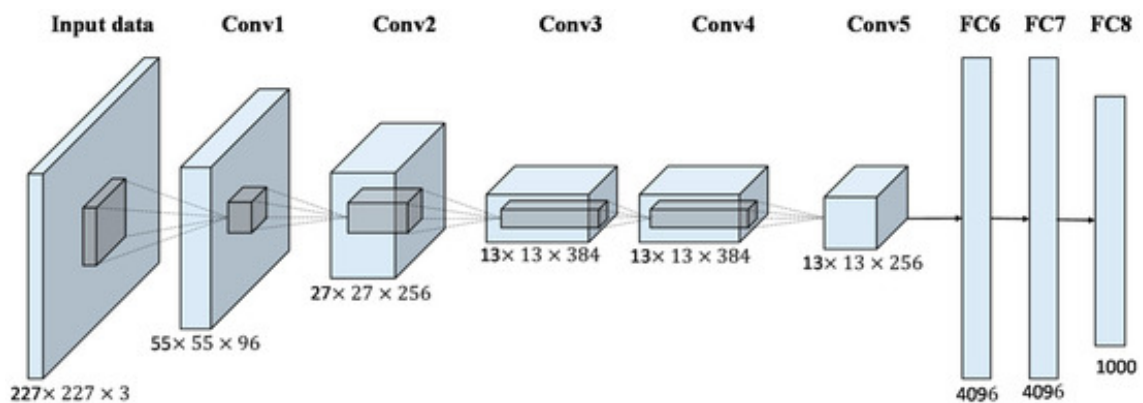
V dnešní době je možné v některých online knihovnách nalézt modely různých architektur sítí, které jsou již předučené na vlastních datech. Těmi je zpravidla databáze ImageNet, která obsahuje 1,2 milionu snímků různých objektů, náležících do 1000 kategorií (např. dopravní prostředky, zvířata, apod.) [33]. Takto natrénovaná síť má své váhy optimálně upravené na původní vstupní data a při řešení nového klasifikačního problému lze převzít její topologii, nebo dokonce její váhy. Možnost volby z širokého výběru již zkonstruovaných topologií sítě, usnadní postup při návrhu vlastní NN. Využití vah z již naučené sítě značně snižuje výpočetní čas, potřebný pro její natrénování, jelikož váhy se pouze přizpůsobí novému souboru dat.

Metoda Transfer Learning spočívá v drobné úpravě předučené sítě. Největší změnou prochází obvykle výstupní, plně propojená vrstva. Ta může být odstraněna a nahradí ji nová vrstva, která obsahuje počet perceptronů, dle námi požadovaného počtu klasifikačních tříd. Na CNN se následně aplikuje nová učební množina dat z vlastní databáze a parametry sítě se podle ní dále upravují dle zvoleného způsobu učení (např. zpětné šíření chyby). V některých případech je také možné určitě hodnoty vah takzvaně zafixovat, tzn. že nebudou podléhat dalším změnám při doučování sítě, a zamezí se tak možnosti přeučení [34]. Zafixování vah se provádí především u nižších skrytých vrstev, neboť tyto vrstvy rozlišují obecnější příznaky v obrazech.

Parametry se tedy dle vlastní učební množiny dat upravují především u vyšších vrstev, které zprostředkovávají detailnější informace o vstupu [34]. V této podkapitole jsou popsány některé dostupné předučené CNN, které budou následně v praktické části otestovány.

AlexNet

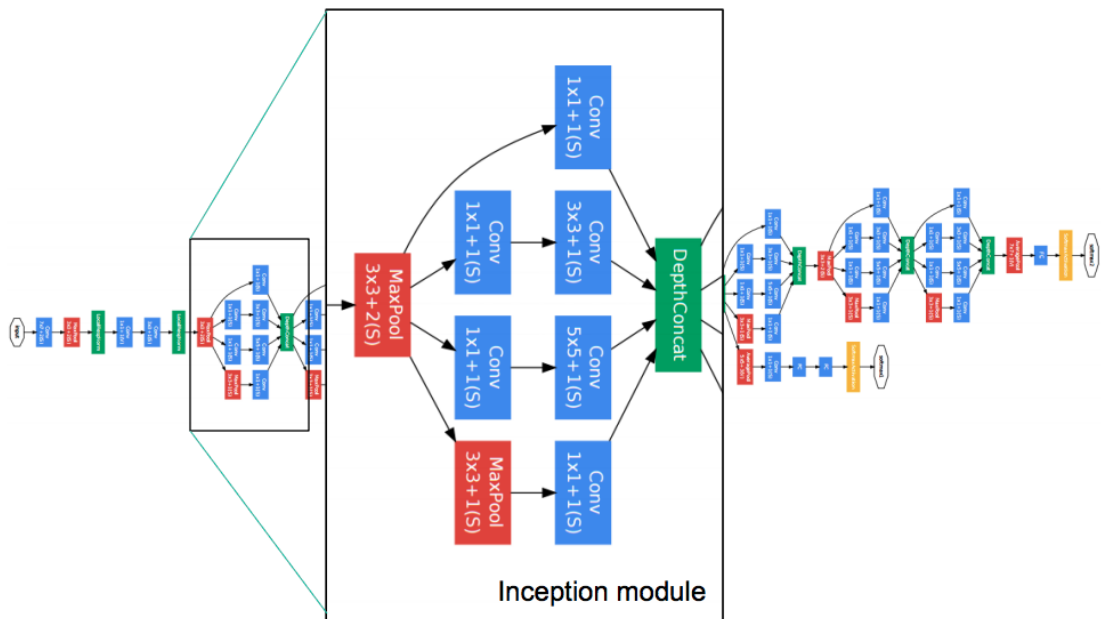
AlexNet je konvoluční síť, navržená Alexem Krizhevskym. V roce 2012 vyhrála na soutěži ILSVRC (ImageNet Large Scale Visual Recognition Competition). Skládá se celkem z 8 vrstev - 5 konvolučních a 3 plně propojených na výstupu. Výstup z každé z těchto osmi vrstev, je řízen nelineární ReLu funkcí. Síť obsahuje 650 000 perceptronů a zpracovává 60 milionů parametrů. První konvoluční vrstva zpracovává vstupní data o rozměrech $227 \times 227 \times 3$ pomocí 96 kernelů o rozměrech $11 \times 11 \times 3$. Výstupní, plně propojená vrstva obsahuje 1000 neuronů pro klasifikaci obrazů z databáze ImageNet [35]. Velikost vstupů každé vrstvy a její další parametry jsou patrné z Obr. 4.5, který popisuje architekturu sítě AlexNet.



Obr. 4.5: Architektura konvoluční neuronové sítě AlexNet [36].

GoogLeNet

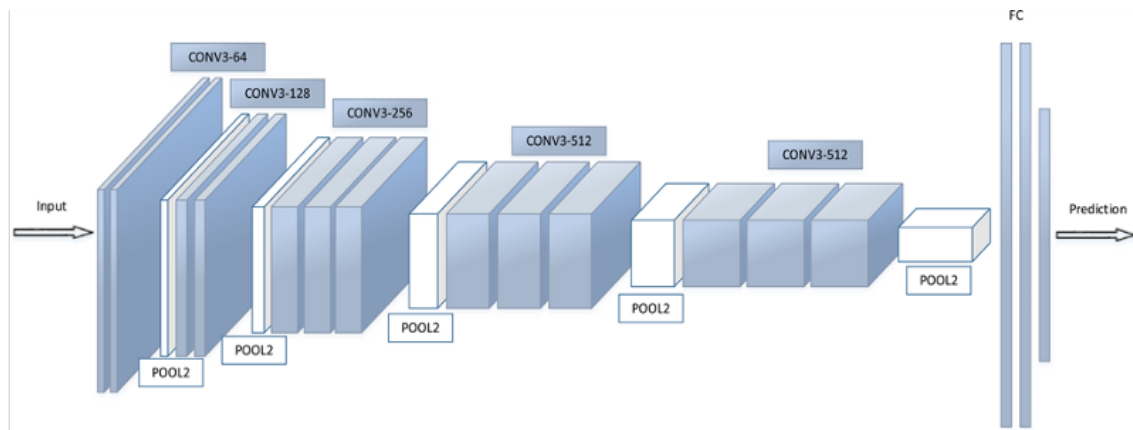
GoogLeNet, neboli Inception v1, je konvoluční síť od společnosti Google. V roce 2014 zítěžila na soutěži ILSVRC. Svou architekturou se výrazně liší od sítí jako jsou AlexNet či VGG. Celkově obsahuje 22 vrstev. V rámci takzvaných Inception modulů jsou do architektury zahrnuty 1×1 konvoluční vrstvy. To vede ke snížení rozměru vstupů, a tedy i výpočetní náročnosti sítě. Z téměř 113 milionů parametrů se díky 1×1 konvoluci stane 5,3 milionů parametrů [37]. Tím je naopak umožněno její prohloubení a rozšíření. Na konci sítě je použito sdružování z globálního maxima namísto plně propojených vrstev [37]. Schéma vrstev sítě GoogLeNet je znázorněné na Obr. 4.6.



Obr. 4.6: Architektura konvoluční neuronové sítě GoogLeNet [38].

VGG-19

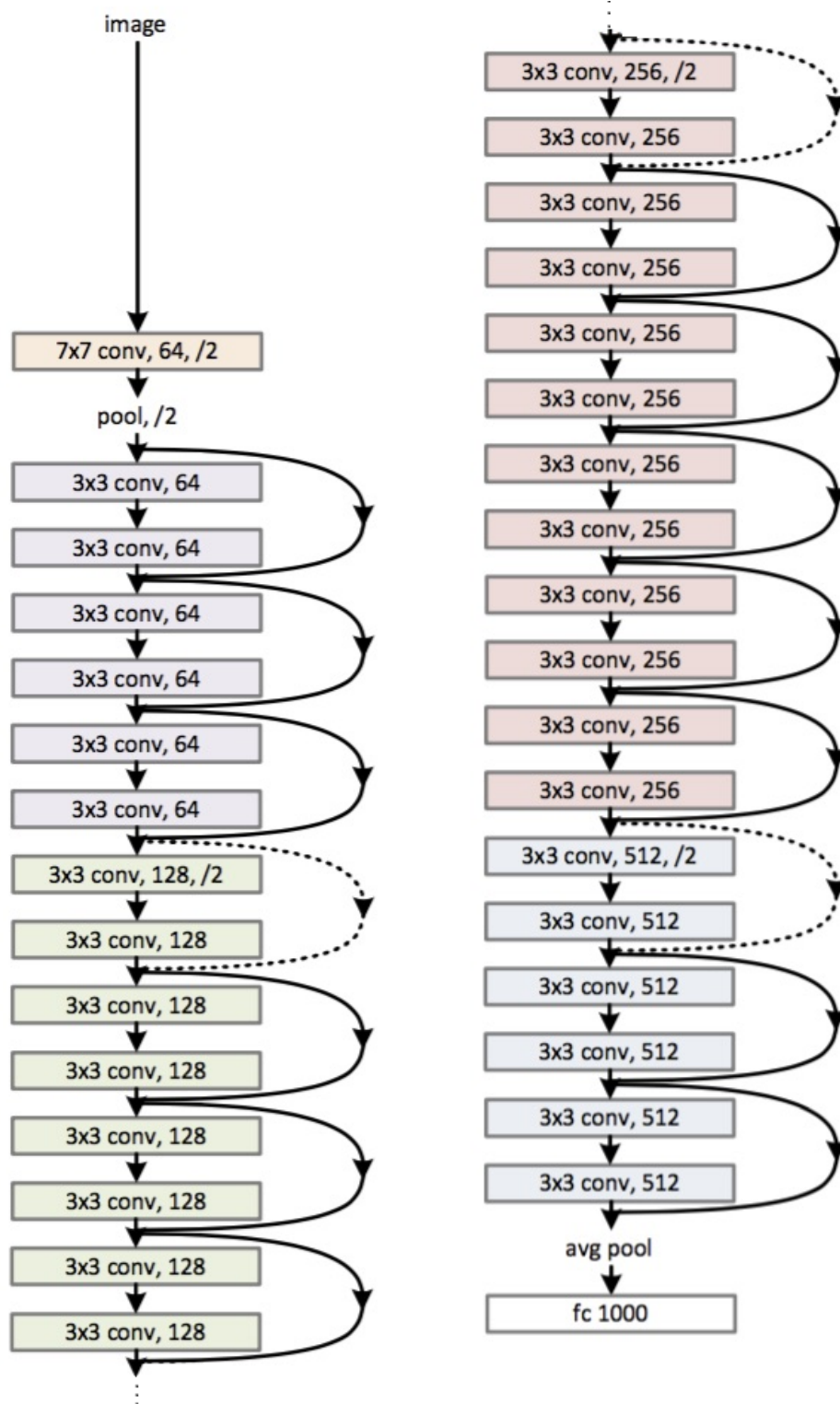
VGG-19 je konvoluční síť, která byla představena v roce 2013 skupinou z Oxfordu, zvanou Visual Geometry Group (odtud VGG). Jedná se o rozšíření AlexNet a obsahuje celkem 19 vrstev. Z nich 16 je konvolučních a 3 plně propojené na výstupu. Jejich výstup je podmíněn aktivační funkcí typu ReLu [39]. Pro snížení počtu parametrů jsou i zde využity max-pooling vrstvy. Na vstupu zpracovává, stejně jako GoogLeNet, obrazy o rozměrech 224x224x3. Síť se nejčastěji používá na obrazovou klasifikaci, případně také jako nástroj pro rozpoznávání obličeje [39]. Existuje také další varianta této sítě, a to VGG-16. Architektura sítě tohoto typu je znázorněna na Obr. 4.7.



Obr. 4.7: Architektura konvoluční neuronové sítě VGG-19 [40].

ResNet

ResNet (Residual Neural Network) je složitá CNN, vyvinutá společností Microsoft. Využívá principu přeskokování vrstev, jejichž váhy se během některých iterací nemění. To vede ke snížení počtu přepočítávaných parametrů, a tedy zjednodušení sítě a urychlení trénování. Tento přístup je analogií k fungování biologických neuronů. ResNet je jedna z nejznámějších CNN a kromě klasifikace obrazů se používá také pro počítačové vidění, detekci objektů a rozpoznávání obličeje [41]. V praktické části této práce budou otestovány dvě různě hluboké předučené sítě tohoto typu, a to ResNet18 a ResNet50, disponující 18 resp. 50 vrstvami. Existují také další varianty, jako například ResNet101 a ResNet152, které však v praktické části nebudou testovány. Architektura sítě typu ResNet je znázorněna na Obr. 4.8.



Obr. 4.8: Architektura konvoluční neuronové sítě typu ResNet [41].

5 Praktická část

Tato kapitola popisuje postup při vypracování praktické části této diplomové práce. Následující podkapitoly představují popis jednotlivých kroků při realizaci programového řešení, které bylo implementováno pomocí softwaru Matlab 2019b.

5.1 Příprava datasetu

Snímky sítnice, které byly pro biometrii použity, vznikly zprůměrováním video-sequenec, pořízené pomocí experimentálního video-oftalmoskopu z Ústavu biomedicínského inženýrství na FEKT VUT v Brně.

Databáze obsahuje snímky 70 osob, přičemž u každé osoby byly nasnímány obě oči, zpravidla nejméně dvakrát. Každý snímek je jednoznačně přidělený odpovídající osobě pomocí identifikačního čísla (ID), a dále také označený informací, zda se jedná o levé či pravé oko. Databáze zároveň obsahuje snímky jak zdravých subjektů, tak také snímky zachycující některé již zmíněné patologie sítnice. Těmto patologiím však není v této diplomové práci věnována pozornost. Schopnost navrženého systému identifikovat identitu, je u nich testována stejně, jako u snímků zdravých očí. K dispozici jsou šedotónové, RGB, a také jejich segmentované varianty, které byly součástí poskytnutých materiálů. Příklad dat, obsažených v databázi je zobrazen na Obr. 5.1.



Obr. 5.1: Příklad dat, zleva: šedotónový obraz, RGB obraz a binární obraz segmentovaného cévního stromu.

Rozdělení vstupních dat

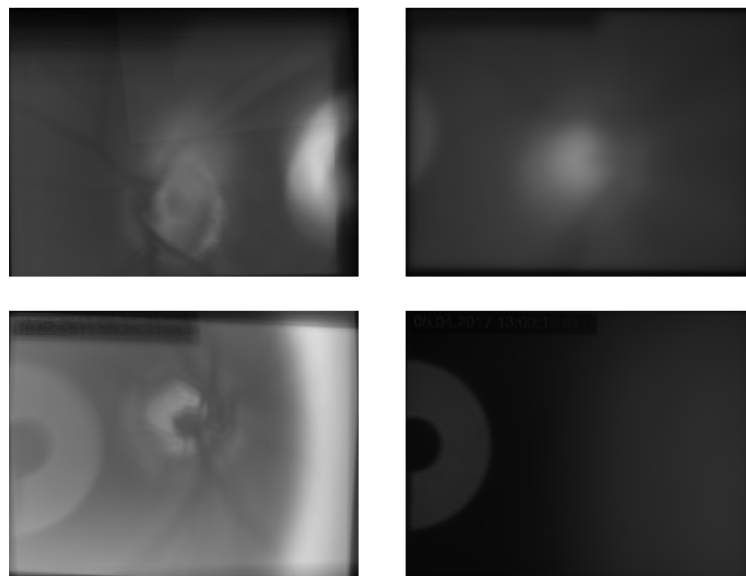
V procesu učení sítě je důležité zvolit vhodný poměr rozdělení dat na trénovací a testovací množinu, nejčastěji v poměru 80 % a 20 %. Tyto dvě podmnožiny by se měly volit jako disjunktní, tedy neobsahující žádný společný prvek [31]. Trénovací množina je dále opět rozdělena v poměru 80 % a 20 %, s tím že 80 % tvoří skutečně trénovací snímky, na kterých se síť učí rozpoznávat příznaky pro klasifikaci.

Zbýlých 20 % je označováno jako validační množina a slouží k průběžnému výpočtu chyby generalizace, a tedy postupnému ověřování průběhu učení sítě. Vhodným rozdělením dat lze navíc předcházet přeučení NN. To by mohlo nastat v případě, že je síť příliš adaptovaná na vyhledávání stejných detailů, jaké se nacházely v učební množině dat. Postrádá tím tedy schopnost pracovat obecněji na nově předložených vstupech [31].

Všechny snímky byly původně součástí jedné množiny. Z té se nejprve manuálně vybraly snímky testovací množiny tak, aby obsahovala od každé kategorie jeden, co možná nejvíce reprezentativní snímek. Rozdělení zbylých dat na trénovací a validační množinu bylo provedeno pomocí algoritmu, který v rámci každého labelu rozděloval data náhodně v poměru 8:2. Tímto způsobem bylo zajištěno, že každá ze tří podmnožin obsahuje z každé kategorie alespoň jeden snímek.

Data musí být navíc uspořádána tak, aby z nich NN dokázala vyčíst všechny potřebné informace. Snímky jednotlivých očí byly proto rozděleny do složek, které nesou název dle přiděleného ID s informací, zda se jedná o pravé (R), či levé (L) oko. Neuronová síť vidí, že všem snímkům stejného oka je přiřazen stejný label (název složky, tedy například 20-L) a díky tomu je schopna si během validace ověřit, zda klasifikuje předložená data do správné kategorie.

Dataset také obsahoval určité množství nekvalitních, až nepoužitelných snímků s výraznými artefakty, které by úroveň klasifikace pouze zhoršovaly. Takovéto snímky byly z datasetu odstraněny, neboť kvalitní databáze je základem úspěšnosti klasifikace pomocí CNN. Některé z artefaktů, které se v databázi nacházely, jsou znázorněny na Obr. 5.2.

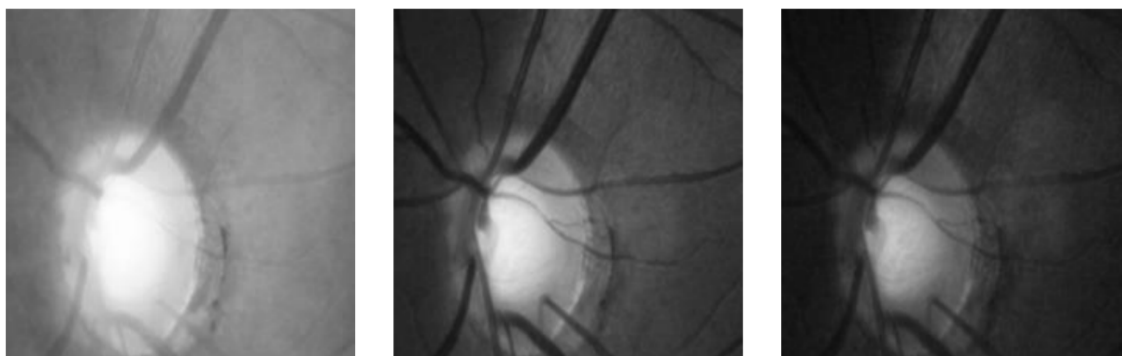


Obr. 5.2: Příklad snímků, které byly z databáze odstraněny.

Augmentace dat pro trénovací množinu

Nástrojem pro úpravu učební sady je tzv. augmetnace dat. Jejím principem je zahrnout do učební množiny kopie některých učebních dat, která jsou modifikována jistou transformací. Často se volí transformace v podobě translace, rotace, zrcadlové převrácení, či jiné nelineární geometrické zkreslení. Tento proces zvýší robustnost navržené sítě a zabrání přeučení z důvodu nedostatku učebních dat [31].

Neuronová síť musí na vstupu obsahovat data jednotného rozměru. To platí i o počtu kanálů, nelze tedy použít zároveň šedotónové a tříkanálové RGB snímky. Aby bylo pro trénování využito maximum dostupných dat, byly všechny RGB snímky rozloženy na šedotónové reprezentace jednotlivých kanálů a jimi následně rozšířena původní databáze šedotónových snímků trénovací množiny. Nejméně reprezentativní byl v případě červeně zbarvené sítnice právě kanál R, který byl ve většině případů přexponovaný. Vybraly se tedy pouze takové snímky, kde bylo cévní řečiště dobře viditelné. Jednotlivé kanály RGB obrazu jsou znázorněny na Obr. 5.3.

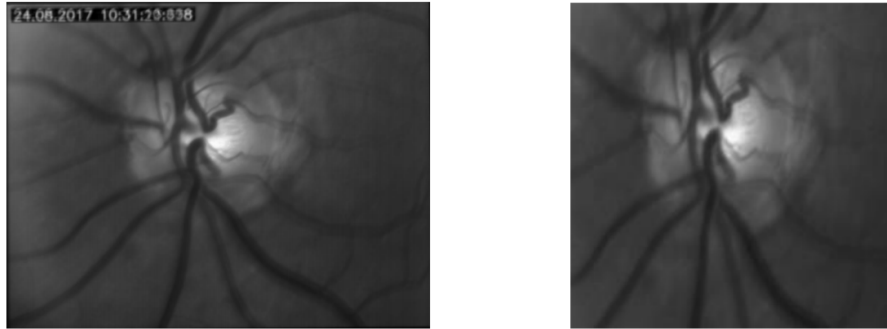


Obr. 5.3: RGB snímek, rozložený šedotónové reprezentace barevných kanálů. Zleva: červený, zelený a modrý kanál.

Oříznutí a úprava velikosti vstupních dat

Poskytnutá data, zaznamenaná video-ofthalmoskopem, obsahovala v levém horním rohu signifikantní informace o čase pořízení snímku. Při učení by se neuronová síť mohla zaměřovat právě na tyto detaily v obraze. Cílem práce je otestovat možnosti biometrie ze struktury cévního stromu sítnice. Pro zajištění důvěryhodnosti provedených metod bylo nutné tuto informaci ze snímků odstranit oříznutím.

Dalším krokem byla úprava velikosti dat, neboť všechny předtrénované sítě mají definovaný rozměr vstupu v první vrstvě své architektury. U většiny sítí se jedná o 224x224, popřípadě 227x227 pixelů. Ukázka úpravy snímků je zobrazena na Obr. 5.4.



Obr. 5.4: Zleva: původní snímek, snímek po úpravě oříznutím a změnou velikosti.

5.2 Výběr architektury CNN

Pro klasifikační problémy se v praxi nejčastěji používají právě konvoluční NN, a proto byly zvoleny pro otestování i v praktické části této diplomové práce. K řešení klasifikační úlohy pro biometrii poslouží metoda Transfer Learning s využitím dostupných předučných modelů CNN. Pro otestování jsem zvolila několik nejznámějších sítí, aby bylo možné z výsledků získat vhodné srovnání jejich možností pro využití v praxi. Podrobnosti o těchto sítích a jejich architektury jsou popsány v Kap. 4.3. Mezi zvolené sítě patří:

- AlexNet
- GoogLeNet
- VGG-19
- ResNet18
- ResNet50

Jak již bylo zmíněno, princip metody Transfer Learning spočívá v odstranění plně propojených vrstev na výstupu sítě. Na jejich místo se napojí nové, plně propojené vrstvy, které obsahují námi požadovaný počet perceptronů, odpovídající počtu klasifikačních tříd.

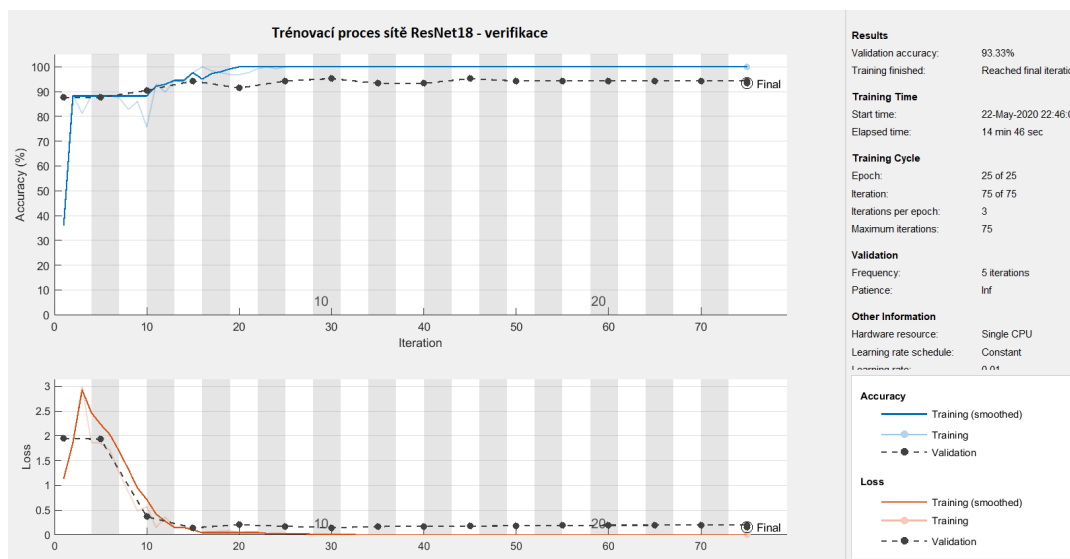
Pro tuto práci byly u všech zmíněných sítí otestovány dva biometrické přístupy, a to verifikace a identifikace. Jako optimalizační algoritmus pro učení byl zvolen stochastický gradientní sestup. Krok pro učení vah byl volen individuálně, dle potřeb jednotlivých sítí. Nejčastěji v rozmezí 0,01 - 0,001, přičemž bylo nutné najít kompromis mezi rychlostí učení a konvergencí k řešení. Trénování bylo ukončeno po dosažení definovaného počtu epoch, s tím že každá síť jich pro dostatečné naučení potřebovala odlišný počet. Frekvence k-násobné validace byla nastavena na hodnotu 5. Na začátku každé trénovací epochy byla data algoritmem náhodně promíchána, aby se zabránilo přeučení.

5.3 Transfer learning na zvolených CNN

Verifikace

Během verifikace má neuronová síť za úkol ověřit předpokládanou identitu, a tu potvrdit či vyvrátit. Při návrhu sítě byla tedy převzata některá z dostupných architektur, včetně zapamatovaných vah a došlo k úpravě jejích vrstev. Konkrétně se jedná o výstupní vrstvu, která původně rozdělovala obrazy na vstupu do 1000 kategorií, obsažených v databázi ImageNet. Původních 1000 perceptronů klasifikační vrstvy se nahradilo pouhými dvěma perceptrony, odpovídající dvěma kategoriím, potřebným pro úlohu verifikace. Tyto kategorie byly nazvány "POVOLENO" a "ZAMÍTNUTO", aby navržená síť představovala jakýsi pomyslný přístupový systém.

Povolený přístup má pouze 9 lidí (identit), tedy celkem 18 očí. Zbýlých 61 identit nemá přístup povolený a na takové vstupy by síť měla reagovat v ideálním případě zařazením do kategorie "ZAMÍTNUTO". Rozdělení do skupiny bylo víceméně náhodné, avšak mělo by simulovat skutečnou situaci zabezpečovacího systému, kdy má přístup obvykle malé procento identit. U takového systému je důležitá především specifická, z hlediska hodnocení biometrického systému proto požadujeme co nejnižší míru chybného přijetí (FAR). Příklad trénovacího procesu je znázorněn na Obr. 5.5.



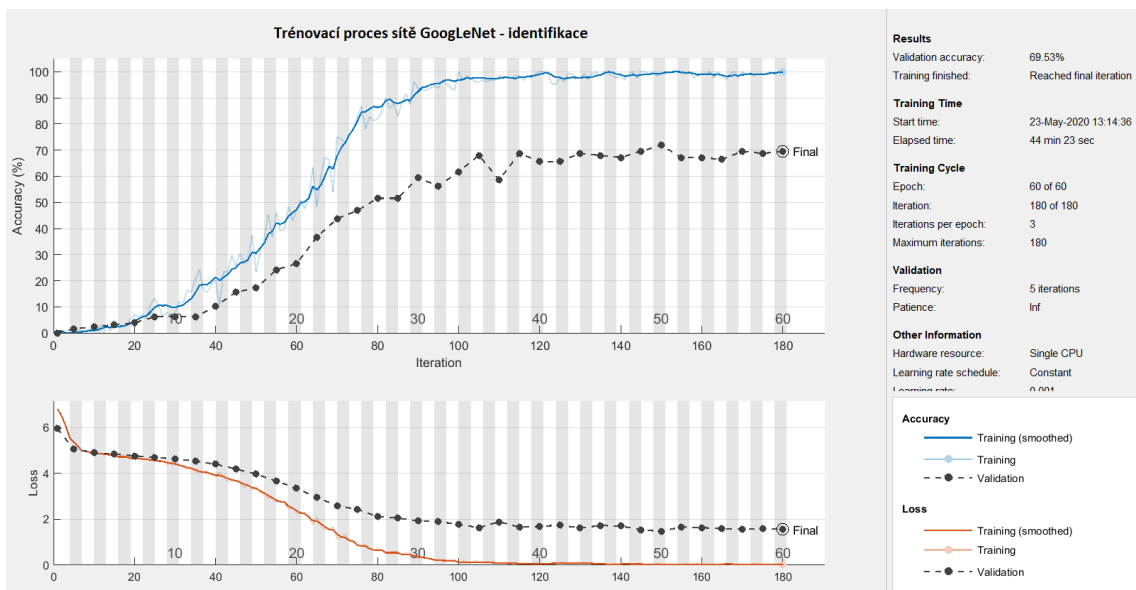
Obr. 5.5: Průběh trénování sítě ResNet18 v režimu verifikace.

Identifikace

Proces identifikace je náročnější, neboť síť má za úkol přiřadit obraz na vstupu konkrétní identitě. Vzhledem k tomu, že sítnice obou očí jsou jiné a nemají žádné společné prvky, nemůžeme předpokládat, že by byla síť schopna na dobré úrovni klasifikovat obě oči pod jednu identitu. Proto síť zvlášť klasifikuje levé a pravé oko stejné osoby, jako dvě různé kategorie.

Postup návrhu sítě byl obdobný jako u verifikace, liší se pouze počtem perceptronů v plně propojených vrstvách sítě. Tento počet je 140, tzn. dvě kategorie pro každou ze sedmdesáti identit. U takového systému jsme schopni vyhodnotit pouze celkovou procentuální přesnost sítě. Ta je určena vzorcem 5.1 [43]. Příklad průběhu trénování sítě je znázorněn na Obr. 5.6.

$$\text{Přesnost klasifikace} = \frac{\text{Počet správně klasifikovaných}}{\text{Celkový počet testovacích snímků}} * 100 [\%], \quad (5.1)$$



Obr. 5.6: Průběh trénování sítě GoogLeNet v režimu identifikace.

5.4 Modifikace sítě ResNet50

V průběhu práce na algoritmu pro Transfer learning se několikrát nabídlo použití Support Vector Machine. Ze zvědavosti jsem se rozhodla ověřit, jak dobře jsou jednotlivé vrstvy CNN schopné extrahovat příznaky pro klasifikaci ze vstupních dat.

Nejdříve byl zvolen předučení model sítě. Tato síť má své váhy pro extrakci příznaků z obrazových dat již přizpůsobené na databázi ImageNet. Na vstup předučení modelu jsem přivedla všechny snímky, které jsem měla k dispozici, ve formě dvou množin. Jednalo se o stejná data, jaká byla použita pro Transfer learning, tedy rozdělená v poměru 8:2 na trénovací a testovací množinu.

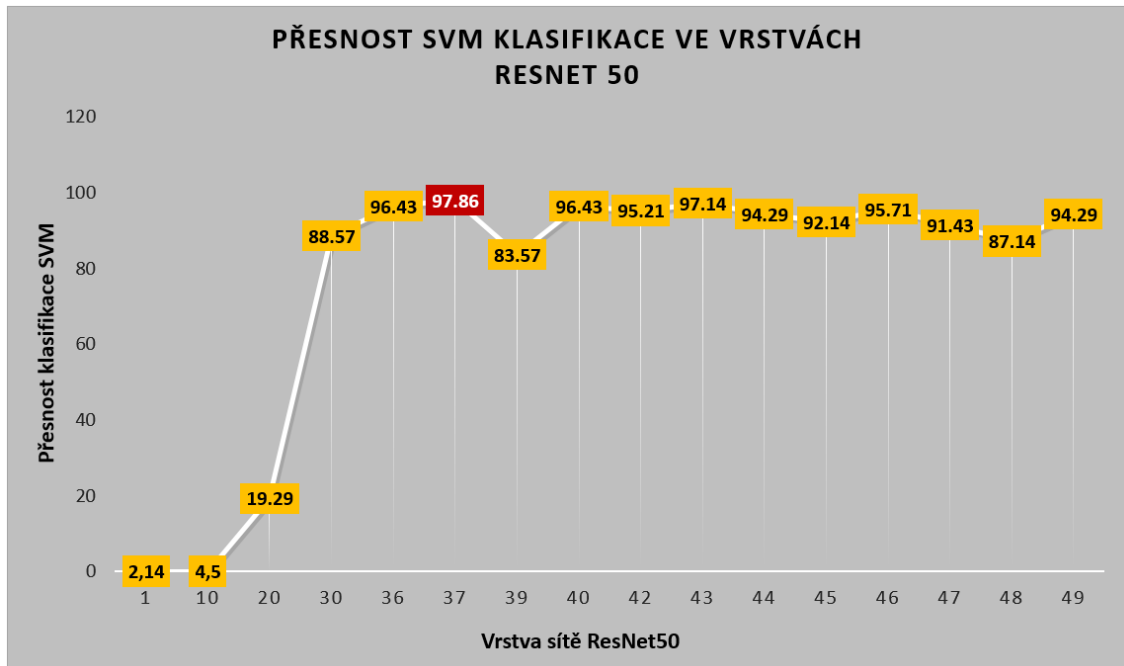
Snímky projdou architekturou daného modelu a jsou konvolucí postupně transformovány. Počet pixelů obrazu se při průchodu architekturou sítě zmenšuje, zatímco roste jeho hloubka. V této formě získáváme reprezentace příznaků, které jsou v každé vrstvě odlišné. V prvních vrstvách sítě jsou informace z obrazu spíše obecné a s postupem do hlubky sítě získáváme více detailů.

Z předučení modelu sítě získám na výstupu obraz, transformovaný operacemi po průchodu jednotlivými konvolučními vrstvami. Následně je možné vyčítat příznaky z různé úrovně sítě, a tedy v různé fázi extrakce otestovat SVM klasifikaci.

Support vector machine (SVM) je klasifikátor, který pro určení oddělovací nadrovin využívá krajní body kategorií v příznakovém prostoru (viz Kap. 3.5). SVM použije trénovací množinu k výpočtu křivky, která vhodně odděluje data v ní obsažená. Touto křivkou je následně proložen příznakový prostor dat z testovací množiny a získán výsledek klasifikace.

V režimu identifikace se pokoušíme klasifikovat více než do dvou tříd, proto byl použit tzv. multiclass klasifikátor. Tento typ SVM umožňuje klasifikaci do více než dvou tříd. Principem je nalézt vhodnou kombinaci vícero lineárních binárních SVM. Tento klasifikátor nebyl trénován žádnými speciálními parametry, taktéž nebyla použita křížová validace. Cílem nebylo SVM co nejlépe vyladit pro klasifikaci, avšak pouze použít jako nástroj pro získání představy o rozložení příznaků v různé fázi průchodu CNN.

Nejhlubší použitou sítí v této práci je ResNet50. Rozhodla jsem se otestovat úroveň kvality příznaků v jednotlivých vrstvách této sítě, abych ověřila, zda je takto hluboká síť pro klasifikaci sítě nutná. Pokud by se prokázalo, že není, může být lepší volbou některá z jednodušších sítí, především kvůli počtu parametrů a času, potřebného pro trénování. Graf, znázorňující přesnost SVM klasifikace s využitím příznaků z jednotlivých vrstev, je znázorněn na Obr. 5.7.



Obr. 5.7: Přesnost SVM klasifikace z příznaků v dané vrstvě ResNet50.

Z grafu na Obr 5.7 je patrné, že nejlépe klasifikoval SVM obrazy vystupující z 37. vrstvy. Toto zjištění se pokusím prokázat modifikací ResNet50 na vlastní síť, kterou neoficiálně nazvu podle počtu vrstev - ResNet37. Architektura ResNet50 bude upravena obdobně, jako během metody Transfer learning, avšak se zásahem do celkové hloubky sítě. Všechny vrstvy za 37. vstvou budou odstraněny a na její výstup se připojí plně propojené vrstvy pro klasifikaci. Plně propojené vrstvy budou opět obsahovat dva perceptrony pro úlohu verifikace a 140 perceptronů pro úlohu identifikace. V rámci experimentu bude testováno, zda síť ResNet s hloubkou 37 vrstev dosáhne lepší přesnosti klasifikace, nežli ResNet18, ResNet50 a ostatní architektury.

6 Výsledky diplomové práce

V této kapitole jsou shrnuty nejlepší dosažené výsledky testování jednotlivých sítí, pro obě biometrické úlohy. Během trénování NN je často nejdůležitější a nejnáročnější příprava dat. V průběhu ladění sítí byl proto dataset experimentálně upravován tak, aby síť předvedla nejlepší možný výsledek. Stejně tak změny parametrů sítí probíhaly experimentálně, neboť každá architektura na ně reagovala jinak.

Pro učení se neprokázaly jako přínosné segmentované snímky, jejich zařazení do trénovací množiny naopak schopnost klasifikace značně snižovalo, a to i v případě, že byly z binárního snímku převedeny na šedotónový (nutné kvůli jednotnosti formátu dat). Z tohoto důvodu nebyly segmentované snímky zahrnuty v učební množině, ani použity pro testování.

Pro trénovací množinu byla použita pouze augmentace s využitím různých kanálů RGB snímků, neboť geometrické transformace se pro augmentaci v případě sítnice neosvědčily. Důvodem může být fakt, že obecně jsou si snímky sítnice velmi podobné. Například v databázi ImageNet nemá rotace snímku květiny či auta takový vliv na výsledek klasifikace, naopak zajistí lepší naučení sítě. Avšak rotací či vertikálním převrácením snímku sítnice, může snadno dojít k záměně s jinou sítnicí. Všechny snímky totiž tvoří jistá pavučina cév se středem v optickém disku. Horizontální převrácení navíc změní polohu optického disku, která může být zásadní pro rozlišení levého a pravého oka.

6.1 Výsledky Transfer learning na zvolených CNN

Trénovací procesy všech sítí jsou zobrazeny v Příloze A. Lze si všimnout, že výsledná klasifikace na testovacích snímcích je úspěšnější, nežli přesnost konečné validace během učení. Tento jev připisuji způsobu, jakým byly tyto množiny dat vytvořeny. Pro testovací množinu byly manuálně vybrány poměrně kvalitní snímky dané sítnice. Validační množina, vzniklá náhodným výběrem algoritmu, tedy pravděpodobně obsahuje snímky s nižší kvalitou. V každém případě dobrá performance na nových datech ukazuje, že modely nejsou přeučené na trénovací množině.

V tabulkách 6.1 a 6.2 jsou přehledně srovnány výsledky jednotlivých sítí. Pro úlohu verifikace se nejvíce osvědčily sítě GoogLeNet, ResNet18 a ResNet37. Tyto správně klasifikovaly 139 z celkových 140 snímků, dosáhly tak přesnosti klasifikace 99,29 %. U všech tří se zároveň podařilo dosáhnout 100% specificity. Ta zajišťuje, že přístupový systém nepovolí vstup nepovolané osobě, což je u takového systému žádoucí.

	VÝSLEDKY TESTOVÁNÍ VERIFIKACE					
Název	AlexNet	GoogLeNet	VGG-19	Resnet18	ResNet37	Resnet50
Celkový počet identit	140					
Počet správně klasifikovaných	138	139	134	139	139	138
Počet nesprávně klasifikovaných	2	1	6	1	1	2
Přesnost klasifikace [%]	98,57	99,29	95,71	99,29	99,29	98,57
Míra chybného přijetí [%]	0	0	0	0	0	0
Specifická [%]	100	100	100	100	100	100
Míra chybného odmítnutí [%]	1,61	0,8	4,9	0,8	0,8	1,61
Senzitivita [%]	88,89	94,44	66,67	94,44	94,44	88,89

Obr. 6.1: Výsledky testovaných sítí v režimu verifikace.

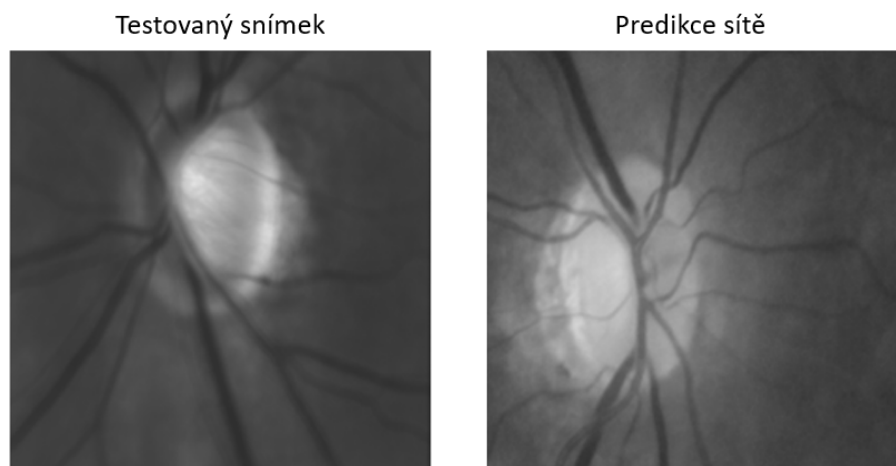
Přestože identifikace je mnohem náročnější klasifikační úloha, i zde se podařilo dosáhnout velmi vysoké přesnosti. Vítězná síť ResNet18 správně přiřadila 138 ze 140 identit, klasifikovala tedy s přesností 98,57 %.

Vlastní síť ResNet37 obstála v úloze verifikace jako jedna z nejlepších. Naopak v režimu identifikace dosáhla nejnižší přesnosti ze všech testovaných architektur. Jelikož ResNet18 i ResNet50 mají při identifikaci dobré výsledky, je pravděpodobné, že to není způsobeno nevhodností architektury. Nízká přesnost klasifikace může být způsobena tím, že se nepodařilo nalézt vhodnou kombinaci parametrů pro učení.

VÝSLEDKY TESTOVÁNÍ IDENTIFIKACE						
Název	AlexNet	GoogLeNet	VGG-19	Resnet18	ResNet37	Resnet50
Celkový počet identit	140					
Počet správně klasifikovaných	134	134	128	138	122	131
Počet nesprávně klasifikovaných	6	6	12	2	18	9
Přesnost klasifikace [%]	95,71	95,71	91,43	98,57	87,14	93,57

Obr. 6.2: Výsledky testovaných sítí v režimu identifikace.

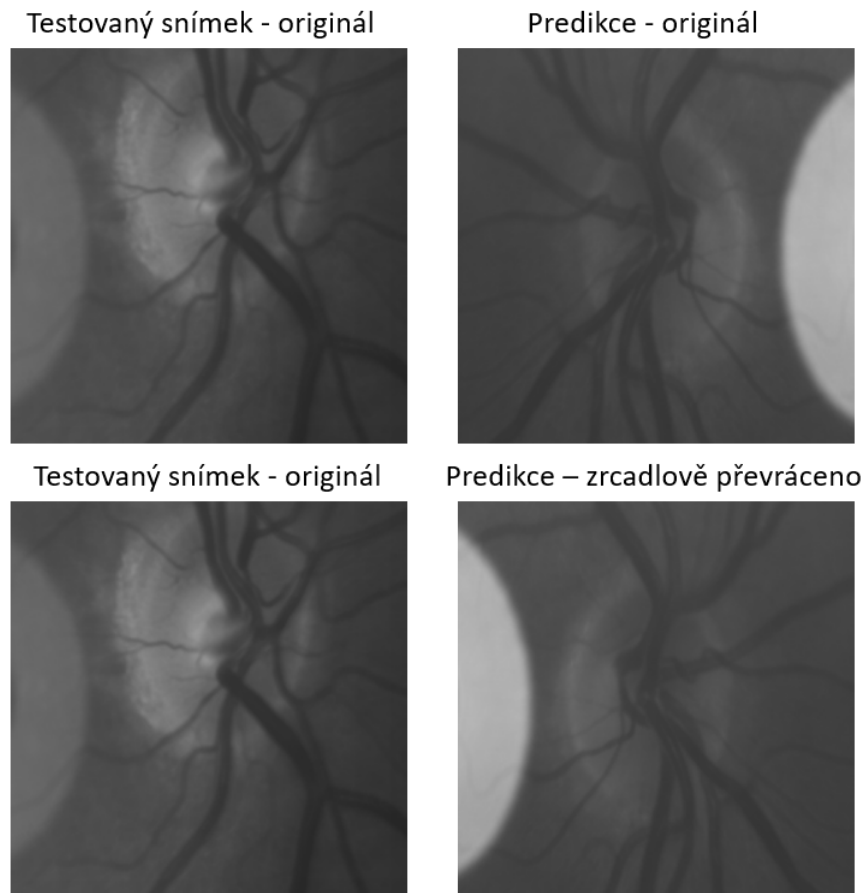
Na Obr. 6.3 vidíme jediný společný snímek, který se žádná z testovaných sítí v režimu verifikace nepodařilo rozpoznat. Pro zajímavost je znázorněn spolu se snímkem, za který jej sítě nesprávně zaměňovaly. Přestože si na první pohled tyto cévní stromy nejsou příliš podobné, CNN extrahuje abstraktní příznaky, které si pravděpodobně podobné jsou.



Obr. 6.3: Příklad špatně klasifikovaného snímku sítě GoogLeNet, ResNet18 a ResNet37 v režimu verifikace.

Sít ResNet18, která dosáhla nejlepší přesnosti klasifikace v režimu identifikace, zaměnila dva snímky. První z nich je totožný, jako u verifikačních sítí a druhý je znázorněn na Obr. 6.4, spolu s horizontálně převrácenou variantou predikce. V tomto

případě je patrné, že snímky jsou si velice podobné, nejen co se struktury cév týče, ale také vzhledem k přítomnému artefaktu.



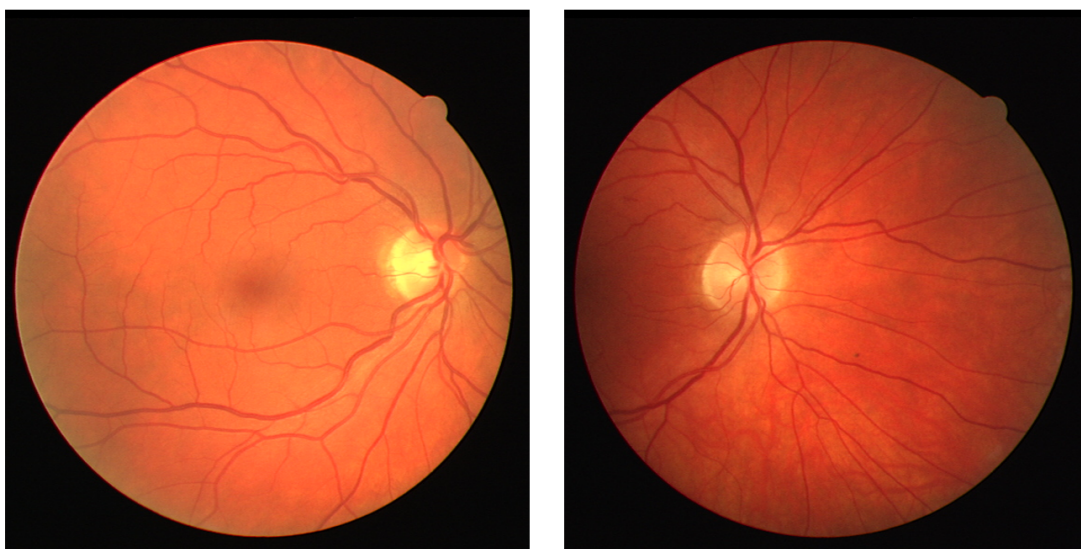
Obr. 6.4: Příklad špatně klasifikovaného snímku sítí ResNet18 v režimu identifikace.

I přes nízké rozlišení snímků, na kterých byly sítě testovány, se podařilo dosáhnout velice dobrých výsledků klasifikace. Pokud by takovýto biometrický systém fungoval v běžné praxi, bylo by hlavním pilířem úspěšnosti systému, zajistit dostatečně obsáhlou databázi kvalitních snímků sítnice. Bylo by samozřejmě nutné zvážit také míru zásahu do soukromí a možnost zneužití osobních údajů.

6.2 Možná rozšíření práce

V průběhu praktické části se postupně nabízely další možnosti, jakými by se mohlo v budoucnu na tuto práci navázat. Jednou z nich je porovnání úspěšnosti algoritmů s kvalitnějšími daty. Existují veřejně dostupné databáze, obsahující snímky sítnice s vysokým rozlišením, pocházející z fundus kamery. Takové porovnání by podpořilo interpretaci výsledků na námi testovaných datech. Snímky ve zmíněných databázích však nejsou nijak označené, bylo by tedy nutné data manuálně rozdělit do kategorií dle vlastního uvážení, popřípadě využít možnosti učení bez učitele. Vzhledem k podobnosti všech fundus snímků by ovšem klasifikace bez učitele pravděpodobně nedosáhla příliš vysoké přesnosti.

Dalším způsobem pro klasifikaci snímků by mohlo být použití takzvaných Siamských sítí. Jedná se o dvě paralelní CNN, které sdílí své parametry. Na jejichž vstupy jsou předložena obrazová data a sítě porovnávají jejich podobnost. U Siamské sítě je však problematické vytvořit vhodnou sadu dat. Pro učení takovéto sítě je nutné na její vstup přivést dvojici snímků (každý snímek vstupuje do jedné z paralelích CNN) a informaci o jejich podobnosti [44]. Množství kombinací pro vytvoření dvojic, by v případě 140 očí, které klasifikujeme, bylo s omezenou časovou dotací na vypracování DP příliš zdlouhavé. Z tohoto důvodu nebyla tato metoda v práci uplatněna. Na Obr. 6.5 je příklad fundus snímků z DRIVE databáze, které by pravděpodobně pro biometrii posloužily ještě lépe, nežli snímky použité pro tuto práci.



Obr. 6.5: Příklad snímků sítnice z databáze DRIVE [45].

Závěr

První kapitoly této práce obsahují teoretický podklad, potřebný pro vypracování praktické části diplomové práce. Jejím cílem bude vytvořit identifikační systém, který využívá biometrická data ve formě snímků oční sítnice.

V teoretické části se tato práce zameřuje na obecný princip biometrických systémů, jejich varianty a také způsoby hodnocení kvality těchto systémů. Dále je zde popsána anatomie lidského oka, nutná pro pochopení principů biometrie sítnice. Jsou představeny některé patologické projevy na sítnici, které se na jejích snímcích projeví a také zařízení, která se ke snímání sítnice využívají.

V následující kapitole jsou představeny různé přístupy a metody, které lze v praxi k rozpoznání sítnice použít. Kapitola obsahuje stručný popis metod, které v práci využity nejsou, dále se podrobněji zaměřuje právě na popis metod, které jsou testovány v rámci praktické části. Neuronovým sítím a metodě hlubokého učení je proto věnována celá následující kapitola.

V praktické části práce je podrobně popsán postup, který byl při návrhu programového řešení uplatněn, přes přípravu datasetu, až po realizaci zvolených metod. Veškeré výsledky testování biometrie ze snímků sítnice jsou přehledně shrnuty a okomentovány v poslední kapitole této práce. Dále jsou nabídnuty další možnosti, kterými by se na práci dalo v budoucnu navázat.

V této práci bylo prokázáno, že snímky sítnice mohou být vhodným nástrojem pro biometrii osob, byť se jedná o snímky s nízkým rozlišením. Algoritmus byl testován v režimu ověření identity (verifikace) a v režimu přiřazení identity (identifikace). V obou případech bylo dosaženo velmi přívětivých výsledků.

Nejlepší úroveň klasifikace v případě verifikace dosáhly sítě GoogLeNet, ResNet18, také vlastní síť ResNet37, které ze 140 snímků správně klasifikovaly 139. Dosáhly tak přesnosti 99,29 % a zároveň 100% specificity. V režimu identifikace nejlépe obstála síť ResNet18 se 138 správně přiřazenými snímky, tedy přesností 98,57 %. Experimentálně modifikovanou síť ResNet37 se nepodařilo natrénovat pro dostatečnou přesnost klasifikace v režimu identifikace. V režimu verifikace však obstála jako jedna ze tří nejlepších architektur.

Literatura

- [1] PANKANTI, Sharath, Ruud M. BOLLE a Anil JAIN. *Biometrics: The future of identification [Guest Eeditors' Introduction]*. Computer [online]. 2000, 33(2), 46-49 [cit. 2020-01-02]. DOI: 10.1109/2.820038. ISSN 0018-9162. Dostupné z: <http://ieeexplore.ieee.org/document/820038>
- [2] PHILLIPS, P.J., A. MARTIN, C.L. WILSON a M. PRZYBOCKI. *An introduction evaluating biometric systems*. Computer [online]. 33(2), 56-63 [cit. 2020-01-02]. DOI: 10.1109/2.820040. ISSN 00189162. Dostupné z: <http://ieeexplore.ieee.org/document/820040/>
- [3] MOHAMED, Soltane. *Multi-modal Biometric Authentications: Concept Issues and Applications Strategies*. 2012, 3. 23-60. Dostupné z: <http://www.sersc.org/journals/IJAST/vol48/3.pdf>.
- [4] WORTHMAN, Ernest. *Biometrics for the IoT*. Semiconductor Engineering, 2015. Dostupné z: <https://semiengineering.com/biometrics-for-the-iot/>
- [5] DYLEVSKÝ, Ivan, JEŽEK, Petr. *Základy funkční anatomie člověka*. 2012, VOŠ Palestra. Dostupné z: <http://vos.palestra.cz/skripta/anatomie/15a2a5.htm>
- [6] DRAHANSKÝ, M., ORSÁG, F., DVOŘÁK, R., HÁJEK, J., VÁŇA et.al. *Biometrie*. Brno: Computer Press, s.r. o, 2011. 294 p. ISBN: 978-80-254-8979-6.
- [7] Eye Physicians and Surgeons, P.A. *Serving Patients in Wilmington with Retinal Diseases: Retinal Detachment, Vascular Disease, Floaters, Flashers*. Dostupné z: <https://www.eyephysicians.com/retinal-diseases>
- [8] VENCKUS, Toni, CRA, VISLISEL, Jesse, MD. 2014, Ophthalmic Atlas Images, The University of Iowa. Dostupné z: <http://eyerounds.org>
- [9] METHA, Sonia , MD. *Overview of Retinal Disorders*. Vitreoretinal Diseases and Surgery Service, Wills Eye Hospital, Sidney Kimmel Medical College at Thomas Jefferson University. Dostupné z: <https://www.merckmanuals.com/home/eye-disorders/retinal-disorders/overview-of-retinal-disorders>
- [10] DOYLE, S., CAMPBELL, M., OZAKI, E. et al. *NLRP3 has a protective role in age-related macular degeneration through the induction of IL-18 by drusen components*. 2012, Nat Med 18, 791–798 Dostupné z: <https://www.nature.com/articles/nm.2717>

- [11] SAINE, P. J., TYLER, M. E. *Fundus Photography Overview*. Ophthalmic Photography: Retinal Photography, Angiography, and Electronic Imaging, 2nd Edition. Butterworth-Heinemann Medical, ISBN: 0750673729 Dostupné z: <https://www.opsweb.org/page/fundusphotography>
- [12] TORNOW, Ralf-Peter, ODSTRČILÍK, Jan, KOLAR, Radim. *Time-resolved quantitative inter-eye comparison of cardiac cycle-induced blood volume changes in the human retina*. 2018, Biomedical optics express 9, 6237-6254. Dostupné z: <https://www.osapublishing.org/boe/abstract.cfm?uri=boe-9-12-6237>
- [13] SADIKOGLU, F., UZELALTINBULAT, s. *Biometric Retina Identification Based on Neural Network*. Procedia Computer Science [online]. 2016, 102, 26-33 [cit. 2020-01-02]. DOI: 10.1016/j.procs.2016.09.365. ISSN 18770509. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1877050916325455>
- [14] RAK, R. *Biometrie a identita člověka ve forezních a komerčních aplikacích*. 1. vydání, Praha: Grada, 2008, 631 s. ISBN 978-80-247-2365-5.
- [15] ODSTRČILÍK, J., JAN, J., GAZAREK, J., KOLÁŘ, R. *Improvement of Vessel Segmentation by Matched Filtering in Colour Retinal Images*. 2009, 4 s., DOI: 10.1007/978-3-642-03891-4₈₇.
- [16] IMRAN, A., LI, J., PEI, Y., YANG, J., WANG, Q. *Comparative Analysis of Vessel Segmentation Techniques in Retinal Images*. IEEE Access, vol. 7, pp. 114862-114887, 2019. DOI: 10.1109/ACCESS.2019.2935912 Dostupné z: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=8804190isnumber=8600701>
- [17] B. MAZUMDAR, Jarina. *Retina based biometric authentication system: A review*. 2018, International Journal of Advanced Research in Computer Science. 9. 711-718. DOI: 10.26483/ijarcs.v9i1.5322. ISSN 09765697. Dostupné z: <http://www.ijarcs.info/index.php/Ijarcs/article/view/5322/4583>.
- [18] BEVILACQUA V., et al. *Retinal Fundus Biometric Analysis for Personal Identifications*. 2008, Lecture Notes in Computer Science. ISBN 978-3-540-85983-3. Dostupné z: http://link.springer.com/10.1007/978-3-540-85984-0_147
- [19] ORTEGA, M, et al. *Biometric authentication using digital retinal images*. International Conference on Applied Computers Science, 2006, 6 p. Dostupné z: <https://www.researchgate.net/publication/238087568>
- [20] ORTEGA, M, et al. *Retinal verification using a feature points-based biometric pattern*. EURASIP J. Adv. Signal Process, 2009, 13 p.

DOI: 10.1155/2009/235746. ISSN 1687-6180. Dostupné z: <https://asp-urasipjournals.springeropen.com/articles/10.1155/2009/235746>

- [21] KARÁSEK, J. *Citlivost metod pro měření podobnosti kvantitativních proměnných*. 2012, roč. 12 [cit. 2014-10-25]. ISSN 1214-9675. Dostupné z: <http://access.feld.cvut.cz/view.php?cislocclanku=2012090003>
- [22] AKRAM, M. U., TARIQ, A., KHAN, S. A. *Retinal recognition: Personal identification using blood vessels*. 2011, International Conference for, pp. 180–184, IEEE. Dostupné z: <https://ieeexplore.ieee.org/document/6148446>
- [23] KAKARWAL, S. N. and R. R. DESHMUKH. *Analysis of Retina Recognition by Correlation and Covariance Matrix*. 2010 3rd International Conference on Emerging Trends in Engineering and Technology. IEEE, 2010, p. 496-499. DOI: 10.1109/ICETET.2010.29.
- [24] HALL, Mark A. *Correlation-based Feature Selection for Machine Learning*. University of Waikato, Department of Computer Science. Dostupné z: <https://www.cs.waikato.ac.nz/mhall/thesis.pdf>
- [25] SARVAIYA, J.N., PATNAIK, S., BOMBAYWALA S. *Image Registration by Template Matching using Normalized Cross-Correlation*. 2009, International Conference on Advances in Computing, Control, and Telecommunication Technologies. DOI: 10.1109/ACT.2009.207. ISBN 978-1-4244-5321-4. Dostupné z: <https://ieeexplore.ieee.org/document/5375779>
- [26] PERWEEN, Nazil, KUMAR, DarshaN, BHARDWAJ, Ishan. *An Overview on Template Matching Methodologies and its Applications*. 2013, International Journal of Research in Computer and Communication Technology, Vol 2, Issue 10, ISSN: 2278-5841 Dostupné z: <https://pdfs.semanticscholar.org/93cf/4a44197e600d4014e234ea6cd3599b962ba4.pdf>
- [27] GUNN, Steve R., et al. *Support vector machines for classification and regression*. ISIS technical report, 1998, 14.1: 5-16. Dostupné z: <http://ce.sharif.ir/courses/85-86/2/ce725/resources/root/LECTURES/SVM.pdf>
- [28] DAVIDSON, P., WAAS, A. *Probabilistic Defect Analysis of Fiber Reinforced Composites using Kriging and Support Vector Machine based Surrogates*. 2018, Composite Structures. DOI: 195. 10.1016/j.compstruct.2018.03.007.

- [29] MALIK, F. *Neural Network Activation Function Types*. 2019, Medium [online] Dostupné z: <https://medium.com/fintechexplained/neural-network-activation-function-types-a85963035196>
- [30] NIELSEN, Michael A. *Neural Networks and Deep learning*. 2015, Determination Press Dostupné z: <http://neuralnetworksanddeeplearning.com/chap6.html>
- [31] YOSHUA BENGIO, IAN J. GOODFELLOW, AARON COURVILLE, *Deep learning*. 2015, Book in preparation for MIT Press Dostupné z: <http://www.deeplearningbook.org/>
- [32] SHARMA V. A., *Understanding Activation Functions in Neural Networks*. 2017, [online] Medium. Dostupné z: <http://www.deeplearningbook.org/>
- [33] ImageNet. 2016, Stanford Vision Lab, Stanford University, Princeton University Dostupné z: <http://image-net.org/about-overview>
- [34] Pan, Sinno Jialin and Qiang Yang. *A Survey on Transfer Learning*. IEEE Transactions on Knowledge and Data Engineering 22 (2010):1345-1359.
- [35] KRIZHEVSKY, A., SUTSKEVER, I, HINTON, G. E. *ImageNet classification with deep convolutional neural networks*. Communications of the ACM. 60 (6): 84–90. DOI:10.1145/3065386. ISSN 0001-0782.
- [36] HAN, X.; ZHONG, Y.; CAO, L.; ZHANG, L. *Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification*. Remote Sens. 2017, 9, 848; <https://doi.org/10.3390/rs9080848>
- [37] SZEGEDY, C. et al. *Going deeper with convolutions*. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594
- [38] XU, J. *An Intuitive Guide to Deep Network Architectures*. 2017, [online] Towards Data Science. Dostupné z: <https://towardsdatascience.com/an-intuitive-guide-to-deep-network-architectures-65fdc477db41>
- [39] SIMONYAN K., ZISSERMAN, A. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2014, University of Oxford
- [40] TAI, D.N., NA, I.S., LEE, G.-S., YANG, H.-J., KIM, S.H. *Tracking by Detection of Multiple Faces using SSD and CNN Features*. 2018, [online] ResearchGate.net.

- [41] FUNG V. *An Overview of ResNet and its Variants*. 2017, [online] Towards Data Science. Dostupné z: <https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035>
- [42] ANAND P. *Residual Network (Resnet)*. 2020, [online] OpenGenus IQ: Learn Computer Science Dostupné z: <https://iq.opengenus.org/resnet/>
- [43] MISHRA A. *Metrics to Evaluate your Machine Learning Algorithm*. 2018, [online] Towards Data Science Dostupné z: <https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234>
- [44] KOCH, G., ZEMEL, R., SALAKHUTDINOV R. *Siamese neural networks for one-shot image recognition*. 2015, ICML deep learning workshop. Vol. 2. Dostupné z: <http://www.cs.toronto.edu/~gkoch/files/msc-thesis.pdf>
- [45] GINNEKEN, B., KERKSTRA, S., MEAKIN, J.. *DRIVE: Digital Retinal Images for Vessel Extraction*. Dostupné z: <https://drive.grand-challenge.org/>

Seznam symbolů, veličin a zkratek

ANN	Umělá neuronová síť - z anglického Artificial Neural Network
CNN	Konvoluční neuronová síť - z anglického Convolutional Neural Network
DP	Diplomová práce
DR	Diabetická retinopatie
EER	Míra vyrovnání chyb - z anglického Equal Error Rate
FAR	Míra chybného přijetí identity - z anglického False Acceptance Rate
FMR	Míra chybné shody - z anglického False Match Rate
FNMR	Míra chybné neshody - z anglického False Non-Match Rate
FRR	Míra chybného odmítnutí identity - z anglického False Rejection Rate
ID	Identifikační číslo
ILSVRC	ImageNet Large Scale Visual Recognition Competition - celosvětová soutěž neuronových sítí
NN	Neuronová síť - z anglického Neural Network
ResNet	Residual Neural Network
RGB	Tříkanálový barevný obraz
SVM	Support Vector Machine
TM	Metoda porovnávání vzorů - z anglického Template Matching
VGG	Konvoluční neuronová síť skupiny Visual Geometry Group

Seznam příloh

A Přílohy	57
A.1 Příložená data a kódy	57
A.2 Obrazové přílohy - přehled trénovacích procesů jednotlivých sítí . . .	58

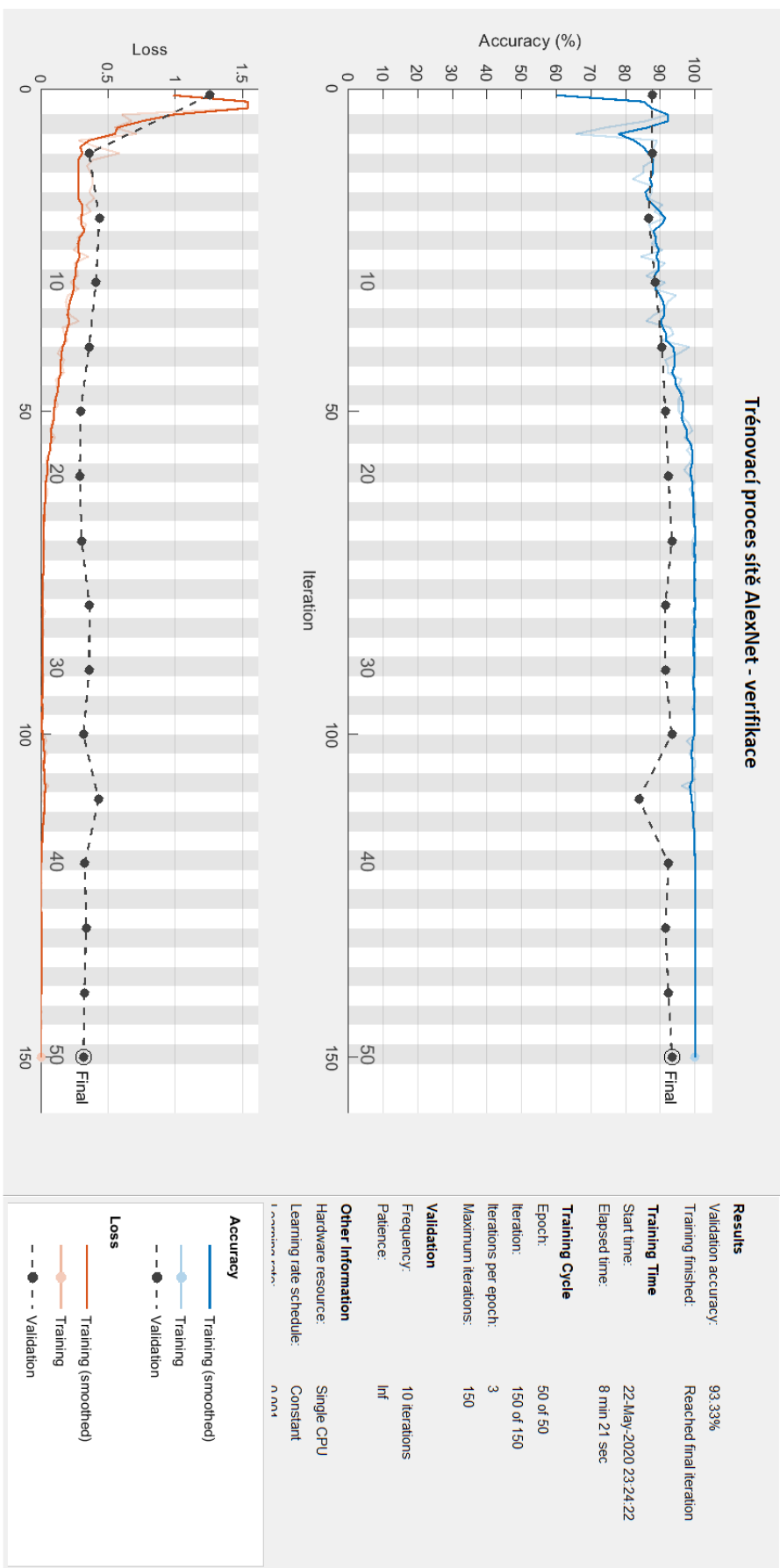
A Přílohy

A.1 Přiložená data a kódy

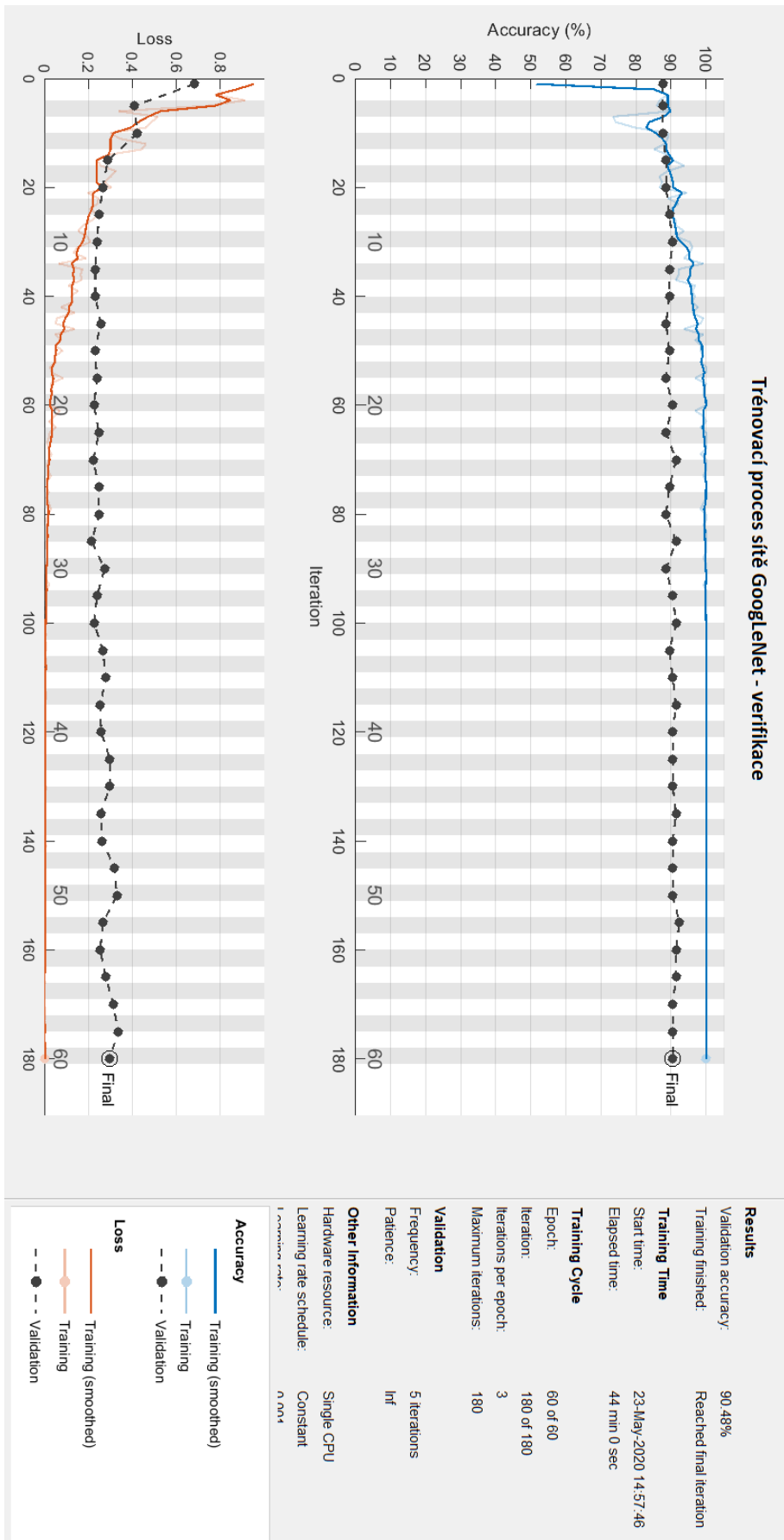
Odevzdané přílohy obsahují zdrojové kódy, vypracované v rámci praktické části DP. Tyto kódy byly vytvořeny v programu Matlab 2019b. Součástí odevzdaných materiálů je také zmenšená verze datasetu, kvůli možnosti spuštění skriptů.

```
/.....kořenový adresář přiložené složky Přílohy Smrčková DP.zip
├── readme.txt
├── data .....dataset pro spuštění přiložených kódů
│   ├── Verifikace
│   │   ├── train ver ..... Trénovací sada pro verifikaci
│   │   └── test ver ..... Testovací sada pro verifikaci
│   └── Identifikace
│       ├── train ID ..... Trénovací sada pro identifikaci
│       └── test ID ..... Testovací sada pro identifikaci
├── alexnet
│   ├── alexnet verifikace.m
│   └── alexnet identifikace.m
├── googlenet
│   ├── googlenet verifikace.m
│   └── googlenet identifikace.m
├── resnet18
│   ├── resnet18 verifikace.m
│   └── resnet18 identifikace.m
├── resnet37
│   ├── resnet37 verifikace.m
│   └── resnet37 identifikace.m
├── resnet50
│   ├── resnet50 verifikace.m
│   └── resnet50 identifikace.m
├── vgg19
│   ├── vgg19 verifikace.m
│   └── vgg19 identifikace.m
└── SVM
    └── svm.m
```

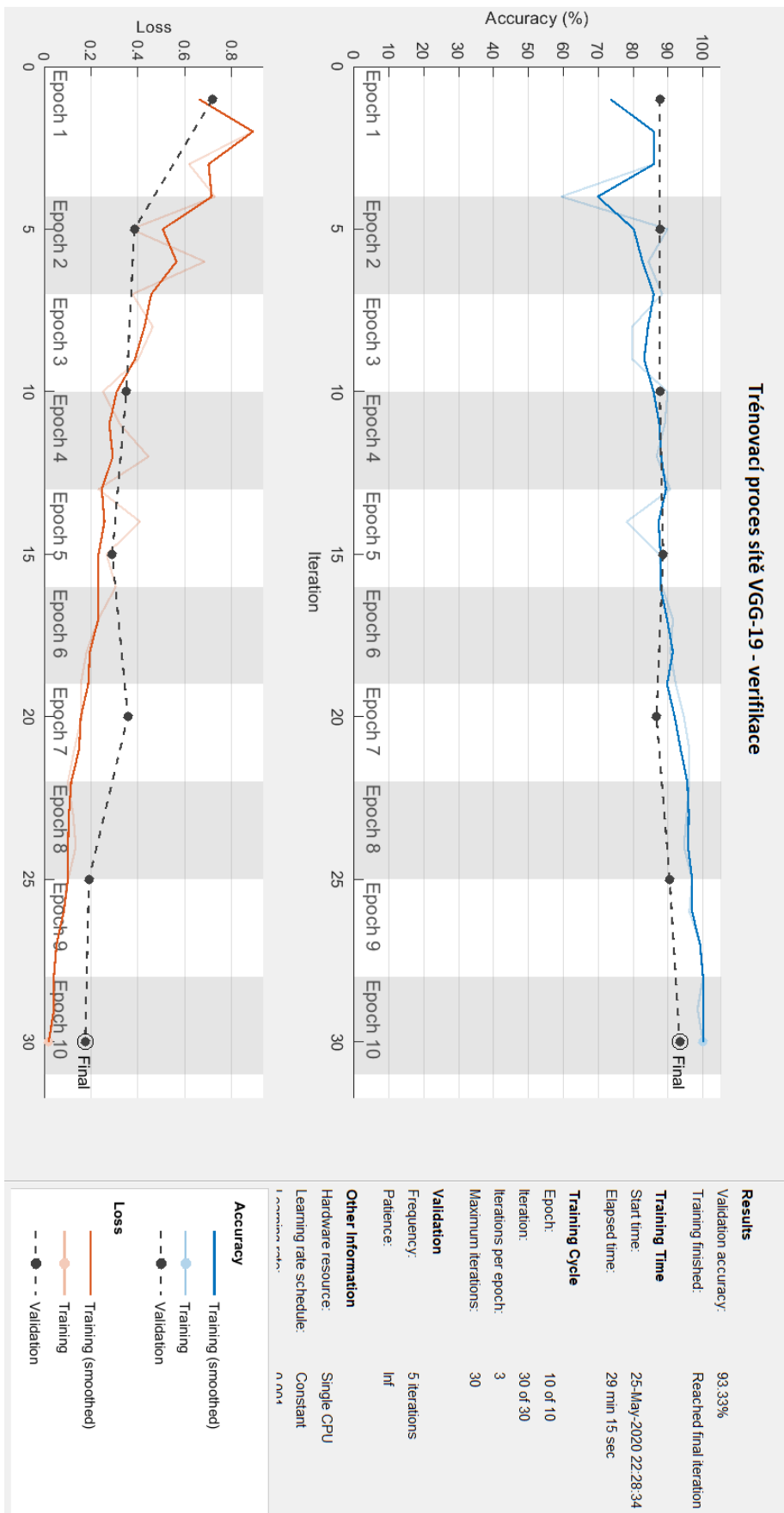
A.2 Obrazové přílohy - přehled trénovacích procesů jednotlivých sítí



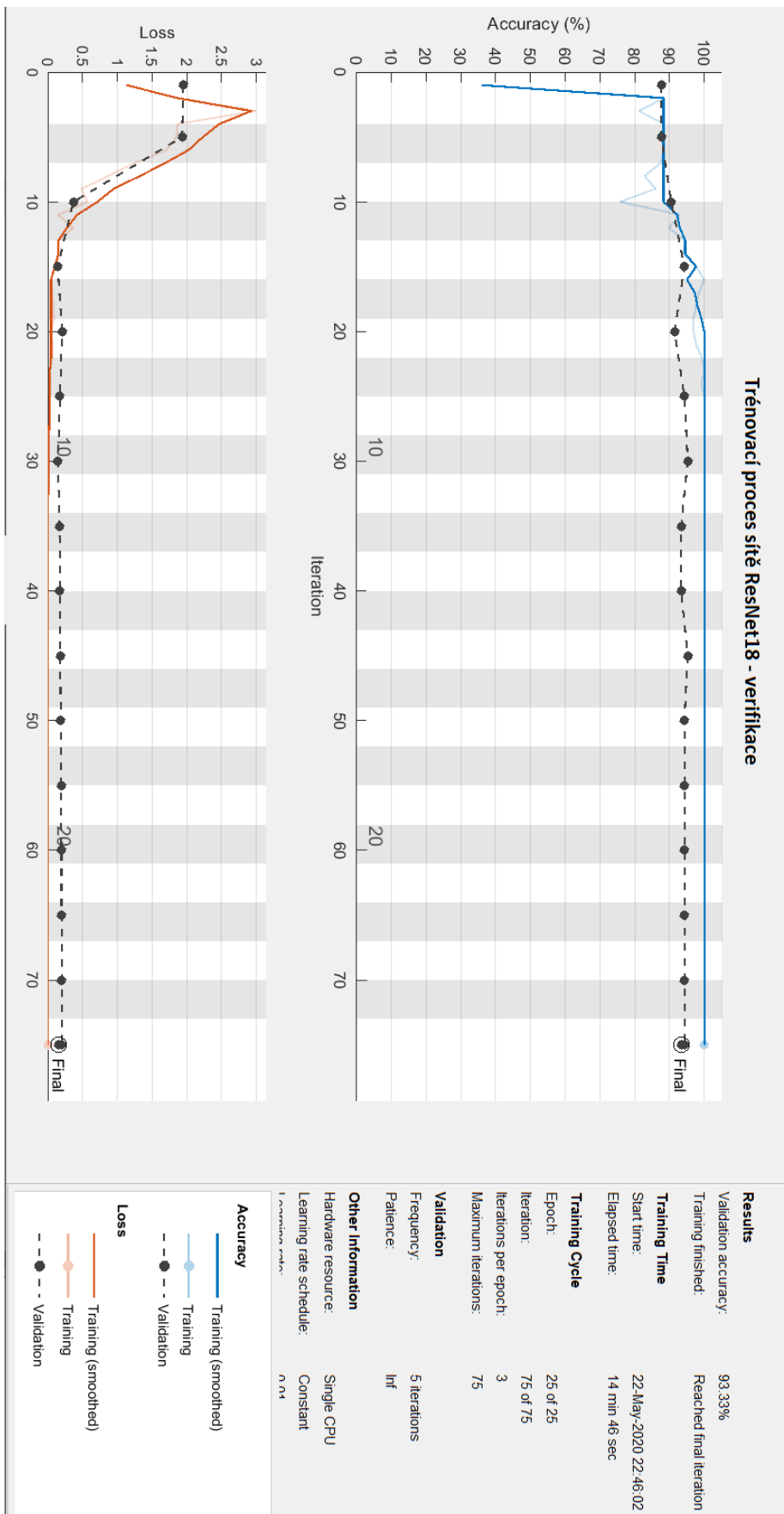
Obr. A.1: Trénovací proces sítě AlexNet v režimu verifikace.



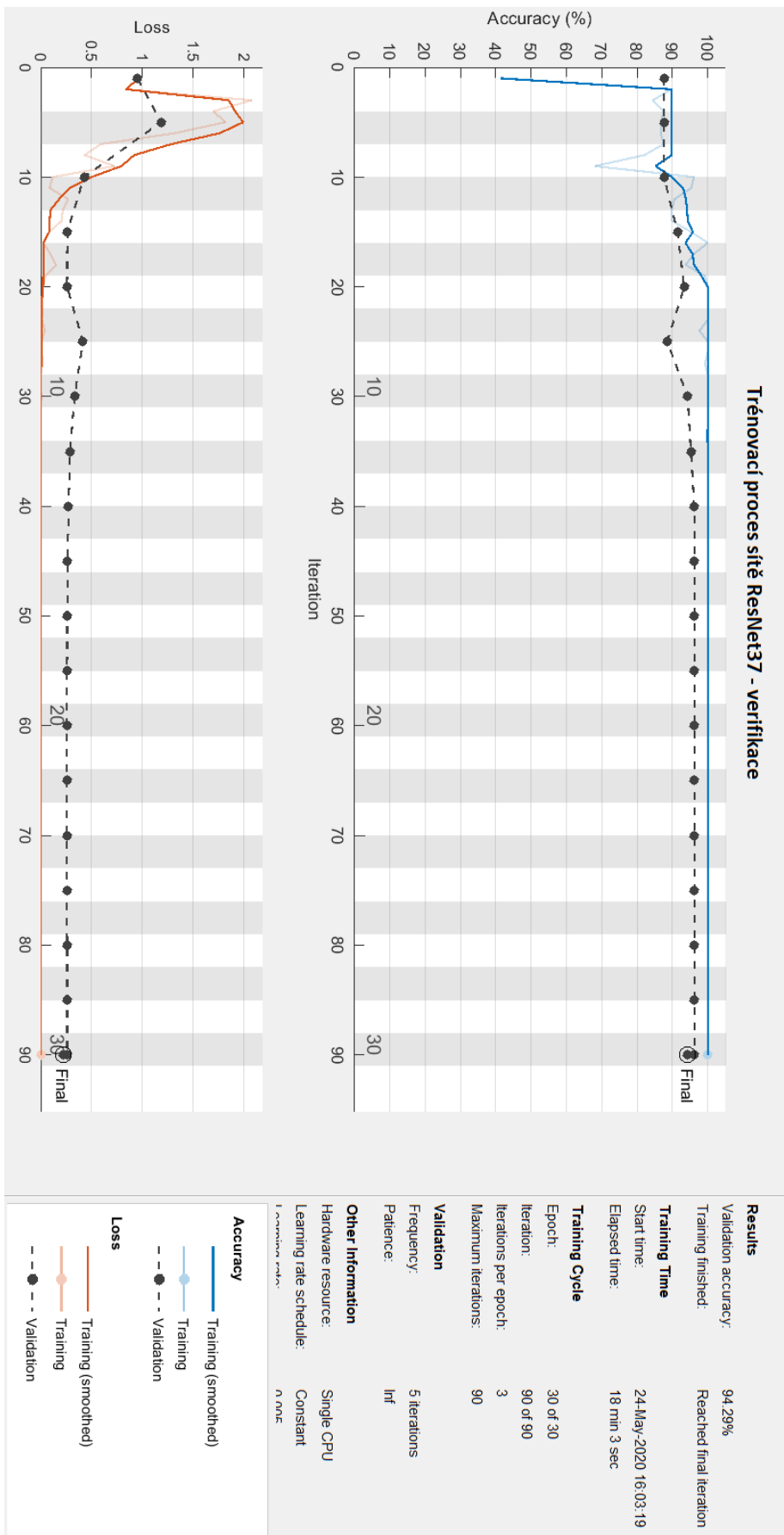
Obr. A.2: Trénovací proces sítě GoogLeNet v režimu verifikace.



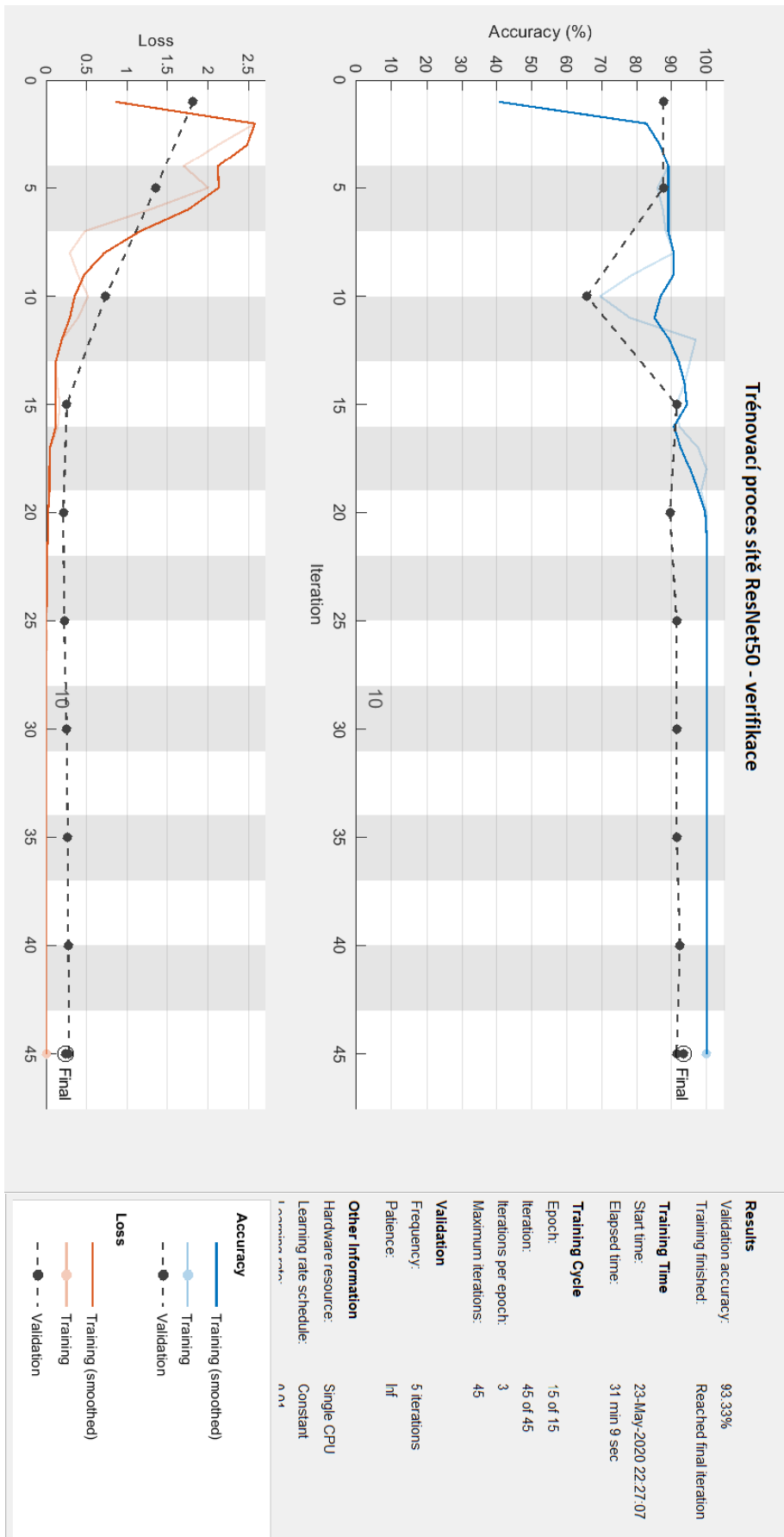
Obr. A.3: Trénovací proces sítě VGG-19 v režimu verifikace.



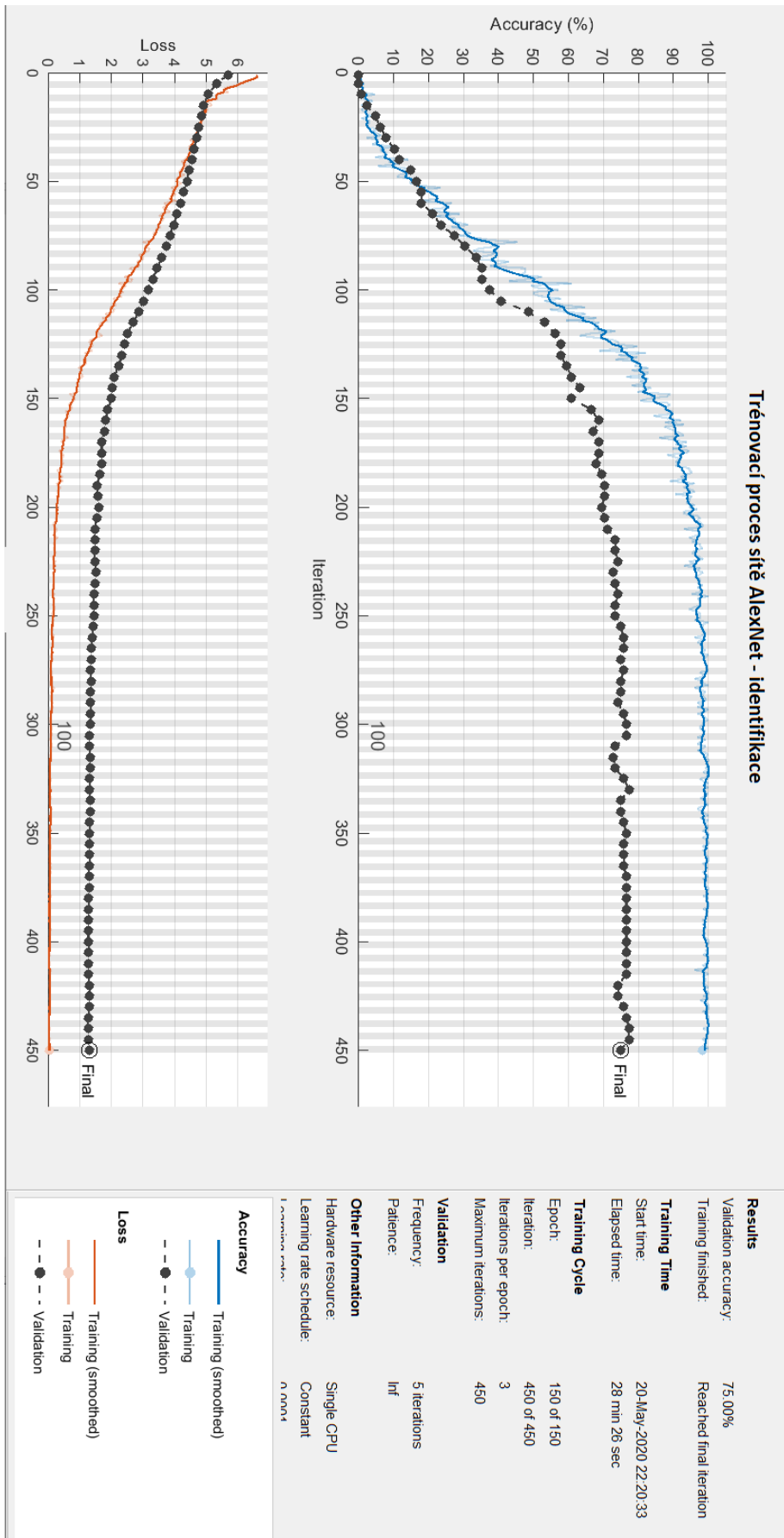
Obr. A.4: Trénovací proces sítě ResNet18 v režimu verifikace.



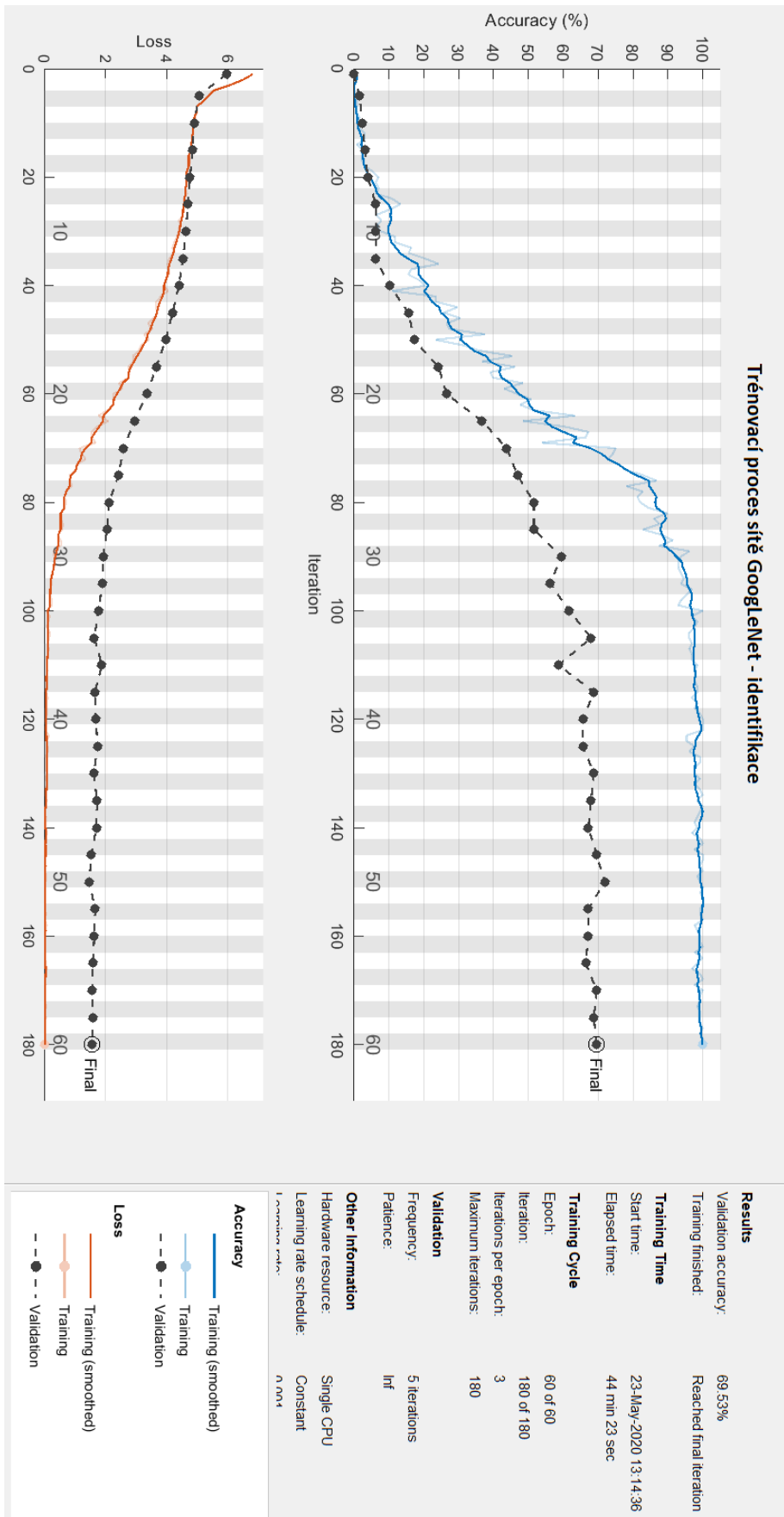
Obr. A.5: Trenovací proces sítě ResNet37 v režimu verifikace.



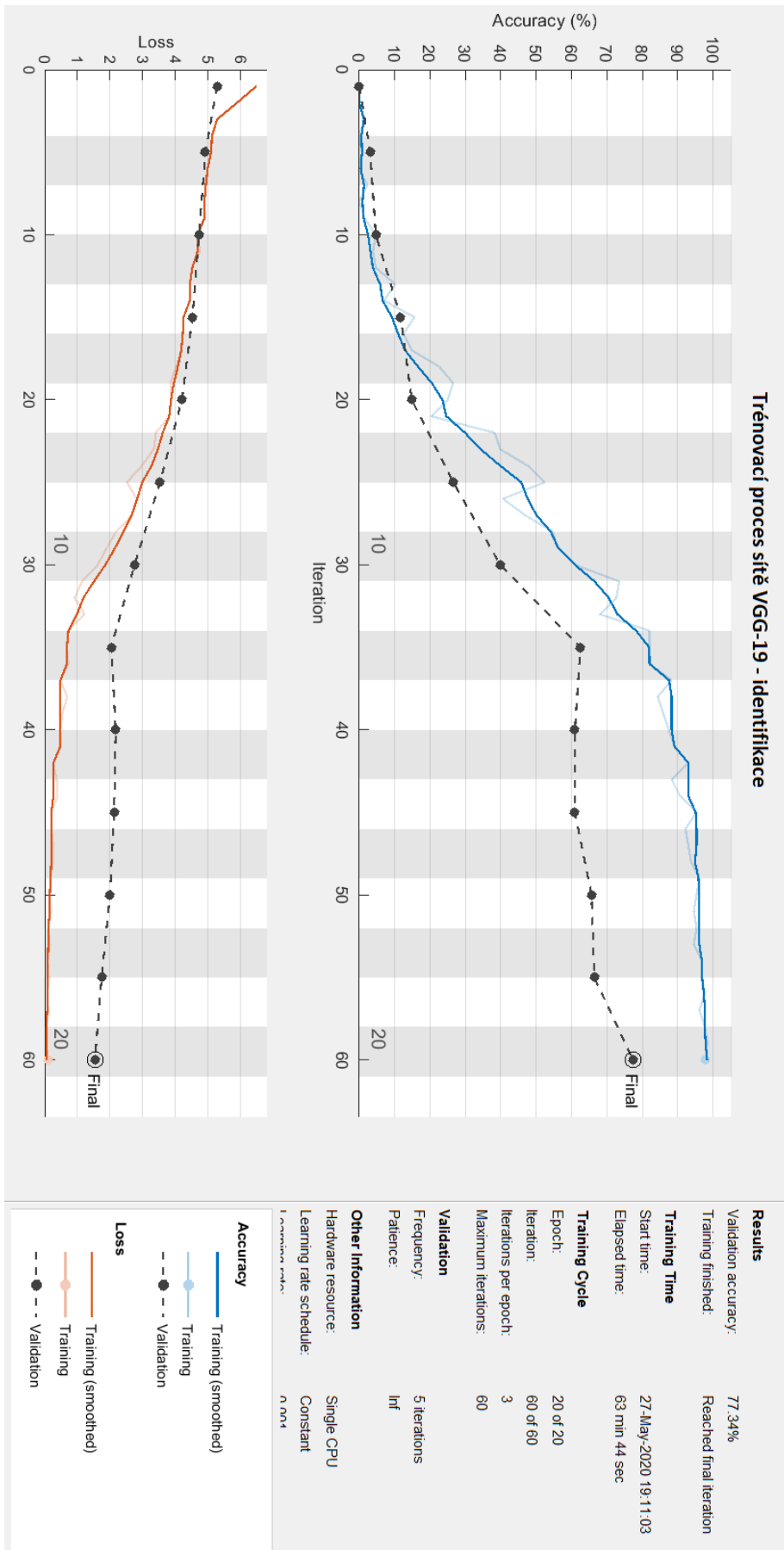
Obr. A.6: Trenovací proces sítě ResNet50 v režimu verifikace.



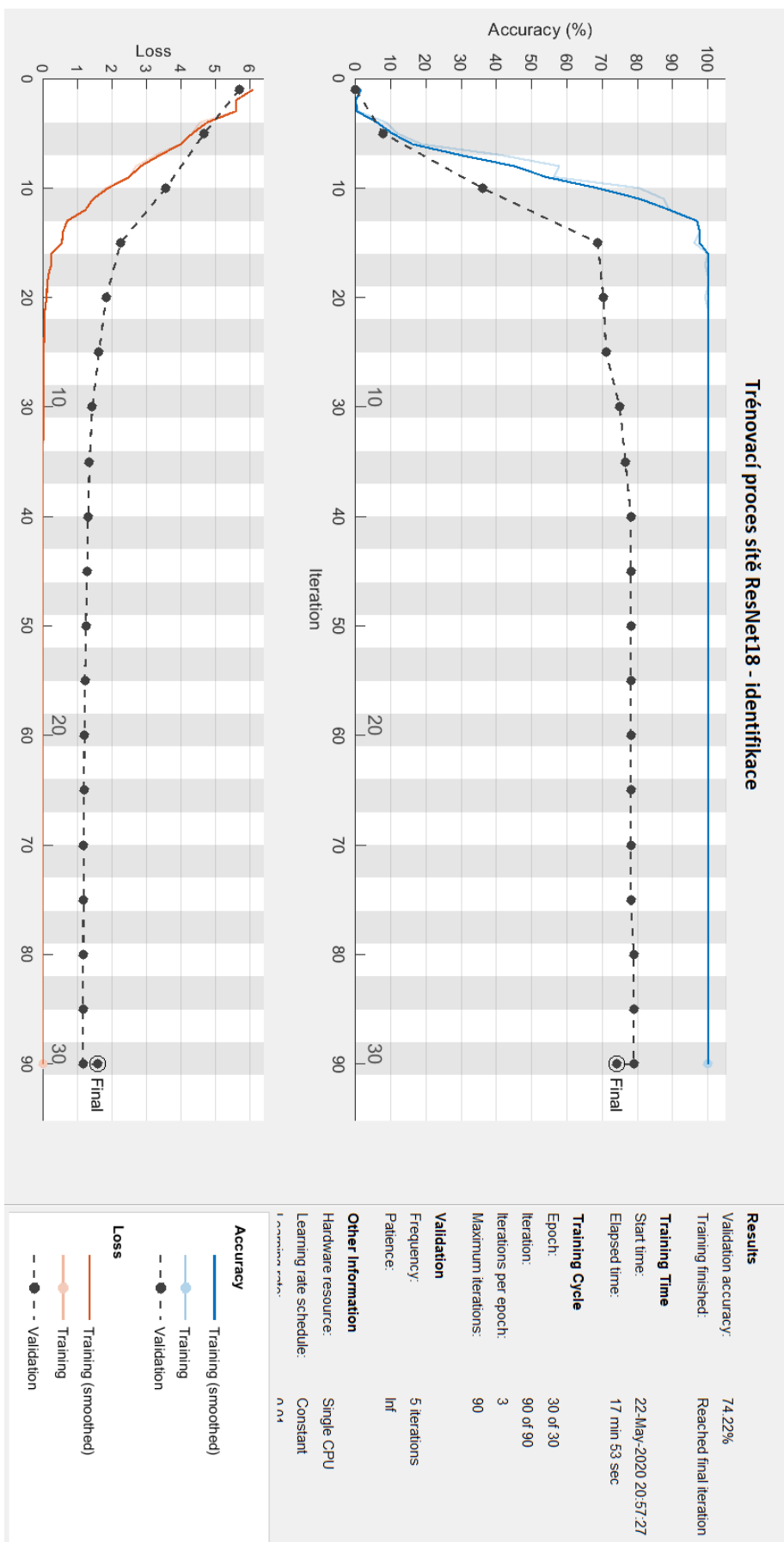
Obr. A.7: Trenovací proces sítě AlexNet v režimu identifikace.



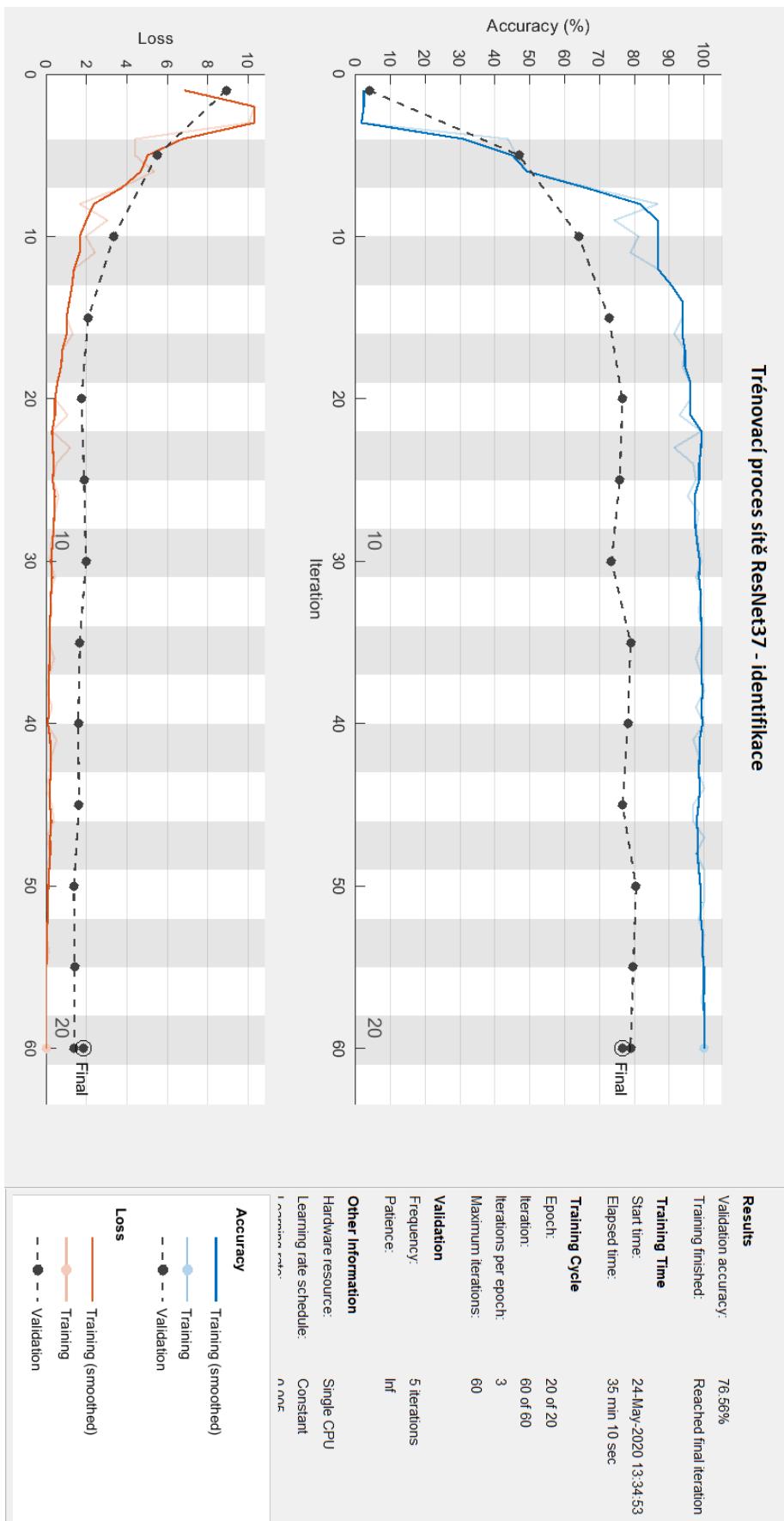
Obr. A.8: Trénovací proces sítě GoogLeNet v režimu identifikace.



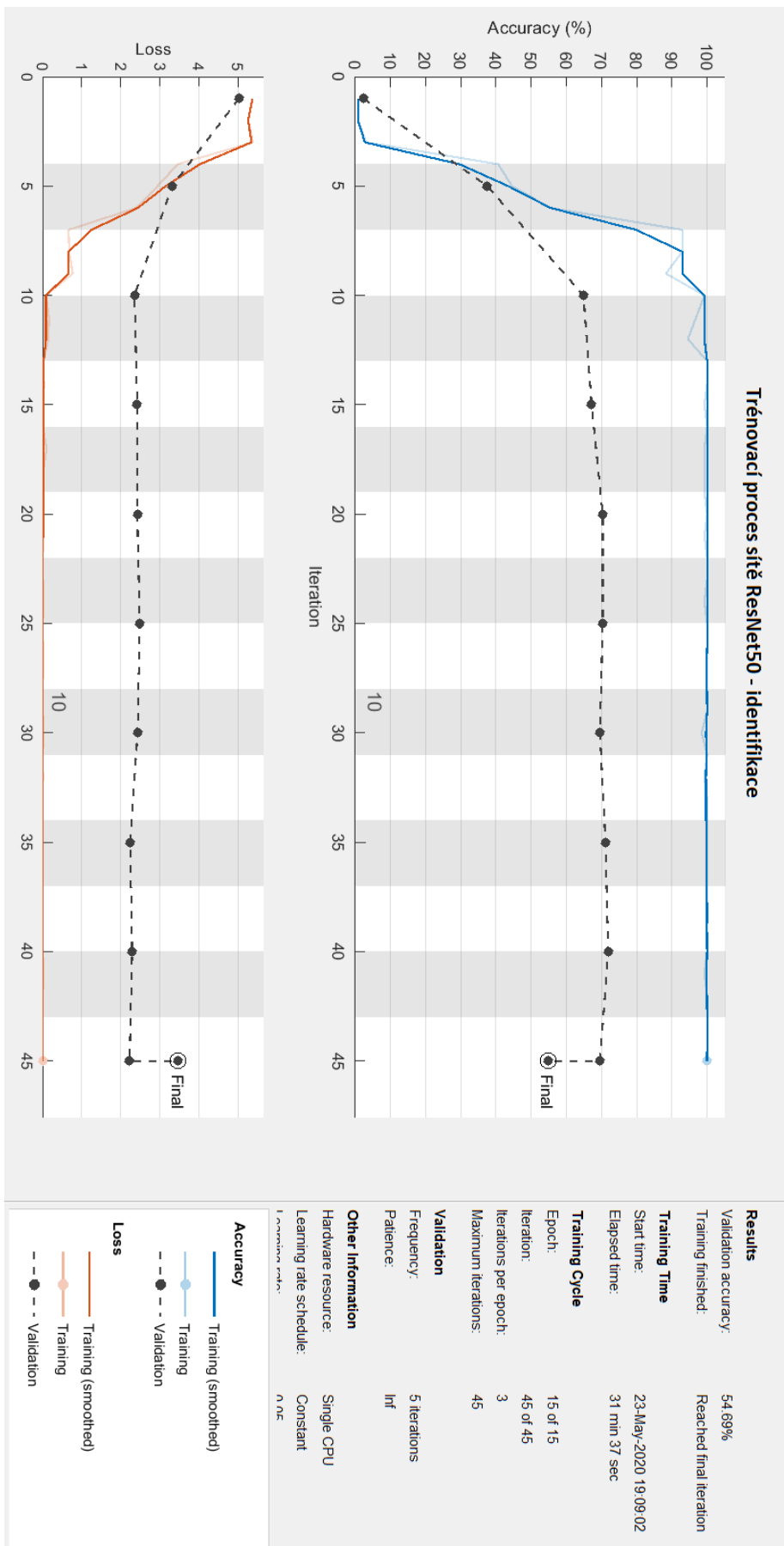
Obr. A.9: Trénovací proces sítě VGG-19 v režimu identifikace.



Obr. A.10: Trénovací proces sítě ResNet18 v režimu identifikace.



Obr. A.11: Trénovací proces sítě ResNet37 v režimu identifikace.



Obr. A.12: Trénovací proces sítě ResNet50 v režimu identifikace.