

Vysoká škola báňská – Technická univerzita Ostrava

Fakulta bezpečnostního inženýrství

Katedra bezpečnostních služeb

**Bezpečnostní řešení kamerové detekce za využití
Raspberry Pi**

Student: Bc. Michal Barbořík

Vedoucí diplomové práce: doc. Mgr. Ing. Radomír Ščurek, Ph.D.

Konzultant diplomové práce: Ing. Petr Vohnický

Studijní program: Požární ochrana a průmyslová bezpečnost

Studijní obor: Technická bezpečnost osob a majetku

Termín odevzdání diplomové práce: 18. 4. 2022

Rád bych poděkoval vedoucímu mé diplomové práce doc. Mgr. Ing. Radomíru Ščurkovi, Ph.D. za odborné vedení práce, poskytnuté informace, ochotu a čas. Nemenší poděkování také patří konzultantovi práce Ing. Petru Vohnickému za jeho cenné rady, připomínky a motivaci k lepším výsledkům a vědecké úrovni práce.

Anotace

Tato diplomová práce se zabývá konstrukcí kamerového systému na bázi jednodeskového počítače od nadace Raspberry Pi s inteligentním rozpoznáváním obličejů. Práce se zabývá teorií biometriky společně se strojovým učením. Dále autor ve své práci představuje použité prostředky, od hardwarových komponentů po využité softwarové prvky jako například detektory či klasifikátory. Další část práce je věnována experimentům, v nichž autor prověřil spolehlivost zvolených algoritmů. V závěru práce je pak finanční porovnání vytvořeného kamerového systému s komerčními produkty. Vytvořený kamerový systém je finančně méně náročný než komerční konkurenti, nicméně výsledky experimentů ukazují slabší obličejeový detekční rámec algoritmu v nočním režimu.

Klíčová slova: Raspberry Pi, strojové učení, hluboké učení, rozpoznání obličejů, jednodeskový počítač, neuronová síť, VSS, bezpečnostní kamera

Annotation

This masters thesis deals with the construction of a camera system based on a single-board computer from the Raspberry Pi Foundation with intelligent face recognition. The thesis deals with the theory of biometrics together with machine learning. Furthermore, the author presents the means used in his work, from hardware components to used software elements such as detectors or classifiers. The next part of the work is devoted to experiments where the author tested reliability of selected algorithms. At the end of the work is a financial comparison of camera system against a commercial product. The created camera system is less demanding than a commercial competitor, however, the results of experiments show a weaker face detection of the algorithm in night mode.

Keywords: Raspberry Pi, machine learning, deep learning, face recognition, single-board computer, neural network, VSS, security camera

Obsah

Obsah	4
Seznam použitých zkratek	6
Úvod a cíle práce	7
1 Literární rešerše a normativní požadavky	8
1.1 Normativní požadavky	10
2 Základní pojmy a definice	14
2.1 Biometrie	14
2.2 Raspberry Pi	15
2.2.1 Raspberry Pi 4 Model B	16
2.3 Strojové učení	18
2.3.1 Druhy strojového učení	20
2.3.2 Data mining, machine learning a deep learning	22
2.3.3 Face recognition system (systém rozpoznání obličeje)	26
3 Použité prostředky	29
3.1 Použitý hardware	29
3.1.1 Raspberry Pi 4 Model B	29
3.1.2 Waveshare RPi IR-CUT kamera	32
3.1.3 Kamerový stojan	35
3.2 Použitý software	37
3.2.1 Python	37
3.2.2 Konvoluční neuronová síť	40
3.2.3 Histogram orientovaných gradientů	43
3.2.4 Haar kaskádující klasifikátor	48
3.2.5 Použité skripty	50
3.3 Selekce detekční metody	55
4 Metodika měření	56
4.1 Laboratorní podmínky	57
4.2 Představení sledovaných veličin a poloh měření	59
4.3 Metody vyhodnocení výsledků	62
4.3.1 Matice záměn	62
4.3.2 ROC prostor	63
4.3.3 Histogram	65

5 Výsledky měření	66
5.1 Výsledky denního měření	66
5.2 Výsledky nočního měření	69
5.3 Vyhodnocení a diskuze	71
6 Finanční vyhodnocení.....	73
6.1 Raspberry pi	73
6.2 Komerční konkurence	74
6.3 Komparace	76
Závěr práce	79
Použitá literatura	80
Seznam obrázků.....	86
Seznam tabulek	88

Seznam použitých zkratek

AI	Aritificial inteligence
API	Application programming interface
CCTV	Closed-circuit television
CNN	Convolutional Neural Network
ČSN	Československá státní norma
DNA	Deoxyribonucleic acid
FPR	False positive rate
FullHD	Rozlišení 1920×1080
GPIO	General purpose input/output
HD	Rozlišení 1280×720
HOG	Histogram of oriented gradients
IoT	Internet of Things
IR	Infrared
LED	Light-emitting diode
OpenCV	Open Source Computer Vision Library
OS	Operační systém
RAM	Random-access memory
SD	Secure digital
TNR	True negative rate
TPR	True positive rate
USB	Universal Serial Bus
VSS	Video Surveillance System

Úvod a cíle práce

Vývoj technologií se každým rokem zrychluje, jsou vytvářeny nové způsoby efektivnějšího řešení starých problémů, ale stejně tak vznikají i nová pole možností, jež mohou mít negativní dopad na naše zájmy. Obor bezpečnosti se již od svého vzniku zabývá problémem lidského činitele jako zdrojem rizik, který je třeba neustále sledovat a snažit se předcházet konání těch jedinců, kteří mají za cíl poškození námi chráněných zájmů, ale i obecné zlovolné jednání. Řešení nebo alespoň regulaci tohoto problému nám mohou poskytnout právě dříve zmíněná technologie. Moderní technologie se dnes zabývají již pokročilejšími koncepty, jako umělá inteligence a strojové učení. Dále jde kupředu i rozvoj samotných hardwarových komponentů, které se stávají dokonalejšími a čím dál kompaktnějšími. Smyslem této práce je propojení těchto dvou faktorů v jeden a jeho zasazení do oboru bezpečnosti. Cílem je navrhnout hardwarové i softwarové řešení bezpečnostní kamery na bázi jednodeskového počítače, následné zhodnocení spolehlivosti, efektivity a stejně tak i finančního zatížení v porovnání s potenciálními konkurenty na českém trhu.

Diplomová práce je rozdělena do dvou pomyslných oddílů. První se zabývá teorií stanoveného problému. Jsou zde uvedeny jednodeskové počítače firmy Raspberry Pi, vysvětleny základy principu strojového učení a jeho odvětví hlubokého učení a rozpoznání obličejů. Druhá část práce se již více zabývá praktickým a experimentálním aspektem tématu. Nejdříve jsou představeny veškeré využité prostředky, které byly nutné pro dosažení stanovených cílů. Zahrnuje to popsání hardwarových komponentů, programovacího jazyka Python, a nakonec i představení potřebných klasifikátorů a detektorů. Práce pak přechází v představení podmínek experimentu společně s metodikou provádění zkušebních testů. Předposlední kapitola je pak věnována výsledkům práce a problematiku zakončuje finanční vyhodnocení a srovnání.

1 Literární rešerše a normativní požadavky

Lucy Hattersley [1] vydala v roce 2022 oficiální příručku k jednodeskovému počítači Raspberry Pi. Jedná se o soubor obrovského množství projektů vytvořených komunitou. Kniha je cenným zdrojem inspirace pro využití tohoto počítače s inovativními technologiemi, stejně tak může posloužit jako určitá forma návodu, protože obsahuje startovací instrukce a tutoriály, jak provést konfiguraci Raspberry Pi.

Jako alternativní knihu k předešlé příručce lze uvést celosvětově populární publikaci Simona Monka [2] vydaná v roce 2014. Tato publikace sice není cílená jako studnice nápadů, zato slouží jako mnohem komplexnější forma návodu, díky kterému lze snadno provést veškerý setup Raspberry Pi. Kniha kromě hardwarového řešení jako je připojení a ovládání sensorů, pohyblivých částí či displejů také zabíhá i do softwarové části projektů na bázi jednodeskového počítače, kde rozebírá jak základy programovacího jazyka Python, tak i pokročilejší věci jako třídy či moduly. Tato kniha, kterou nelze nazvat jinak než kuchařkou, sloužila jako kritický článek pro sestavení kompletního kamerového systému, který je představen v této práci. V roce 2019 napsal Jason Brownlee [3] knihu pojednávající o počítačové vizi a klasifikaci obrázků. V této knize představuje koncepty hlubokého učení a neuronových sítí v kontextu jazyka Python a mimo jiné zde také uvádí konvoluční neuronovou síť, která značnou částí přispěla k naplnění cíle této diplomové práce. Závěr knihy je také věnován objektové a obličejobě detekci, což je stěžejním pilířem tématu této závěrečné práce. Také publikace od Fabia Manganielleho [4] má za cíl představit strojové učení. Není tak obsáhlá, co se umělé inteligence týče, oproti knize předešlé, zato si klade za cíl realizaci detekci osob přímo na Raspberry Pi jak pomocí TensorFlow, tak OpenCV, který je důležitou knihovnou využívanou v detekčním skriptu diplomové práce.

Možnou variabilitu a potenciál využití jednodeskového počítače Raspberry Pi představil Matthew Poole (2015) [5], kde ve své knize představuje způsob, jak využít tento počítač jako možnou domácí ústřednu. Je zde uvedeno potenciální zapojení různých prvků jako jsou čidla a sledovací kamery. Další využití nalezneme v práci Raju A Nadafa et al. [6]. V tomto článku se autoři snaží vytvořit chytré zrcadlo, které by zároveň fungovalo jako bezpečnostní systém v případě absence majitele obydlí. V práci Nadafa je představena obličejobě detekce za pomocí OpenCV a Haar klasifikátoru, jež hraje důležitou roli i v této diplomové práci. V roce 2018 pak přišel Muhammad Saijid a kolektiv [7] s návrhem vytvoření smart security pro bezpečnostní

složky za využití Raspberry Pi, který bere v kontext jeho limitovanou výpočetní sílu a zohledňuje to při vývoji. Saijid s týmem vytvořili detekční framework, který má za cíl rozpoznání výrazu obličeje a tím potenciálního odhalení podezřelých aktivit. Tento framework s kombinací jednodeskového počítače má za cíl přispět ke konceptu smart security a ulehčit práci příslušníkům bezpečnostních složek. Pro demonstraci variability Raspberry Pi lze taktéž uvést práci Wesleyho J. McBrida a Jasona R. Courtera [8], kteří sestavili chytrý systém pro vzdálené sledování ptáků, za využití Raspberry Pi jako výpočetní jednotky, doplněný o energii ze solárních panelů.

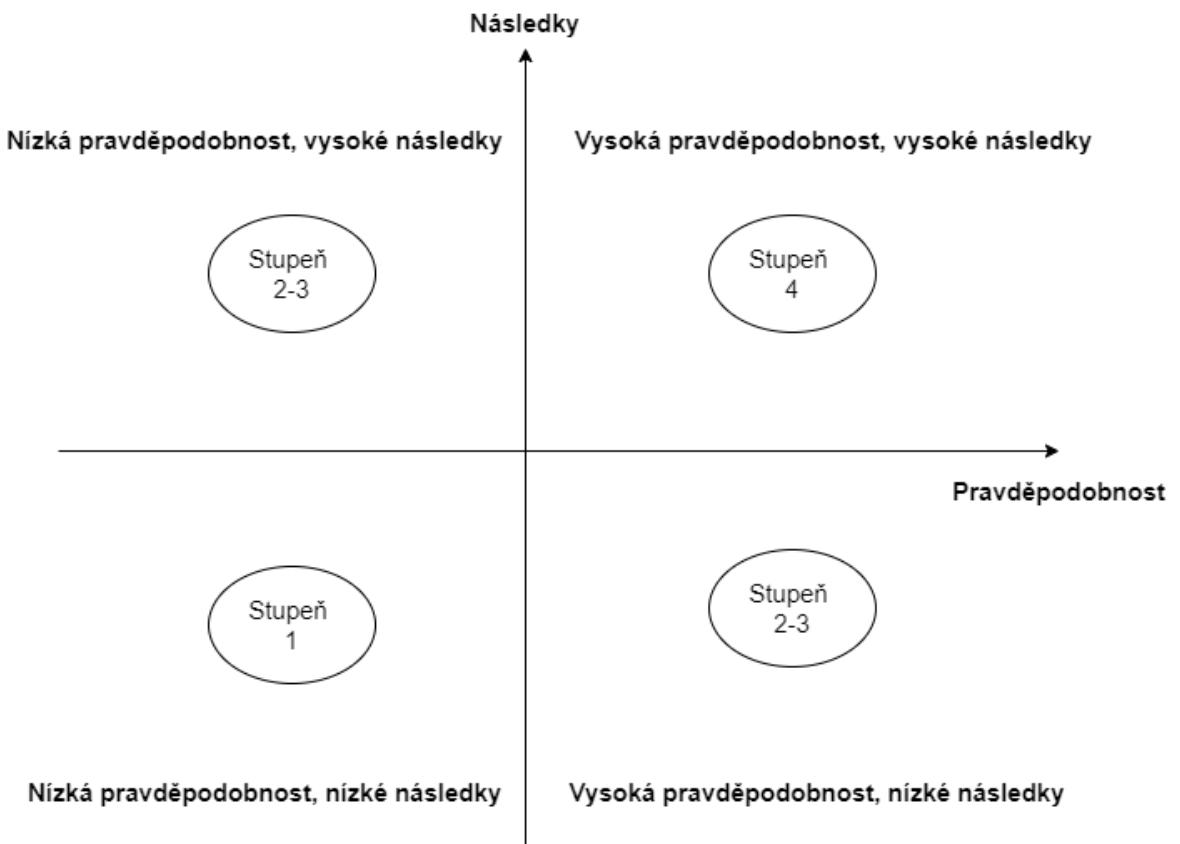
V roce 2020 vydal Ian Berle [9] publikaci zabývající se technologií obličejoblého rozpoznání v kontextu práv a soukromí. Autor v této knize představuje problém, jež tvoří tato technologie, a jaký zásah již má do současného světa a do našeho soukromí. Dále taktéž diskutuje vztah práva s ochranou dat a sledováním. Další téma, které autora znepokojuje v kontextu obličejoblé detekce, je státní paternalismus a v závěru autor uvádí další potenciální problémy budoucnosti v tomto kontextu. V roce 2021 pak Smith a Miller [10] taktéž vydali publikaci týkající se práva vztaženého k biometrickým metodám, nicméně se v této knize zabývají nejen právem pro obličejobou rekognici, ale i etikou při využívání jiných metod identifikace, jako třeba využití otisků prstů či identifikace podle deoxyribonucleic acid (DNA). Biometrické systémy jako možný komponent bezpečnostních systémů taktéž rozebírá Radomír Ščurek [11] ve své práci, kde uvádí formy biometrických technologií a zaobírá se jejich výhodami v kontextu bezpečnostních technologií, stejně jako možnostmi měření jejich spolehlivosti. Ve stejném roce pak Ščurek [12] vydává publikaci zaměřenou čistě na biometrické technologie v bezpečnostní praxi. Je zde vytvořen úvod do základních pojmu biometrie a následuje rozbor jednotlivých biometrických technologií využívaných v security aplikacích. Ščurek se zde také venuje metodám profilace osob a biosignálům neboli způsobům, jak získat od člověka biometrické informace.

1.1 Normativní požadavky

Kamerové systémy se dnes již označují jako Video Surveillance Systems (zkráceně VSS) oproti zastaralému označení Closed Circuit Television (CCTV). Tento moderní název také používá i aktuální balíček norem, který byl postupně vydáván od roku 2015 do roku 2019. Normy pro VSS nesou označení ČSN EN 62676-X-Y, kdy X je nahrazeno příslušným číslem, jež označuje část tohoto balíku a písmeno Y pak případným číslem dílu dané části, pokud je vůbec dělena. Balík norem je taktéž nazýván jako Dohledové videosystémy pro využití v bezpečnostních aplikacích. Norma je dělena na pět částí: první je věnována systémovým požadavkům, druhá video přenosovým protokolům, třetí analogovému a digitálnímu video rozhraní, ve čtvrté jsou uvedeny pokyny pro aplikaci a v poslední páté části jsou specifikace dat a kvality obrazu pro kamerové zařízení.[40]

V kontextu této práce se jeví první část tohoto balíku jako nevhodnější kandidát na bližší specifika předpokladů pro vhodný kamerový systém. První část je taktéž dělena na dva díly, kdy první díl se věnuje systémovým požadavkům obecně a druhý pak už konkrétněji výkonovým požadavkům na video přenos. Představen bude tedy první díl s označením ČSN EN 62676-1-1, který obsahuje několik důležitých náležitostí pro VSS.

Norma ČSN EN 62676-1-1 Dohledové videosystémy pro použití v bezpečnostních aplikacích – Část 1-1: Systémové požadavky – Obecně je členěna do osmi kapitol, ke kterým jsou přidruženy tři přílohy, dvě normativní a jedna informativní. Dokument je standardně uveden předmluvou, obsahem a taktéž úvodem. První kapitola je pak věnována rozsahu platnosti, druhá obsahuje citované dokumenty a třetí termíny, definice a zkratky. Čtvrtá kapitola je zaměřena na funkční popis VSS. Jedná se spíše o informaci, v nichž je popsáno fungování VSS ve schématu tří funkčních bloků. Prvním blokem je video prostředí, které reprezentuje snímání obrazu scény, následné zpracování získaných obrazů a předání vizualizace operátorovi. Druhým blokem je správa systému, která zastupuje uživatelské rozhraní pro správu aktivit a dat v rámci VSS, což je spojeno s komfortem využívání systému, stejně jako funkcionalitou i bezpečností. Poslední funkční blok je bezpečnost systému, která reprezentuje samotnou integritu systému i dat. Ačkoliv je tato kapitola pouze informativní, slouží jako vhodný nástroj pro uvedení kontextu fungování VSS. Další, pátá kapitola, se pak již zaobírá požadavky na systém, konkrétně stupně zabezpečení.[40]



Obrázek 1 Schéma rozdelení stupňů zabezpečení [40] [autor]

Každý VSS má přiřazený stupeň zabezpečení v hierarchické stupnici stanovené normou od 1 do 4. Tyto stupně zabezpečení pak pracují se dvěma faktory pro určení tohoto stupně: jedním je míra rizika, jež je stanovena pravděpodobností incidentu a druhým je množství potenciálních škod. První stupeň je označený jako nízké riziko, kdy VSS nemá žádnou ochranu a žádná omezení přístupu. Druhým je nízké až střední riziko, kdy má VSS nízkou úroveň ochrany a přístupu. Třetím stupněm je střední až vysoké riziko, kde je VSS s vysokou úrovní ochrany a jeho omezení přístupu je taktéž vysoké. Čtvrtým a posledním stupněm je vysoké riziko, kdy má VSS velmi vysokou úroveň ochrany i omezení přístupu. Schéma začlenění stupňů ve smyslu pravděpodobnosti a následků lze vidět v Obrázku 1.[40]

Šestá kapitola je označena jako funkční požadavky a je to obsáhlá část dokumentu, která obsahuje značné množství funkčních požadavků na VSS. V této práci budou zmíněny jen ty nejpodstatnější v kontextu zaměření práce. První značným požadavkem je ukládání nebo nahrávání snímků. Požadavky jsou rozděleny dle přiděleného stupně zabezpečení a jsou zobrazeny v Tabulce 1.[40]

Tabulka 1 Požadavky na ukládání [40]

VSS musí být schopen	Stupně zabezpečení			
	1	2	3	4
Zálohování dat nebo redundantní ukládání			X	X
Provozu zabezpečeného proti poruše nebo automatického přepnutí z jednoho paměťového média na druhé				X
Reakce na aktivační mechanismus s maximální latencí		1 s	500 ms	250 ms
Přehrání snímku z úložného prostoru s maximální latencí po incidentu nebo aktuálního záznamu			2 s	1 s

Další důležitý požadavek, stanovený v normě, je v sekci bezpečnosti systému. Jedná se o ochranu proti neautorizovanému přístupu. Norma stanovuje, že přístup musí být stanoven autorizačním schématem, které obsahuje několik úrovní přístupu. Úrovně jsou členěny následovně:[40]

- Úroveň 1: přístup jakoukoliv osobou, požadované funkce nesmí mít omezení.
- Úroveň 2: přístup jakýmkoliv uživatelem, funkce ovlivňující provoz systému beze změny nastavení.
- Úroveň 3: přístup administrátora systému, funkce ovlivňující konfiguraci systémových dat.
- Úroveň 4: přístup pro servisního technika nebo výrobce, přístup za účelem změny systémového návrhu či za účelem údržby.

V Tabulce 2 lze vidět specifikace, které funkce musí být dostupné na určitých přístupových úrovních bez závislosti na úrovni zabezpečení. P – povoleno, NP – nepovoleno.

Tabulka 2 Přístup k funkcím systému na základě přístupové úrovně [40]

Funkce	Přístupové úrovně			
	1	2	3	4
Konfigurace systému	NP	NP	P	P
Změna individuálních autorizačních kódů	NP	P	P	P
Přiřazování a mazání uživatelů a autorizačních kódů úrovně 2	NP	NP	P	P
Obnovení do továrního nastavení	NP	NP	P	P
Upgrade systému	NP	NP	P	P
Spouštění nebo vypnutí VSS	NP	NP	P	P

Předposlední kapitola je věnována třídám prostředí, které stanovují, jaké podmínky dané prvky v této třídě musí být schopny zdolat. Prvky vyšších tříd samozřejmě lze suplementovat za prvky nižších tříd. Třída prostředí I je určena pro uzavřené prostory, omezené na obytné či kancelářské prostředí. Jedná se o prostory, ve kterých je udržována stálá teplota v rozmezí +5 °C až +40 °C s relativní vlhkostí 75 % bez kondenzace. Třída prostředí II je určena pro uzavřené prostory, nicméně s obecným určením. Jedná se o prostory, ve kterých není udržována konstantní teplota, ačkoliv by měla být v rozmezích -10 °C až +40 °C s relativní vlhkostí 75 % bez kondenzace. Třída prostředí III je určena pro vnější prostory, které jsou kryté před deštěm a sluncem nebo pro vnitřní prostory s extrémními podmínkami. Je to tedy takový prostor, kde prvky nejsou vystaveny plnému vlivu počasí. Teplota by měla být v rozmezí -25 °C až +50 °C s relativní vlhkostí 75 % bez kondenzace. Výjimkou zde je, že 30 dní v roce je povolena vlhkost v rozmezí 85 % - 95 % bez kondenzace. Třída prostředí IV je určena pro vnější prostory s obecnějším záběrem, podobně jako třída II. Jedná se o prvky, které jsou již plně vystaveny vlivům počasí. Příklad možných rozmezí je totožný jako u předešlé třídy.[40]

Poslední, osmá kapitola je pak věnována dokumentaci samotné. Zde je zařazena dokumentace systému jako takového, jednotlivé provozní instrukce a dokumentace prvků.[40]

2 Základní pojmy a definice

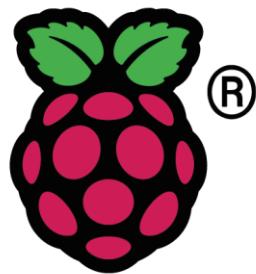
Jako vhodný úvod do komplexní problematiky této práce se jeví představení biometrie jako vědní disciplíny v kontextu bezpečnostního oboru. Kapitola dále pokračuje představením konceptu jednodeskových počítačů Raspberry Pi a v závěru jsou uvedeny základní poznatky o strojovém učení, které je klíčovým prostředkem využívaným v této práci.

2.1 Biometrie

Biometrie je vědní obor, jehož cílem je studium živých organismů s převážným zaměřením na ten lidský. Zabývá se měřením biologických vlastností člověka společně s behaviorálními znaky, tedy způsobem, jakým se chová a jaké má návyky a vzorce. Jedná se o soubor metod, jak rozpoznávat lidskou bytost na základě jejích typických vlastností. Ideálním příkladem je rozpoznávání obličejů, kdy lidský mozek rozpoznává jednotlivce na základě rysů v obličeji. Biometrické metody hledají v námi poskytnutém vstupu možné znaky a vzorce, které pak následně analyzují a rozhodují o výsledku. Pro správný výklad následujícího textu zabývajícího se biometrií, je vhodné také představit pár základních pojmu, jež jsou kritické pro správné porozumění studovaného problému. První pojem je rekognice neboli rozpoznávání. Jedná se o termín, který určuje rozpoznání osoby na základě vhodné tělesné vlastnosti. Verifikace neboli ověření, je označení pro proces, kdy se stanovený biometrický systém snaží potvrdit totožnost jedince za pomoci sejmutého vzorku a jeho porovnání se vzorkem, jež byl dříve zapsán v systému. V kontextu biometrie hovoříme o principu jeden na jednoho. Pojem, který se váže na verifikaci, je identifikace. Taktéž se jedná o proces, při kterém se snaží systém určit totožnost neznámého jedince, nicméně sejmutá biometrická informace je nyní srovnávána s celou databází referenčních vzorků a jedná se tedy o princip jeden na mnoho. Posledním termínem, jež se často vyskytuje v kontextu biometrie, je autentizace, nebo také někdy uváděna jako legalizace. Je to pojem, který je úzce svázán s termínem rozpoznání, nicméně výstupem toho procesu je rozhodnutí nebo status, zdali je rozpoznávaný jedinec zadaný v systému jako oprávněný či neoprávněný. Metody autentizace se v základu uvádějí tři, a to použitím hesla, předmětu nebo právě využitím biometrického prvku, které mají výhodu životní neměnnosti za běžných okolností.[11, 12]

2.2 Raspberry Pi

Raspberry Pi je název, který je používaný pro dnes již obsáhlou sérii jednodeskových počítačů produkovaných Raspberry Pi nadací, jejíž logo (viz Obrázek 2) se stalo ikonou mezi nadšenci a hobbisty, ale stejně tak i mezi profesionály využívající je při každodenní činnosti. Jedná se o charitu sídlící ve Spojeném království a mající za cíl edukaci obyvatelstva v poli počítačových technologií a taktéž co nejvíce umožnit přístup a rozvoj v tomto směru. Zcela první model Raspberry Pi byl uveden do prodeje v roce 2012 a od té doby bylo vydáno mnoho variací i opakovaných výrob stejných modelů. Lidé po celém světě využívají tyto počítače pro výuku programování, stavění hardwarových projektů, vytvoření domácího automatizačního systému, anebo například pro implementaci Kubernetes či Edge computingu a dokonce jsou využívány v industriálních aplikacích.[1, 2, 17]



Obrázek 2 Logo Raspberry Pi [15]

První model disponoval procesorem s jediným jádrem o frekvenci 700 MHz a pouhých 256 MB RAM paměti. V kontrastu se začátky, nejnovější model Raspberry Pi 400 uvádí na trh již opravdu plnohodnotný způsob nahradu klasického počítače, jak jej známe. Tento model je osazen čtyřjádrovým procesorem o výkonu 1,8 GHz, dále 4 GB RAM pamětí, u které byl využit její nový typ – LPDDR4, což umožnilo výrazné zvýšení její rychlosti. Počítač je zabudován do klávesnice (viz Obrázek 3), což z něj činí opravdu jednoduchý nástroje pro masové využití například ve školách díky schopnosti plug-and-play.[1, 2, 16, 17, 18]

Raspberry počítače využívají pro svůj běh nejčastěji distribuci Linux operačního systému (OS) ve variantě Debian OS, který je nazývána taktéž jako Raspbian či Raspberry Pi OS. Tyto počítače jsou taktéž osazeny general purpose input/output (GPIO) piny, které

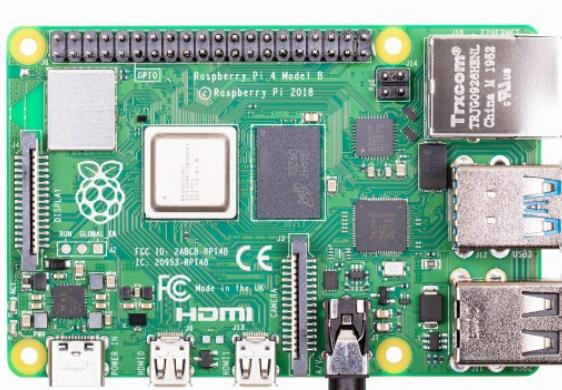
umožňují ovládání elektronických komponentů a často jsou využívány pro konstrukci Internet of Things (IoT).[1, 2, 16, 17, 18]



Obrázek 3 Raspberry Pi 400 [16]

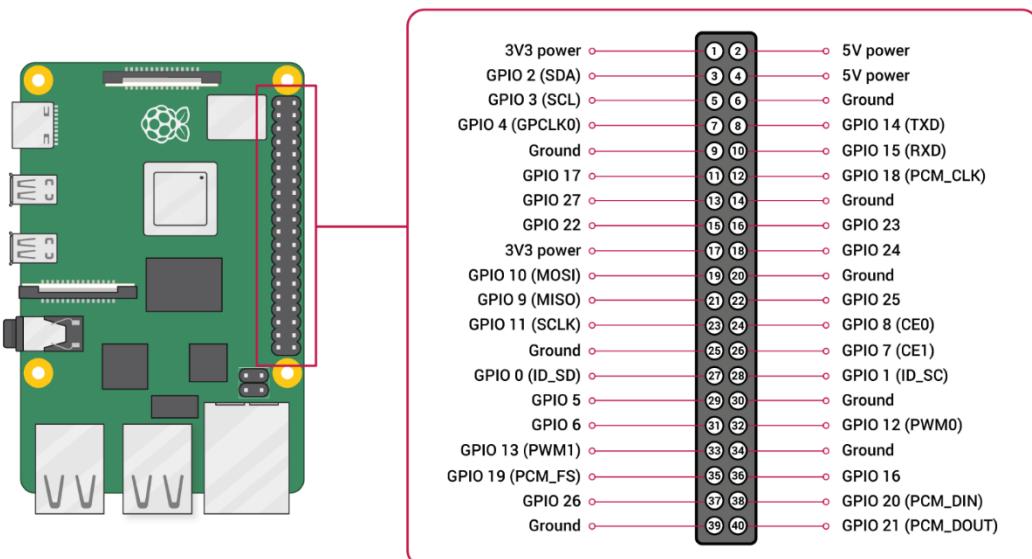
2.2.1 Raspberry Pi 4 Model B

Tento počítač bude následovně detailněji představen, jelikož byl zvolen jako ideální kandidát pro dosažení cílů práce. Uvedený model byl uveden na trh v červnu roku 2019. Počítač byl vybaven čtyřjádrovým procesorem (ARM Cortex-A72) o pracovní frekvenci 1,5 GHz a 64 bitech, random-access-memory (RAM) pamětí o hodnotách 2 až 8 GB, grafikou VideoCore VI s rozlišením 4Kp60 a režimem dvou monitorů, Wi-Fi se standardem 802.11ac, Bluetooth 5 a gigabitovým ethernet portem.[1, 2, 17, 19]



Obrázek 4 Raspberry Pi 4 Model B – pohled shora [19]

Deska plošných spojů je samozřejmě také osazena velkým množstvím užitečných příslušenství (pohled shora lze vidět na Obrázku 4), nicméně rozměry si zachovává stejné, jako předešlé modely. První znatelná výhoda oproti předešlým modelům je povýšení napájení na konektor USB-C, což zpřístupní napájení náročnějších periférií a taktéž vyšší výkon samotného počítače bez nutnosti speciálního napájecího zdroje. Dále byly přidány dva microHDMI porty, oproti staršímu jednomu HDMI. Jak již bylo zmíněno výše, Raspberry disponuje Ethernetovým připojením, vedle něj jsou pak universal seriál bus (USB) konektory, kdy dva jsou verze 3.0 a dva 2.0. Mimo jiné je deska osazena konektorem pro připojení displeje (MIPI DSI) a konektorem pro připojení kamery (MIPI CSI), společně s čtyřpolovým 3,5 mm jackem pro výstup zvuku a kompozitního videa. V poslední řadě je důležité zmínit, že jako každý model je i tento osazen 40 GPIO pinů pro ovládání a napájení jiných zařízení, jejichž rozložení zůstalo stejně jako u předcházejícího modelu. Schéma rozložení těchto pinů a jejich funkce lze vidět v Obrázku 5.[1, 2, 19]

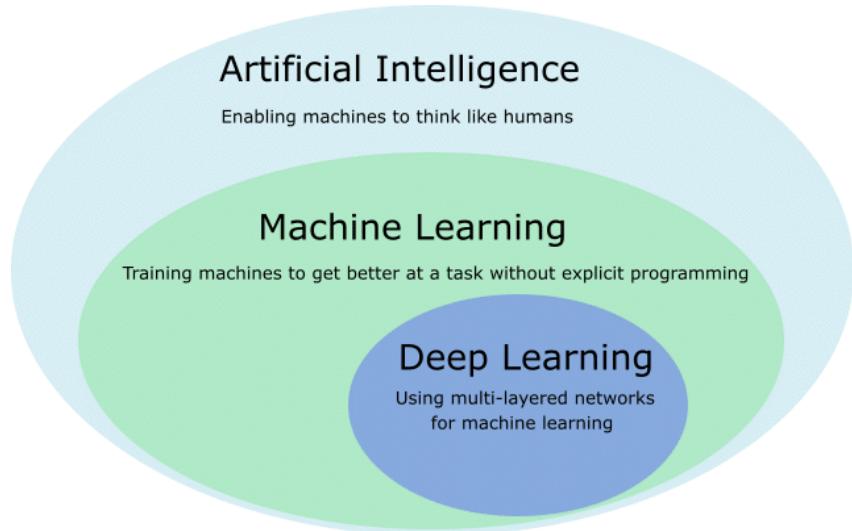


Obrázek 5 Rozložení GPIO pinů u nejnovější generace modelů [20]

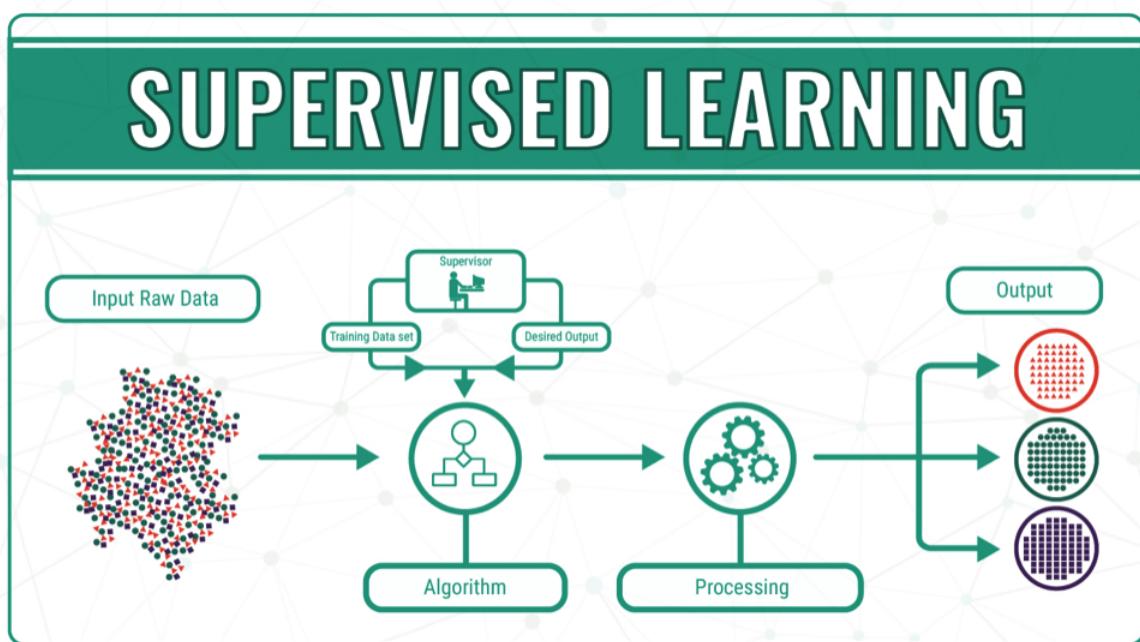
2.3 Strojové učení

Pod pojmem strojové učení (v angličtině označované jako machine learning) si lze představit metodu datové analýzy, která automatizuje stavbu analytického modelu. Jedná se tedy o odvětví umělé inteligence, stavějící na myšlence, že samotný systém je schopen se učit z poskytovaných dat, rozpoznávat určité vzory a zároveň utvářet rozhodnutí s minimálním zásahem člověka.[3, 22]

Dnes již strojové učení, díky novým technologiím, není stejným odvětvím, jako dřív. Toto pole odbornosti bylo zrozeno na základech rozeznávání vzorů a teorie, že počítače se mohou samy učit bez nutnosti naprogramování pro výkon daného účelu – výzkumníci tedy chtěli dosáhnout toho, že umělá inteligence se bude učit z poskytnutých dat. Z toho plyne poznatek, že iterativní aspekt strojového učení je klíčem k úspěchu, protože jak jsou nová data představována daným modelům, jsou pak tyto modely schopny se nezávisle a samostatně na ně adaptovat. Učí se z předešlých výpočtů, aby vyprodukovaly spolehlivá, a především opakovatelná rozhodnutí a výsledky. Pro laika se může zdát pojem umělá inteligence a strojové učení jako totožné věci, nicméně tomu tak není, ačkoliv pojmy spolu souvisí. Umělá inteligence (artificial intelligence, AI) je samostatný vědní obor, který se zaobírá možnostmi, jak umožnit strojům a počítačům mimikovat lidské chování a rozhodování – jednoduše řečeno provádět lidské úkony. Strojové učení je pak součástí této problematiky jako celku. Jedná se již o samotnou metodu, která učí stroje se učit. Hledají se určité vzory v datech a na jejich základě se vytváří závěry. Speciální částí strojového učení je tzv. deep learning neboli hluboké učení. Této problematice bude věnována jedna z následujících kapitol. Jednoduché rozvržení lze vidět v Obrázku 6.[3, 13, 22]



Obrázek 6 Schéma oboru umělé inteligence. V překladu: Artificial intelligence – Umělá inteligence, umožňuje stojům myslet jako lidé; Machine Learning – Strojové učení, trénování strojů za účelem jejich zlepšení bez explicitního naprogramování; Deep Learning – Hluboké učení, využití mnohovrstvých sítí pro strojové učení [21]



Obrázek 7 Schéma průběhu učení s učitelem. V překladu: Supervised learning – Učení s učitelem; Input raw data – Hrubá vstupní data; Supervisor – Učitel; Training data set – Trénovací datový soubor; Desired output – Požadovaný výstup; Algorithm – Algoritmus; Processing – Zpracování; Output – Výstup [24]

2.3.1 Druhy strojového učení

Za nejrozšířenější a nejčastěji používané druhy strojového učení se považuje učení s učitelem (supervised learning), učení bez učitele (unsupervised learning), kombinace učení s učitelem a bez učitele (semi-supervised learning) a zpětnovazebné učení (reinforcement learning) někdy označované jako učení posilováním.[3, 13, 23]

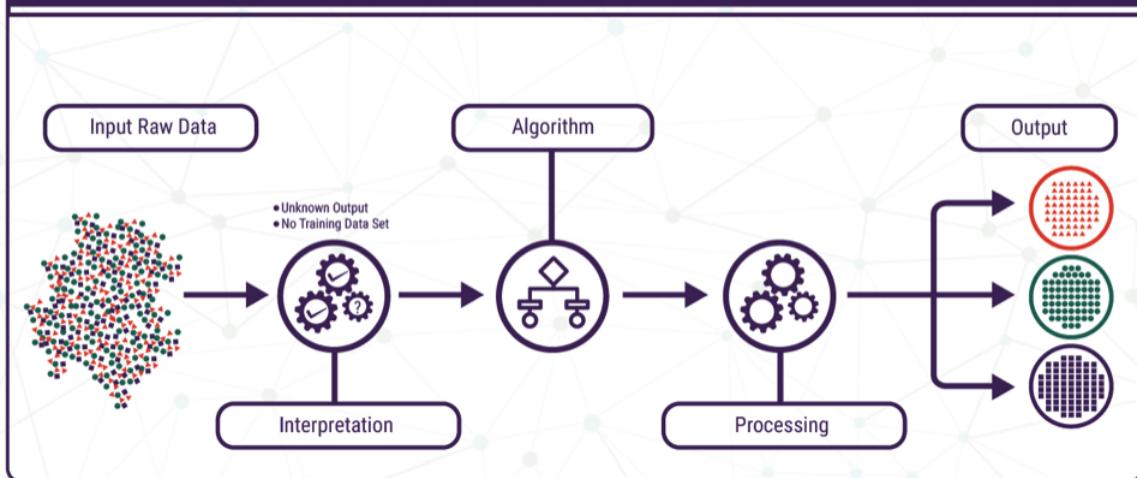
Učení s učitelem je způsob, při kterém je stroj učen příklady. Operátor poskytne danému algoritmu známý datový soubor, obsahující požadované vstupy a výstupy, a tento algoritmus musí vytvořit způsob, kterým dosáhne požadovaných výsledků, tedy jak ze zadaných vstupů dostane zadané výstupy. Zatímco operátor zná správné odpovědi, algoritmus identifikuje vzorce v poskytnutých datech, učí se jejich pozorováním a vytváří predikce. Tyto utvořené predikce pak prezentuje operátorovi, který jej případně opraví, byla-li by některá z nich nesprávná. Tento proces se pak nadále opakuje, dokud není dosaženo vysoké přesnosti a výkonu. Schéma celého průběhu učení lze vidět na Obrázku 7.[3, 13, 23]

Do tohoto způsobu učení spadá:

- Klasifikace (classification): V klasifikačních úlohách musí daný program dojít k závěru z pozorovaných hodnot a rozhodnout, do které kategorie nová pozorování spadají. Příkladem může být filtrace spamů v emailu.[13, 23]
- Regrese (regression): V regresních úlohách musí program odhadnout a pochopit vztah mezi jednotlivými proměnnými. Regresní analýza se zaměřuje na jednu závislou proměnnou a sérii dalších měnících se proměnných, využívá především pro predikce.[13, 23]
- Předpovídání (forecasting): Předpovídání je proces utváření predikcí o budoucnosti na základě minulých a současných dat a využívá se nejčastěji pro analýzu trendů.[13, 23]

Druhý způsob je kombinace učení s učitelem a bez učitele, které je velmi podobné předešlému způsobu s tím rozdílem, že jak označená, tak neoznačená data jsou poskytnuta danému algoritmu. Označená data jsou ta, která mají nějaké významné štítky či označení, aby jim mohl algoritmus porozumět. Neoznačená, jak již z názvu vyplývá, tyto značky nemají. Kombinací těchto technik lze dosáhnout toho, že algoritmus je schopen se naučit označovat neoznačená data.[3, 13, 23]

UNSUPERVISED LEARNING



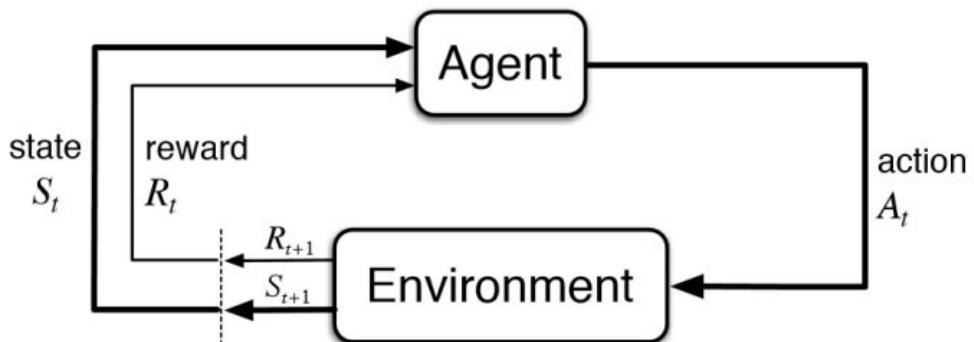
Obrázek 8 Schéma průběhu učení bez učitele. V překladu: **Unsupervised learning – Učení bez učitele; Interpretation – Interpretace; Unknown output – Neznámý výstup; No training data set – Bez trénovacího datového souboru [24]**

Při učení bez učitele algoritmus studuje data pro následnou identifikaci vzorů. Není zde žádný klíč k odpovědi ani operátor, který by předal instrukce. V tomto případě stroj usiluje o určení korelace a vztahu analýzou poskytnutých dat. Při tomto druhu učení je algoritmus ponechán, aby interpretoval velké datové soubory a následně je řešil odpovídajícím způsobem. Snaží se je organizovat určitým způsobem pro popis jejich struktury. To může znamenat sestavení dat do shluků nebo seřazení dat tak, aby vypadala více organizovaně. Čím více dat posuzuje, tím se jeho schopnost udělat rozhodnutí na tato data postupně zvyšuje a stává se více efektivnější. Schéma celého průběhu lze opět vidět na Obrázku 8.[3, 13, 23]

Je zde zahrnuto:

- Shlukování (clustering): Shlukování zahrnuje seskupení podobných dat na základě definovaných kritérií. Bývá často využíváno pro segmentaci dat do několika skupin a provádění následných analýz na každém datovém souboru pro zjištění možných vzorů v nich.[13, 23]
- Zmenšení rozměru (dimension reduction): Snižuje množství proměnných, které jsou brány v potaz pro nalezení přesně požadované informace.[13, 23]

Učení posilováním se zaměřuje na striktně regulovaný učící proces, kdy vzdělávanému algoritmu je poskytnu několik sad akcí, parametrů a konečných hodnot. Po stanovení pravidel se algoritmus snaží objevovat různé možnosti a příležitosti, monitoruje a vyhodnocuje každý výsledek pro stanovení jednoho, který je nevhodnější. Tato technika vyučuje pomocí stylu pokus omyl. Učí se ze zkušeností z minulosti a snaží se adaptovat svůj postup jako odpověď na danou situaci pro dosažení nejlepších možných výsledků. V tomto druhu učení se vyskytuje tři komponenty agent (učenec, nebo ten, co dělá rozhodnutí), prostředí (vše s čím agent interaguje) a akce (vše co může agent dělat). Schéma závislostí těchto tří komponentů lze vidět v Obrázku 9.[3, 13, 22, 23]



Obrázek 9 Schéma učení posilováním. V překladu: Agent – Agent/učenec; Action – Akce; Environment – Prostředí; Reward – Odměna; State – Stav [24]

2.3.2 Data mining, machine learning a deep learning

Ačkoliv všechny tyto tři metody (data mining, machine learning a deep learning) mají stejný cíl, a to extrakci vhledů, vzorů a vztahů, které mohou být použity pro utváření rozhodnutí, přesto se liší ve svých způsobech přístupu a také i schopnostech.

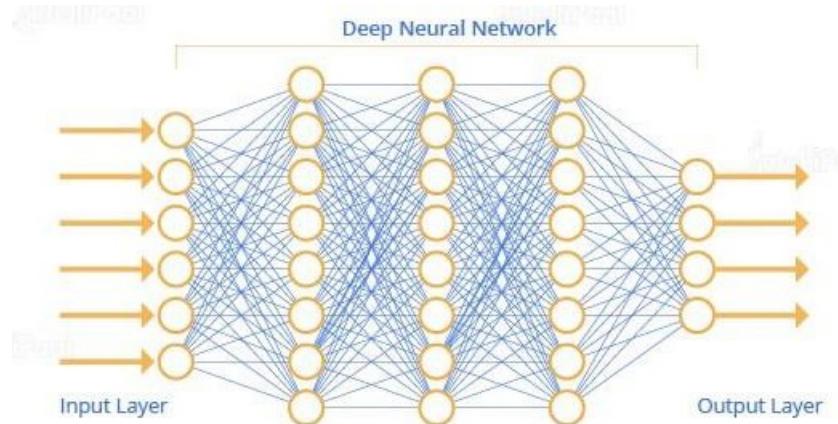
Data mining je považováno za jakýsi supersoubor mnoha různých metod, jak získat vhled či náhled do studovaných dat. Může zahrnovat jak tradiční statistické metody, tak i novější strojové učení. Data mining uplatňuje metody z širokého záběru různých oblastí, aby byl schopen identifikovat vzorce, které předtím nebyly schopen rozpoznat. Často zde bývají zahrnutý statistické algoritmy, strojové učení, analýza textu, analýza časových řad a nespočet jiných oblastí analytiky. Tato metoda mimo jiné zahrnuje studium a zároveň i následnou aplikaci uchovávání a manipulace dat.[22, 41]

Strojové učení má se statistickými modely stejný cíl: porozumět struktuře dat a usadit teoretickou distribuci na data, která jsou dobře pochopená. U statistického modelu existuje teorie, že model je matematicky podložen, to však vyžaduje, aby data také splňovala určité silné předpoklady. Oproti tomu strojové učení bylo vyvinuto na základě schopnosti počítačů prozkoumat data pro jejich strukturu i přes fakt, že neexistuje teorie, jak by struktura daných dat měla vypadat, a to díky jejich možné multidimenzionalitě anebo neúměrné korelací. Testem pro strojové učení je chyba ověření na nových datech, nikoliv teoretický test, který prokáže nulovou hypotézu. Vzhledem k tomu, že strojové učení je používáno téměř výhradně jako iterativní proces pro učení z dat, lze tento proces učení lze jednoduše automatizovat. Jednotlivá opakování jsou pak použita na zpracovávaná data, dokud není dosaženo robustního vzorce.[3, 22]

Deep learning neboli hluboké učení je podsložkou strojového učení, které dokáže vyvinout velkou sílu a flexibilitu za pomocí studování světa způsobem hierarchického vrstvení konceptů a abstrakcí. Toto pole je tedy strojové učení, které bylo inspirováno funkcionalitou mozkových buněk a vytváří takzvané umělé neuronové sítě. Funguje jednoduše tak, že vezme datová spojení mezi jednotlivými umělými neurony a upraví je na základě datového vzorce. Z logiky věci také vyplývá, že je potřeba více neuronů, pokud máme větší objem datových výstupů. Tato metoda automaticky obsahuje učení na mnoha úrovních abstrakce, a tudíž umožňuje systému naučit se komplexní mapování funkcí bez závislosti na jakémkoliv specifickém algoritmu. Hlavním rozdílem tedy mezi klasickým strojovým učením a hlubokým učením je ten, že u dřívějšího musíme manuálně poskytnout určité vlastnosti či funkce do systému, aby mohl fungovat. V hlubokém učení se tyto věci automaticky extrahují pro klasifikaci, což má ale za následek obrovský požadavek na množství dat pro vytrénování deep learning algoritmu a tím pádem přesnost výstupu u této metody je závislá na množství trénovacích dat.[3, 25, 26]

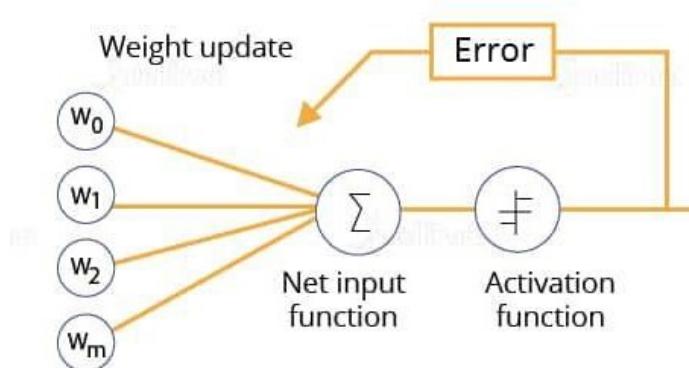
Pro lepší pochopení fungování hlubokého učení je vhodné rozebrat strukturu a princip umělých neuronových sítí. Ty jsou složeny ze tří vrstev. První vrstva je takzvaná vrstva vstupu. Ta je využívána pro příjem vstupních dat z externích zdrojů a jejich následného předávání do skryté vrstvy neuronové sítě. Tato vrstva neprovádí žádné výpočty. Druhou vrstvou je skrytá vrstva, která se skládá buď z jedné anebo z několika skrytých vrstev a tvoří hlavní výpočetní jednotku s určitým druhém matematické funkce. Po tom, co jsou veškeré výpočty hotovy,

výstupy z nich jsou přesunuty do výstupní vrstvy. Ve výstupní vrstvě jsou prováděny finální výpočty, a především je zde předáván výstup do vnějšího světa.[3, 25, 26, 42]



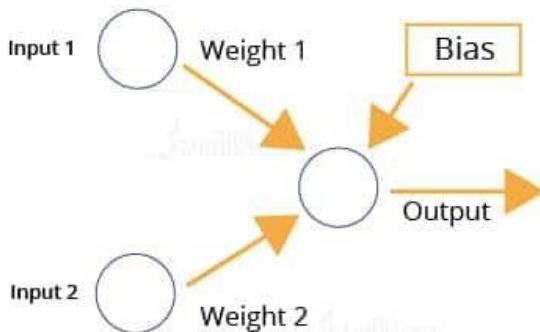
Obrázek 10 Příkladná struktura umělé neuronové sítě. V překladu: Deep neural network – Hluboká neuronová síť; Input layer – vstupní vrstva; Output layer – výstupní vrstva [25]

Schéma potenciální struktury ukazuje Obrázek 10, kde vlevo můžeme vidět uzly vstupní (input) vrstvy, uprostřed jsou jednotlivé vrstvy uzlů skryté vrstvy a napravo pak uzly výstupní (output) vrstvy. Jednotlivé uzly jsou navzájem propjeny spojením, přes které spolu interagují. Tyto spoje jsou asociovány s reálným číslem, které se nazývá váha spojení. Váhy jsou nejdříve inicializovány s náhodnou hodnotou, a tudíž zde může být velký rozdíl mezi skutečnými hodnotami a těmi odhadovanými. Vlivem toho je vyžadováno více iterací pro zvětšení robustnosti.[3, 25, 26, 42]



Obrázek 11 Fungování iterací upravování vah. V překladu: Weight update – Úprava vah; Error – Chyba; Net input function – Funkce součtu vstupů; Activation function – Přenosová funkce [25]

V Obrázku 11 je znázorněna jedna iterace úpravy váhy. Pokud i přes přiřazení nové váhy a provedení výpočtu v přenosové funkci (activation function) není výsledek tím požadovaným, tak se vrací zpět na začátek a upraví se opět váha spojení pro dosažení optimálního výsledku. Takto postupně pokračuje, dokud nedosáhne nejlepšího možného výsledku. Mimo jiné váhy spojení taktéž určují, jak rychle se spustí přenosová funkce a obecně je lze označit jako ukazatele důležitosti daného vstupu do funkce.[3, 25, 26, 42]

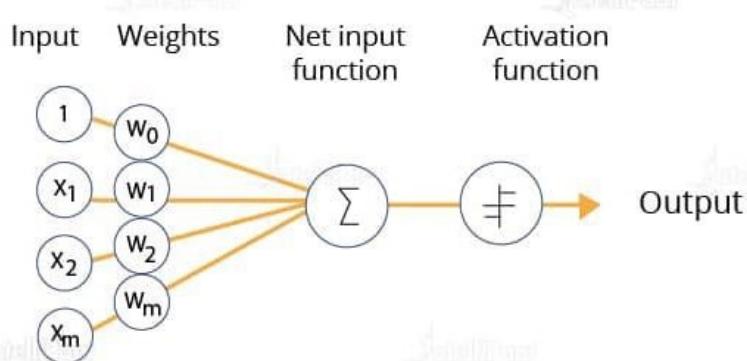


Obrázek 12 Schéma všech veličin pro výpočet výstupu. V překladu: Input – Vstup; Weight – Váha; Bias – Zaujatost; Output – Výstup [25]

Společně s hodnotami vstupů a hodnotami jednotlivých vah spojení se využívá ještě jedné veličiny, a to zaujatosti (bias), která je prakticky konstantou pro vyrovnání výsledku (může být jak záporná, tak kladná). Pokud bychom chtěli z Obrázku 12 vytvořit matematickou formulaci, jak zjistit hodnotu výstupu, mohla by vypadat jako rovnice (1).[3, 25, 26, 42]

$$VÝSTUP = \sum (\text{váha spojení} * \text{vstup}) + \text{zaujatost} \quad (1)$$

Samozřejmě ve skutečném světě bývají data zpravidla 3dimenzionální a často mají mnoho vlastností, které je třeba brát v potaz (viz Obrázek 13). Příkladem toho může být rozpoznání obrázku auta. Samotné auto lze promítnout do 2-D prostoru (výška, délka), nicméně se s objektem auta jako takovým pojí mnoho dalších parametrů, a proto se tento výpočetní úkol vyvíjí v poměrně komplexní situaci. Abychom snížili tuto komplexitu, využijeme právě již dříve uvádzovaných přenosových funkcí. Mezi nejčastěji používané funkce patří lineární přenosová funkce, sigmoidální přenosová funkce, přenosová funkce hyperbolické tangenty a tzv. RELU funkce.[3, 25, 26]



Obrázek 13 Schéma exemplárního reálného vstupu do funkce [25]

2.3.3 Face recognition system (systém rozpoznání obličeje)

Jedna z definic označuje pojem systém rozpoznání obličeje jako technologii, která je schopna identifikovat a verifikovat osobu z digitálního obrázku či ze snímku videa. Tato definice je poměrně obecná a neříká nám příliš mnoho informací o tom, jak takový systém doopravdy funguje. V této kapitole tedy bude představeno, co doopravdy je technologie rozpoznání obličeje a následně bude uveden stručný, a především obecný průběh celé této metody. Není možné specifikovat přesné fungování, jelikož možností, jak provádět rozpoznání obličeje je mnoho a každý má svá specifika. Nicméně obecný vzor průběhu zůstává stejný.[3, 12, 28]

Rozpoznání obličeje můžeme chápat jako biometrickou softwarovou aplikaci, která je schopna unikátní identifikace a verifikace identity dané osoby za pomocí kontrastování a analýzy vzorů založených na obličejobvých rysech osoby. Obecně tedy vytváří mnoho šablon cíleného individuálního obličeje a následně je porovnává s již existujícími obrázky známých obličejů ve vytvořené databázi. S přílivem novějších technologií tato metoda bývá také označována jako biometrická aplikace založená na umělé inteligenci, protože AI bývá v těchto případech hojně využíváno.[3, 11, 12, 28]

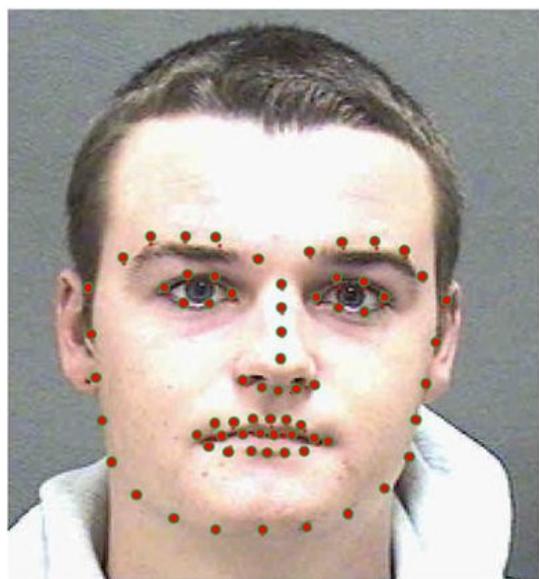
Mezi základní obory, na kterých tato metoda staví a které přispívají k jejímu vývoji, patří strojové učení, rozpoznávání vzorů, obličejobvá analýza a hluboké učení. Jak již bylo v dřívějších kapitolách rozebráno, hluboké učení je metoda, jak vytvořit umělou inteligenci, která by byla schopna se sama učit dělat správná rozhodnutí a hledat řešení na dané problémě, tudíž se jedná o inspiraci, která čerpá z fungování rozhodování lidského mozku. V kontextu rozpoznání obličejů tedy mluvíme o schopnosti lidí umět rozpoznat obličeji a stejně tak si jej

umět zařadit do určitého kontextu či souvislostí, které jsou v mozku jedince již založeny.[3, 11, 28]

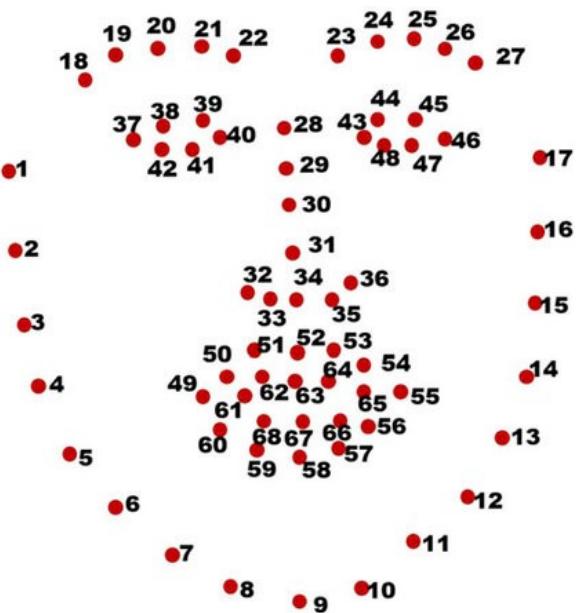
V praxi rozpoznáváme osoby a jejich obličeje na základě určitých tvarů a markantů, kterými disponují. Pokud tento atribut převedeme do technologické roviny, stává se nám z toho systém rozpoznání obličeje, který ale funguje na mnohem systematičejším, a především více matematickém základě a procesech, které jsou následně rozděleny na základy algoritmů.[11, 12, 28]

První fází bývá obecná detekce obličeje, případně více obličejů, v obrázku. Jedná se o jejich pouhou lokalizaci ve snímku. Druhá fáze je obličejobá analýza, kdy software přečte geometrii obličeje dané osoby. Zde se detekují klíčové faktory jako orientační či uzlové body (viz Obrázek 14), do kterých spadá například vzdálenost mezi očima osoby nebo vzdálenost od čela po bradu apod. Jakmile jsou tyto obličejobá orientační body identifikovány, jsou následně použity pro vytvoření tzv. podpisu obličeje. Fáze třetí je převedení obrázku na data. Jakmile je analýza obličeje hotová, je převedena na tzv. obličejobá podpis, což lze chápat jako určitý matematický vzorec. Tento podpis se skládá z kódu a je následně adresován jako otisk obličeje podobně, jako je označován například otisk prstu. Poslední, čtvrtou fází je nalezení shody. Nyní již dříve vytvořený numerický kód je porovnán se známou databází obličejobá otisků. Tato databáze by měla být tvořena kolekcí obrázků s jejich identifikací, aby mohla být provedena jejich komparace.[3, 12, 28]

Pokud vezmeme v potaz kontext práce, a především obor jejího zaměření, jeví se jako vhodné také zmínit i nemalé problémy, jež se pojí s užíváním technologie rozeznávání obličejů. Jako první problém jistě každému vyvstane soukromí. Pro užití této technologie je nutné sesbírat a následně i uchovávat data o dané osobě, což se jeví jako značné bezpečnostní riziko v kontextu nynější právní úpravy. Velké datové úniky začínají být častým problémem dnešní doby a bohužel ani tato technologie není výjimkou pro tento typ útoku. Dalším problémem může být špatné vyhodnocení. Rozpoznávací software porovnává vzorky s těmi, které má uložené v databázi, nicméně ne vždy se tento proces může zdát tak jednoduchý. Problemy a chyby se mohou vynořit kvůli nevhodné kvalitě snímaného obrazu a v důsledku chybného vyhodnocení obličejobá markantů. Tento fakt může vyústit ve špatné vyhodnocení, a tudíž vytvářet další nepříjemné situace a bezpečnostní hrozby.[9, 10, 11, 12, 28]



(a)



(b)

Obrázek 14 Příklad z knihovny dlib - (a) označené obličejové markanty, (b) rozmištění a očíslování jednotlivých bodů [27]

3 Použité prostředky

V této kapitole budou představeny veškeré prostředky použité pro dosažení cílů práce a provedení simulačních pokusů pro ověření efektivity rozpoznávacích algoritmů.

3.1 Použitý hardware

Tato podkapitola je věnována veškerému hardwaru, který byl klíčový pro realizaci této práce. Je zde představen jednodeskový počítač Raspberry Pi 4, kamera Waveshare RPi IR-CUT a v poslední části pak stojan, který byl vyroben pro účely testování kamerového systému a rozpoznávacího algoritmu.

3.1.1 Raspberry Pi 4 Model B

Tento jednodeskový počítač již byl představen v jedné z předcházejících kapitol, proto zde nebude příliš rozebíráno. Spíše bude předveden model, který byl použit pro tuto práci a k němu přidané komponenty.



Obrázek 15 Fotografie Raspberry Pi 4 s pasivním chladičem [autor]

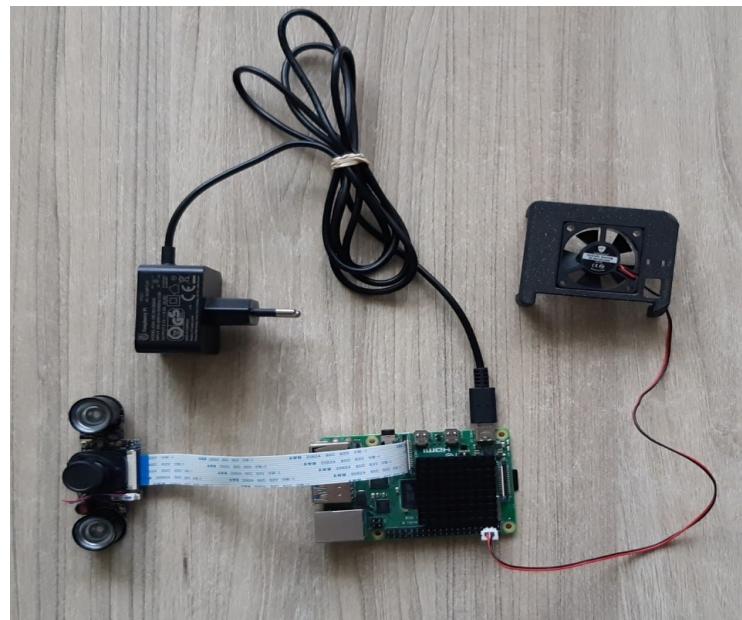
Model výše uvedeného počítače byl zvolen s 8 GB RAM pamětí pro zajištění co nejplynulejšího chodu a rezervní výpočetní paměti v případě náročnějších operací. Počítač je koncipovaný jako výpočetní jádro celého kamerového systému. Tento minipočítač je 56,5 mm vysoký, 85,6 mm dlouhý a váží kolem 46 g bez doplňků. Je osazen 64 gigabyte (GB) secure digital (SD) kartou, na které je Raspbian OS, čímž je zajištěna identická funkcionality, jakou lze znát z běžných desktopů či laptopů. Přístup do systému je podmíněn přihlášením za pomocí uživatelského jména s heslem. To platí jak pro přímý, tak i vzdálený přístup přes internetovou síť. Vzhledem k náročnosti prováděných operací bylo nainstalováno pasivní chlazení (které lze

vidět na Obrázku 15) společně s aktivním chlazením ve formě větráku (které lze vidět na Obrázku 16), který je připojen a kontrolován pomocí GPIO pinů. S kamerou je počítač propojen pomocí speciálního konektoru MIPI CSI, určeného pro připojení kamery na základní desku. Navíc je ještě kamera propojená drátovou propojkou s jedním GPIO pinem pro umožnění kontroly nad nočním režimem kamery. Na piny je také připojeno aktivní chlazení v podobě větráku se dvěma kably. Černý je propojen se zemí a červený k jednou ze dvou možností napájení. Větrák může pracovat v jednom ze dvou režimů – buď při sníženém výkonu při napojení na 3,3 V, kdy je větrák tišší, a nebo na plný výkon při připojení na 5 V.[19]

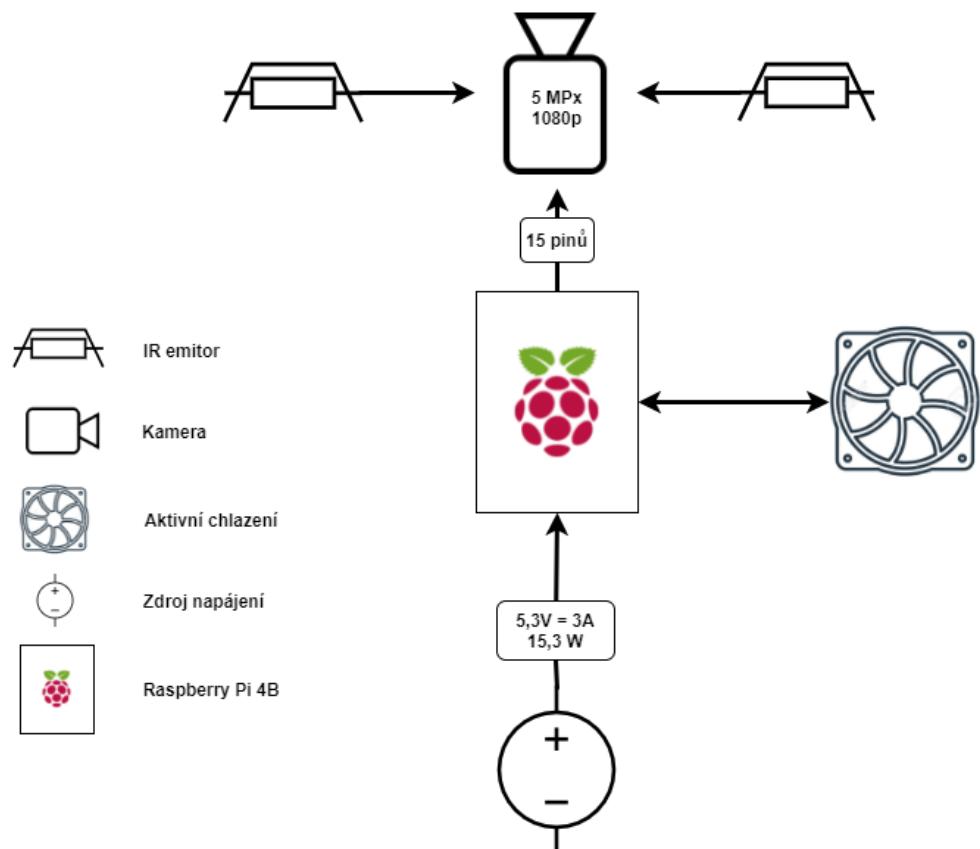


Obrázek 16 Aktivní chladící zařízení pro Raspberry Pi [autor]

Celé schéma zapojení jednotlivých prvků lze vidět v Obrázku 18, na Obrázku 17 lze potom vidět skutečné zapojení komponentů.



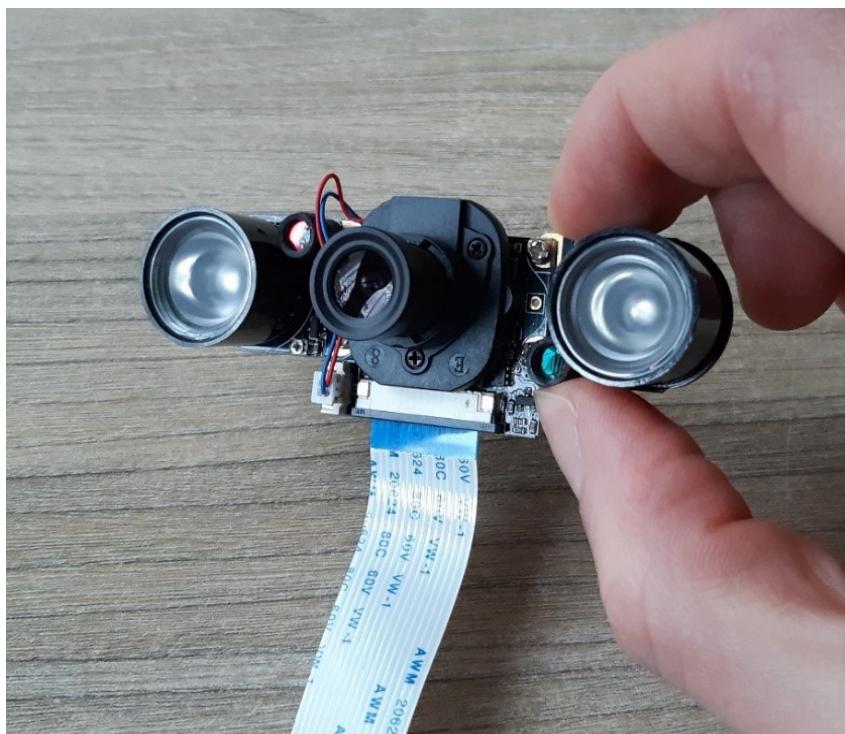
Obrázek 17 Fotografie zapojení jednotlivých komponentů [autor]



Obrázek 18 Schéma zapojení kamerového systému včetně legendy nalevo [autor]

3.1.2 Waveshare RPi IR-CUT kamera

Jako snímací zařízení byla pro kamerový systém vybrána právě tato kamera od firmy Waveshare (lze vidět na Obrázku 19). Jedná se o kameru, která podporuje všechna zařízení Raspberry Pi, ikdyž se nejedná o stejného výrobce. Kamera je koncipována jak do denního, tak i nočního prostředí, disponuje totiž vestavěným a odnímatelným IR-CUT filtrem, který eliminuje zkreslení barev za denního světla. Kamera pak sama sleduje dostatečné osvětlení prostředí a v případě snížené viditelnosti přepne na noční snímání obrazu ve formě infračerveného záření. Součástí kamery jsou také dvě infračervené LED, které emitují požadované záření a umožňují fungování i v absolutní tmě. Kamera je osazena 5megapixelovým snímačem OV5647 společně s možností manuálně nastavitelné vzdálenosti ostření.[29]



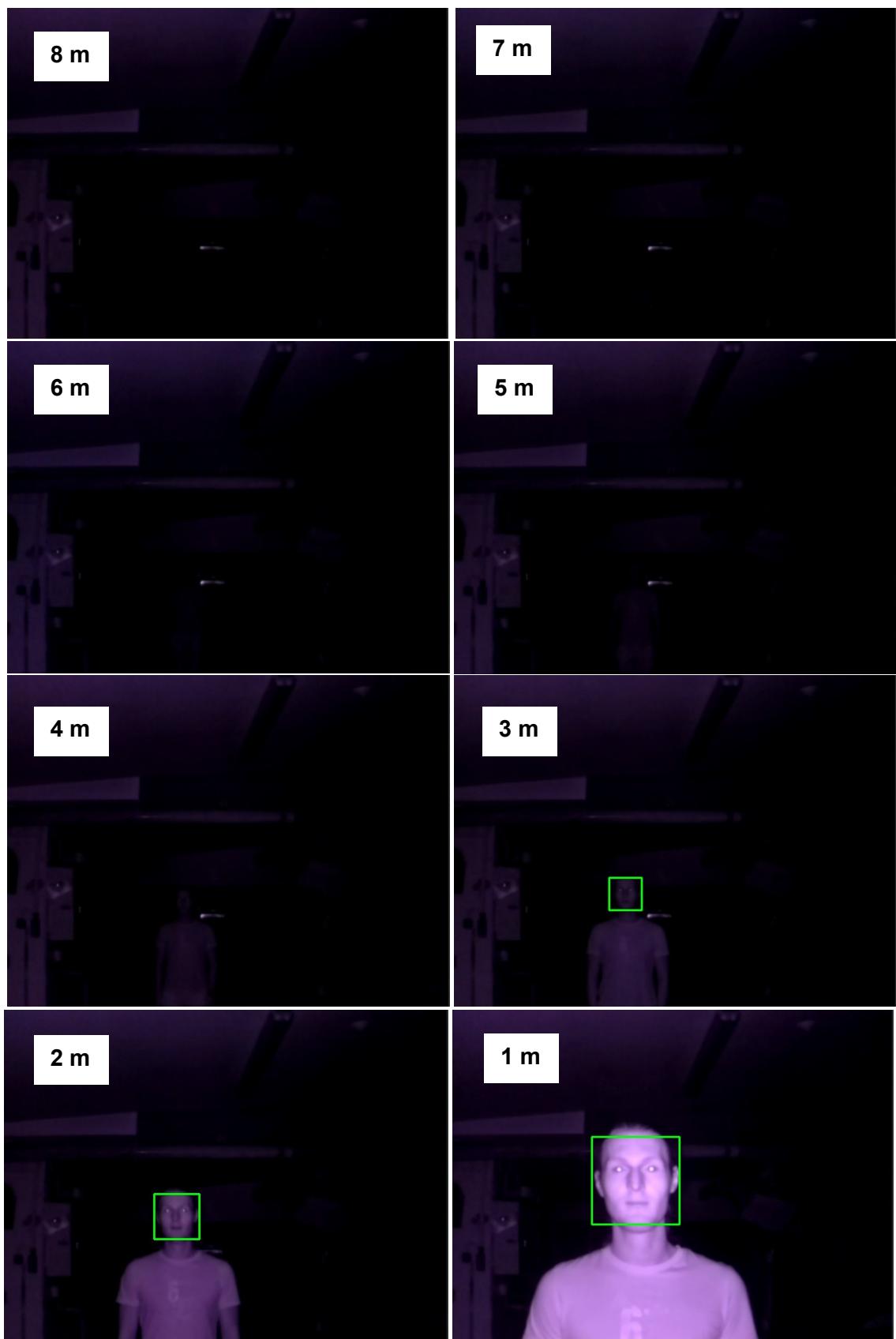
Obrázek 19 Kamera Waveshare RPi IR-CUT [autor]

Rozměry samotného kamerového modulu jsou 31 mm výšky na 32 mm šířky, po připojení přídavných LED komponentů se šířka kamery zvětší na 80 mm. Kamera je propojena s počítačem 15pinovým plochým kabelem přes rozhraní CSI. Důvodem, proč byl tento model kamery preferován oproti oficiálnímu kamerovému modulu od firmy Raspberry je absence pomocného příslivku poskytovaného ke kameře. Výrobce Raspberry poskytuje takéž kameru s IR-CUT filtrem, nicméně je třeba dodat podporu příslivku z jiného zařízení. Waveshare kamera

oproti tomu byla dodána již s IR LED, které lze jednoduše napojit na kamerový modul a není tedy třeba složité konstrukce přísvitu.[29]

Pro zjištění možných podmínek pro noční testování a jeho proveditelnost byly provedeny předem testy viditelnosti v nočním režimu a následně vytvořena série fotografií, ilustrující dosah infračerveného přísvitu poskytovaného kamerou. Subjekt začínal ve vzdálenosti 8 metrů od kamery a na každém celém metru byl proveden záznam snímku. Tedy byla pořízena fotografie na osmém metru, dále na sedmém, šestém atd. Konečná fotografie byla pak zaznamenána ve vzdálenosti jednoho metru od kamery. Celý soubor fotografií lze vidět na Obrázku 20.

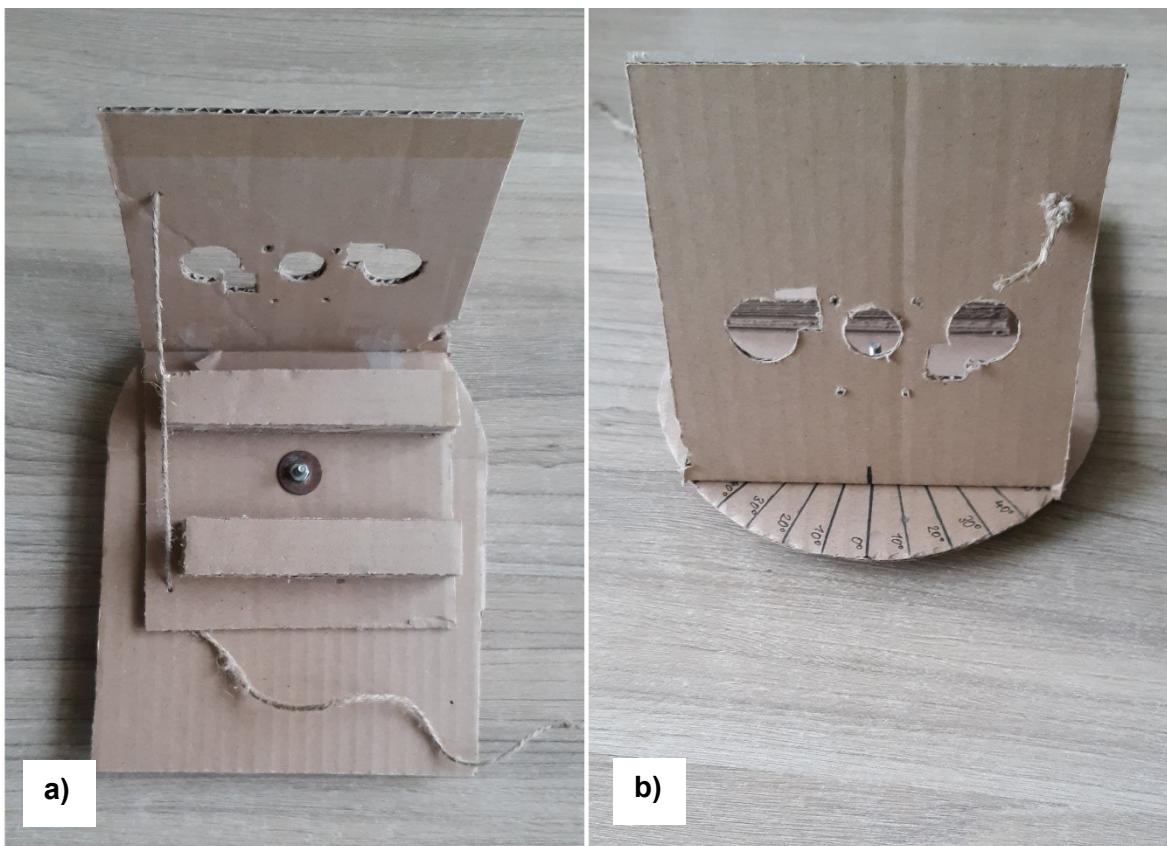
Závěrem nočních přípravných testů je efektivní vzdálenost, ve které bude následné noční testování osob prováděno. Tato hodnota byla stanovena na vzdálenost 4 metrů od kamerového systému. V rámci přípravných testů byly také provedeny testy na detekci osob, které ukázaly, že tato detekce je natolik efektivní, že vytyčený testovací prostor nebyl dostatečný pro zjištění vzdálenosti, kde již algoritmus nebyl schopný osobu jako objekt detektovat.



Obrázek 20 Výsledky testů vzdálenosti detekce obličeje v noci [autor]

3.1.3 Kamerový stojan

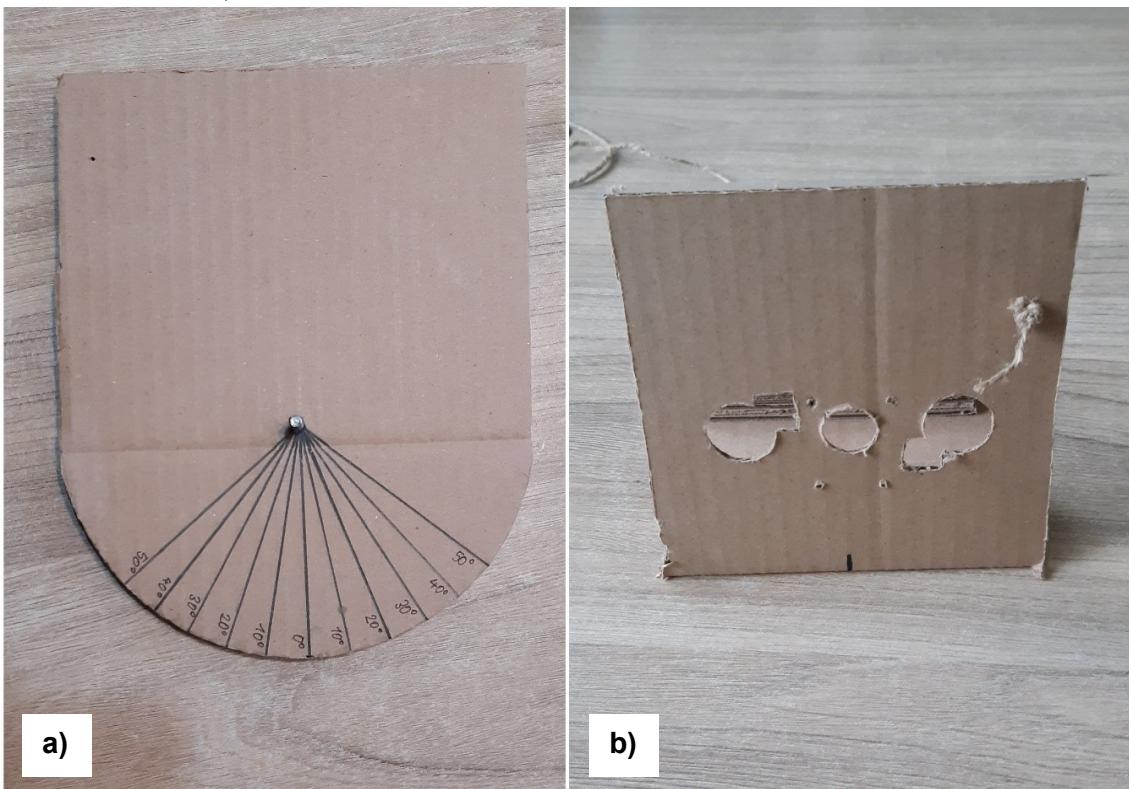
Pro účely testování byl vyroben a zkompletován polohovatelný stojan pro kamerové zařízení. Požadavkem byla schopnost polohování, které by vyhovovalo požadavkům měření. Měl by tedy být schopen určité aracetace okolo osy a také mít schopnost záklonu. Pro výrobu byl zvolen kartonový papír jakožto nejjednodušší možný materiál na zpracování a formování konstrukce stojanu. Zhotovený stojan lze pak vidět na Obrázku 21. Jedná se o pouhý prototyp, určený čistě pro účely testování. Pro užití v praxi by bylo vhodné vytvoření praktičejšího modelu v počítačové grafice a následný tisk na 3D tiskárně.



Obrázek 21 Stojan pro kameru – a) zadní pohled, b) čelní pohled [autor]

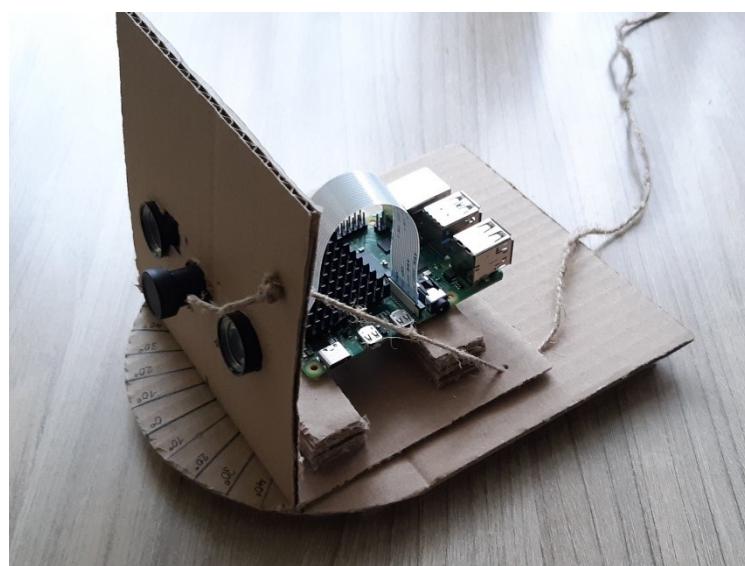
Experimentální stojan je tvořen dvěma komponenty: držákem na kameru (viz Obrázek 22 část a) a pevnou základnou (viz Obrázek 22 část b), se kterou je držák spojen šroubem a maticemi a na níž je promítnuto měřítko úhlů pro pozdější zajištění vytočení samotné kamery. V čele držáku byly vyříznuty otvory jak pro samotný objektiv, tak na IR LED a články snímající osvětlení prostoru. Dále byla na spodní části čela (jak lze vidět v Obrázku 21 části b) vytvořena značka symbolizující přesný střed čočky kamery, čelo držáku bylo rovněž ukotveno k jeho

spodní části pomocí lanka, které má zajistit možnost aretace v záklonu či předklonu (viditelné v Obrázku 21 části a).



Obrázek 22 Fotografie základny (a) a držáku na kameru (a) [autor]

Úhel, pod kterým je čelo umístěno, je pak nutno měřit externím nástrojem, jako například úhloměrem. Plně osazený stojan lze vidět v Obrázku 23.



Obrázek 23 Plně osazený stojan na kameru [autor]

3.2 Použitý software

Tato podkapitola je věnována softwarovým prvkům, které byly kritické pro sestrojení skriptů použitých pro cíl práce. Nejdříve je stručně představen využitý programovací jazyk Python společně s uvedením knihoven, které tvoří základní stavební kameny fungování skriptů. Následuje deskripce fungování použitých algoritmů, mezi které spadá histogram orientovaných gradientů (HOG), konvoluční neuronová síť (v angličtině convolution neural network, dále jako CNN) či Haar kaskádující klasifikátor, a v závěru jsou představeny veškeré skripty použité v práci a též je diskutován princip jejich fungování.

3.2.1 Python

Pokud bychom chtěli některý z aktuálně využívaných programovacích jazyků označit za vycházející hvězdu, tak jednoznačným kandidátem na tento titul bude Python. Je to jazyk, který je využívaný opravdu na všech pracovních místech, v každém odvětví, v mnoha firmách a projektech, ať už amatérských, revolučních či velikých projektech jako raketoplány do vesmíru. Je používán profesionálními softwarovými vývojáři, i lidmi, kteří jej využívají pouze jako pomocný nástroj ke své jinak nevývojářské činnosti.[31]

Python je vysokoúrovňový programovací jazyk a byl vytvořen Guidem van Rossumem. Jeho tvůrce klade velký důraz na jednoduchou čitelnost kódu, což se projevilo i v designu Pythonu a jeho filozofii. Tento jazyk je poměrně přívětivý i co se týče programovacích paradigm a podporuje jich hned několik: od imperativního a procedurálního, přes funkcionální, až po v dnešní době populární objektově orientovaného programování, které bez debat přispívá k jeho snadnému užívání a míry komplexnosti osvojení. Python je dynamicky typovaný jazyk a má v sobě zahrnutý garbage collector. Často je označován jako jazyk, který má „baterie v balení“, jelikož jeho standardní knihovna je opravdu široká a podporuje mnoho funkcionalit.[30, 31]

Python je běžně využíván pro vývoj webových stránek a softwaru, automatizaci úkolů, datové analýzy a datovou vizualizaci. Jelikož se tento jazyk dá poměrně lehce naučit, byl rychle osvojen lidmi bez programovacích znalostí, jako jsou účetní nebo vědečtí pracovníci, pro plnění každodenních úkonů, jako například správa finančního sektoru apod. Nejčastější využití:[31]

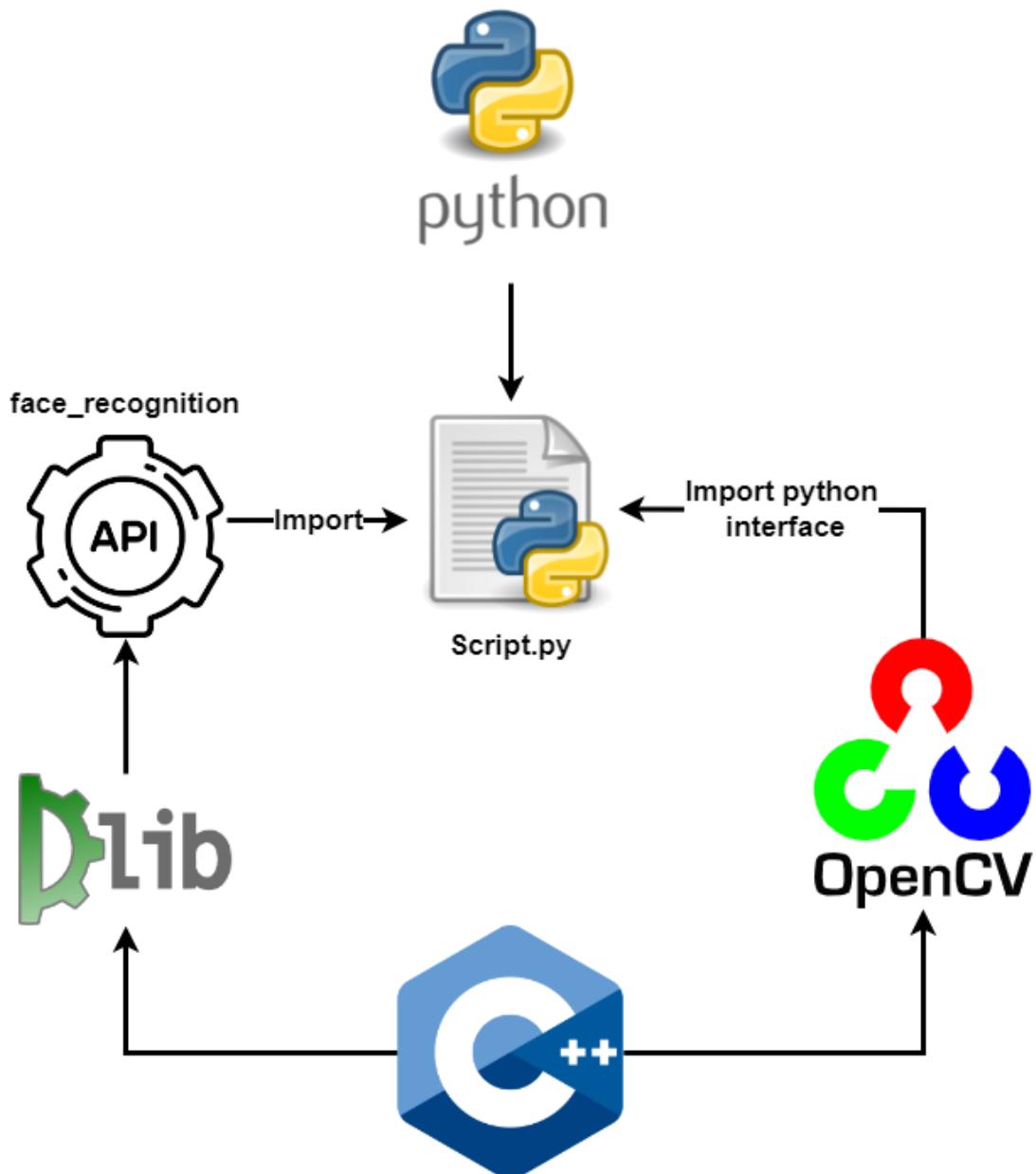
- webový vývoj,
- softwarové testování a prototypování,
- obecná automatizace a skriptování,
- datová analýza a strojové učení,
- každodenní úkony.

Právě pro tyto jeho přednosti byl Python vybrán jako jazyk vhodný pro cíl práce. Jeho jednoduchý syntax je poměrně snadno pochopitelný i pro neznalé programátory, je opravdu všeobecný a lze jej využít pro mnoho úkolů. Je přívětivý pro začátečníky, je open source, což z něj činí ideální případ pro vědecké, školní ale i komerční projekty, má obrovskou a náromocnou komunitu a taktéž jeho velký archiv modulů a knihoven byl velkým kritériem, které hrálo roli při jeho výběru. Následně budou představeny ty nejdůležitější knihovny, které byly použity jako základní stavební kámen pro softwarovou část práce.

Při pohledu do schématu v Obrázku 24 lze vidět ty nejdůležitější knihovny a moduly, které jsou kritické pro dosažení stanoveného cíle. Jak je vidět, nejsou tyto komponenty psány v jazyce Python, nýbrž v C++. Ačkoliv byl zvolen jako cílený jazyk Python, není tento fakt problém. Jelikož jsou tyto knihovny široce využívané a velice populární, jejich kompatibilita je zajištěna pomocí tzv. application programming interface (API), neboli programovacích rozhraní, které umožňují provádět volání funkcí z těchto knihoven v požadovaném jazyce. Co se týče knihovny Open Source Computer Vision Library (OpenCV), tak kompatibilita s Pythonem je již zabudována v knihovně samotné, respektive již obsahuje API. Pro knihovnu Dlib byl pak vytvořen balíček přímo pro její funkcionality rozpoznávání obličeje, mající příznačný název Face Recognition. Pomocí tohoto balíčku lze taktéž provádět volání na určité funkce knihovny.

Dlib je moderní balíček nástrojů a algoritmu pro vytváření komplexního softwaru a jeho následné využití pro řešení problémů v životě. Je využíván jak v průmyslu, tak i v akademické sféře v různých polích expertíz, zahrnující mobilní telefony, robotiku, vestavěná zařízení a nebo velké, vysokovýkonné výpočetní prostředí. Díky open source licenci je Dlib volně a zdarma k využití pro jakýkoliv projekt. Dlib byl vyvinut Davisem E. Kingem v jazyce C++ a je silně inspirován vývojem „design by contract“. Tato inspirace jde vidět v jeho balíčku nástrojů, který obsahuje komponenty pro data mining, networking, vlákna, uživatelská grafická rozhraní,

datové struktury, lineární algebru, numerické optimalizace, Bayesovské sítě, ale hlavně pro tuto práci důležitý image processing se strojovým učením, na které se v posledních letech tento balíček soustředil.[32, 33]



Obrázek 24 Struktura nejdůležitějších knihoven při stavbě skriptu [autor]

V návaznosti na Dlib je nutné představit knihovnu Face Recognition, která hraje velkou roli při používání Dlib nástrojů pro obličejobou rekognici. Jedná se o jednoduché aplikační programovací rozhraní, které umožňuje volání funkcí a komunikaci s kódem psaným v C++. Tato knihovna disponuje hned několika funkcemi, mezi které patří nalezení všech obličejů

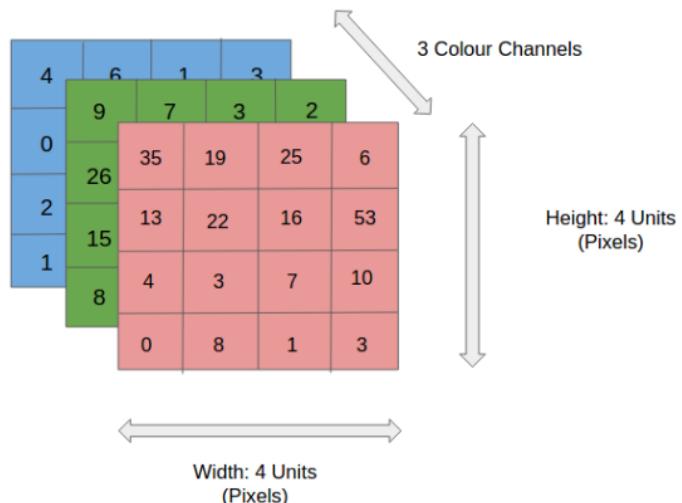
v poskytnutém obrázku, nalezení a manipulace obličejových rysů a pro cíl této práce nejdůležitější funkce rozpoznání obličejů v předaném snímku.[34]

Poslední knihovnou zobrazenou ve schématu je OpenCV. Jedná se tedy o open source knihovnu, jejíž hlavním účelem a zaměřením je počítačová vize společně se strojovým učením. Cílem jejího sestrojení bylo poskytnutí infrastruktury pro aplikace mající za cíl počítačovou vizi a zrychlit využívání strojového vnímání v komerčních produktech. Tato knihovna je využívána jak firmami, státními službami, tak výzkumnými a akademickými institucemi. Obsahuje přes 2500 optimalizovaných algoritmů, mezi nimiž lze nalézt jak klasické, tak moderní algoritmy pro počítačovou vizi a strojové učení. Mezi funkce, které lze díky této knihovně využít patří například sledování pohybujících se objektů, extrakce 3D modelů objektů, sledování pohybu kamery, nalezení podobných obrázků v databázi obrázků, odstranění červených očí či sledování pohybu očí, detekce a rozpoznání obličejů, obecné nalezení objektů a klasifikace lidských činností ve videu. Ačkoliv je OpenCV psán nativně v C++, tak disponuje rozhraním pro jazyk Python, tudíž je jeho zakomponování do skriptů vytvořených v rámci práce poměrně jednoduché. [35]

3.2.2 Konvoluční neuronová síť

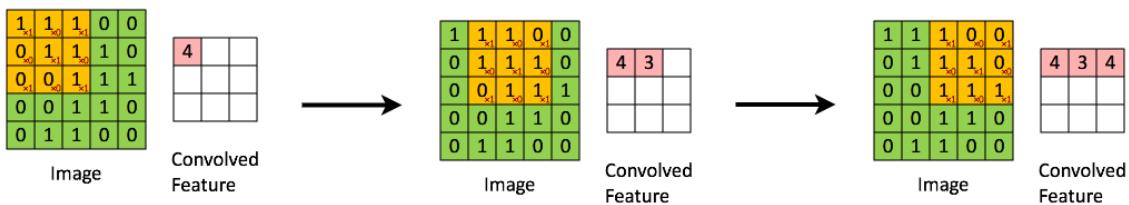
CNN (někdy také ConvNet), neboli česky konvoluční neuronová síť, je algoritmus založený na základě hlubokého učení, který je schopen představený obrázek zpracovat, tedy vytvořit váhy a zaujatosti pro různé aspekty a objekty v obrázku, a následně je schopen jej rozlišit od jiných možných obrázků. Předzpracování, které vyžaduje CNN je mnohem nižší, pokud jej srovnáme s ostatními klasifikačními algoritmy. Ačkoliv u primitivních metod jsou filtry vytvářeny ručně, pokud poskytneme CNN dostatečné trénování, má pak schopnost se tyto filtry či charakteristiky sama naučit. Architektura, kterou je CNN navržena, je silně inspirována, téměř lze říci, že mimikuje, konektivitou shluků neuronů v lidském mozku a lze vidět motivaci v mozkové kůře, která má na starost vizuální zpracování podnětů.[36, 43]

Pokud se podíváme na obrázek v počítači, tak v jádru je tento obrázek reprezentován jako soubor binárních hodnot, tedy jedniček a nul. To vede k nápadu zploštění tohoto obrázku do vektoru, který se dá následně použít pro jednoduché klasifikátory. Tento postup se může jevit jako validní u méně náročných obrazů, ale jakmile je zpracováván obrázek mnohem komplexnější, tato metoda selhává se svou hodnotou přesnosti klasifikace. Problémem je právě závislost jednotlivých pixelů a určitá hloubka obrazu. CNN je schopna zachytit prostorové a časové závislosti v obrázku skrze aplikaci různých relevantních filtrů. Jinými slovy lze říci, že tato neuronová síť je schopná být vytrénována, aby pochopila sofistikovanost představeného obrázku.[36, 43]



Obrázek 25 Příklad rozboru obrázku na jeho barevné vrstvy. V překladu: Colour Channels – Barevné kanály; Height – Výška; Units – Jednotky; Pixels – Pixely [36]

V Obrázku 25 vidíme příklad toho, jak by mohl být vstupní obrázek rozložen na jeho jednotlivé dimenze. Jedná se o obrázek 4 pixely na výšku a šířku a o třech barevných kanálech. Je to klasický RGB obrázek, který byl pouze rozvrstven na příslušné barvy. Samozřejmě existuje mnoho dalších barevných formátů jako BRG, HSV, stupně šedi apod. Lze si představit, jak výpočetně náročný by byl obrázek ve vyšší kvalitě, například 8K (to odpovídá zhruba 7680 x 4320 pixelů). Pro tento účel je sestrojena právě CNN, která má za úkol zredukování obrázku do takové formy, kterou je již pak jednoduché zpracovat, ale při stálém zachování vlastností, které jsou kritické pro zaručení správné predikce.[36, 43]

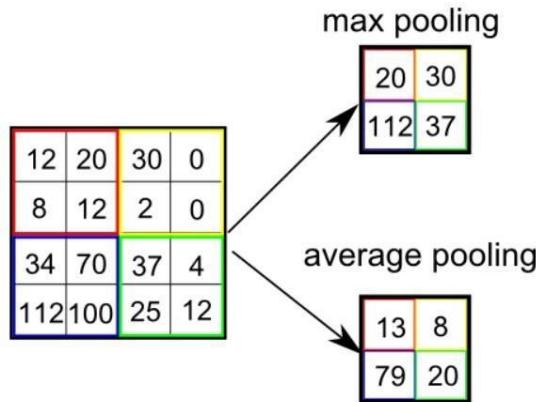


Obrázek 26 Konvoluce obrázku. V překladu: Image – Obrázek; Covolved Feature – Konvolvovaná vlastnost [36]

V Obrázku 26 lze vidět příkladnou konvoluci jednoho barevného kanálu obrázku. Obrázek je zde reprezentován zelenou tabulkou 5x5, element, který vykonává konvoluci je nazýván jádro nebo filtr a je reprezentován žlutou částí tabulky. Jeho velikost je 3x3. Filtr pokaždé provede násobení matice čísel, které zrovna označuje, výsledek se zapíše do tabulky konvolucí a posune se v tabulce o 1 doprava, v případě okraje tabulky pak zpět na začátek a o řádek níže. Takto pokračuje, dokud není celý obrázek zpracovaný. Smyslem této konvoluční operace je extrakce vysokoúrovňových vlastností z obrazu. CNN téměř nikdy není limitován na pouze jednu konvoluční vrstvu. Nejčastěji se právě první vrstva využívá pro zachycení nízkoúrovňových vlastností jako hrany, barvy, orientace gradientu apod. S přibývajícími vrstvami se architektura více adaptuje na vysokoúrovňové vlastnosti, což následně poskytuje souhrnnější chápání obrázku, stejně jako tomu je i u živočichů.[36, 43]

Z této operace pak mohou vyjít dva druhy výsledků. Jeden je, kdy konvoluce vlastnosti zredukuje její dimenze v porovnání s původním vstupem (lze vidět v Obrázku 26). Druhým pak je zachování dimenzí tak, jaké byly na vstupu. Toho se dosahuje aplikací tzv. validní (v případě redukce dimenzí) a nebo stejné (v případě zachování dimenzí) vycpávky.[36, 43]

Kromě konvoluční vrstvy existuje společně s ní ještě sdružovací vrstva. Ta je zodpovědná za redukci prostorové velikosti vlastnosti, která již prošla konvolucí. To má za cíl snížení výpočetní náročnosti, která je vyžadována pro zpracování dat skrze dimenzionální redukci. Navíc je to užitečná funkce, jak vytáhnout dominantní vlastnosti, které jsou rotačně a pozicičně invariantní. Existují dva druhy sdružování: maximální a průměrné. Maximální sdružení vrací maximální hodnotu z oblasti, kterou zrovna překrývá filtr a průměrné naopak vrací průměr překrytých hodnot. Názornou ukázkou lze vidět v Obrázku 27.[36, 43]



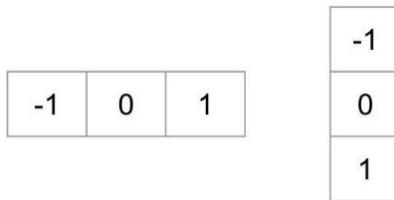
Obrázek 27 Exemplární výsledky sdružovacích metod. V překladu: Max pooling – maximální sdružení; Average pooling – průměrové sdružení [36]

Konvoluční a sdružovací vrstvy dohromady tvoří N-tou vrstvu celé CNN. V kontextu komplexity studovaných obrázků se pak určuje množství těchto vrstev, aby bylo zajištěno zachycení všech nízkoúrovňových detailů. Nicméně je třeba nutné počítat s tím, že čím více vrstev vytvoříme, tím náročnější bude požadavek na výpočetní sílu. Po provedení výše popsaných procesů je model schopen plně pochopit jednotlivé vlastnosti obrázku. Nyní již jen zbývá výstup zploštit a předat pro klasifikační účely.[36, 43]

3.2.3 Histogram orientovaných gradientů

HOG je deskriptor hojně využívaný v počítačovém vidění a při zpracování obrázků za účelem detekce objektů. Tato technika počítá výskyty orientace gradientů v lokalizovaných částech obrázku. Pod pojmem deskriptor si lze představit reprezentaci obrázku, která jej simplifikuje a zároveň provede extrakci užitečných informací a eliminuje ty vedlejší. Pojmy užitečné a vedlejší jsou zavádějící, pokud nejsou uvedeny v kontext situace. V tomto případě považujeme za užitečný výsledný vektor, který nám tento deskriptor je schopen na konci operace poskytnout. Tento vektor sice pro nás není užitečný za účelem prohlížení obrázku, ale je velmi užitečný pro účely rozpoznání obrazu a detekci objektů, jelikož může být později použit pro konkrétní klasifikátor s jasným účelem. Je taktéž vhodné stanovit, co jsou pro nás užitečné vlastnosti HOG deskriptoru. U tohoto deskriptoru je distribuce směrů jednotlivých gradientů pro nás tou užitečnou vlastností. Gradienty obrázku jsou užitečné z toho důvodu, že jejich velikost je vyšší okolo hran a rohů, kde jsou oblasti náhlé změny intenzity. S tím se pojí fakt, že právě hrany a rohy v obrázku pro nás nesou mnohem více informací o zkoumaném tvaru objektu než ploché oblasti.[37, 44]

Pro získání histogramu orientovaných gradientů je třeba udělat několik procesů, které budou následně popsány. Nejdříve je nutné provést předpřípravu, musí se tedy zachovat poměr stran 1:2 a ideální rozměr obrázku by měl být 64x128 pixelů pro co nejefektivnější výsledky z deskriptoru. Jiné rozměry jsou možné, je ale nutné zachovat dříve stanovený poměr. Následně spočítáme gradienty v obrázku. Jako první je nutné spočítat horizontální a vertikální gradienty – toho dosáhneme jednoduše pomocí filtračních jader zobrazených v Obrázku 28.[37, 44]



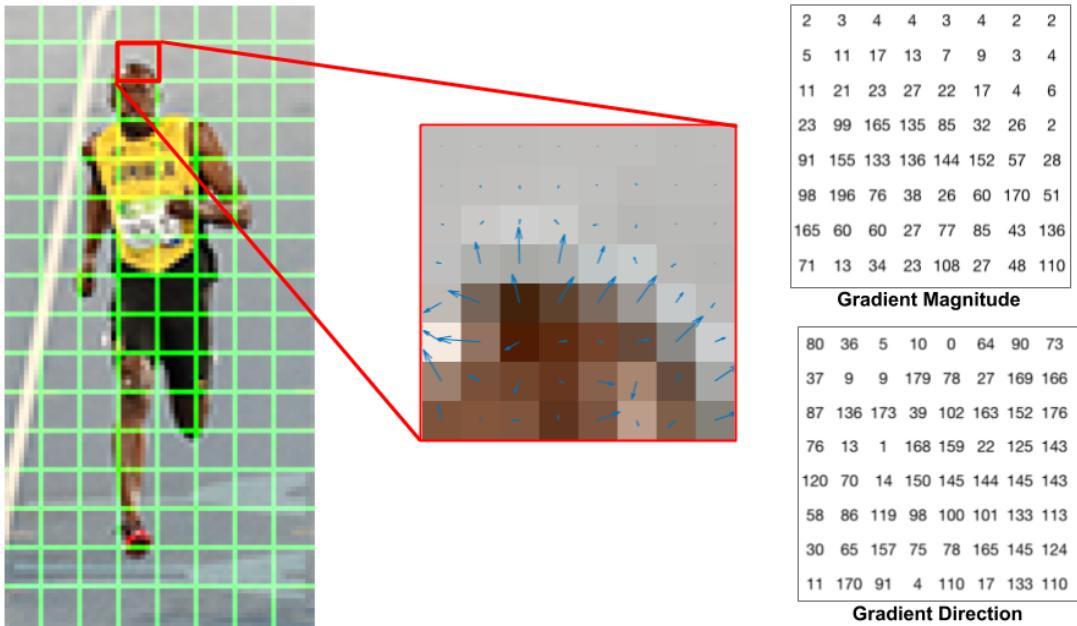
Obrázek 28 Filtrační jádra pro gradienty [37]

Dále zjistíme velikost a směr gradientu pomocí vzorce (2) a (3).[37, 44]

$$g = \sqrt{g_x^2 + g_y^2} \quad (2)$$

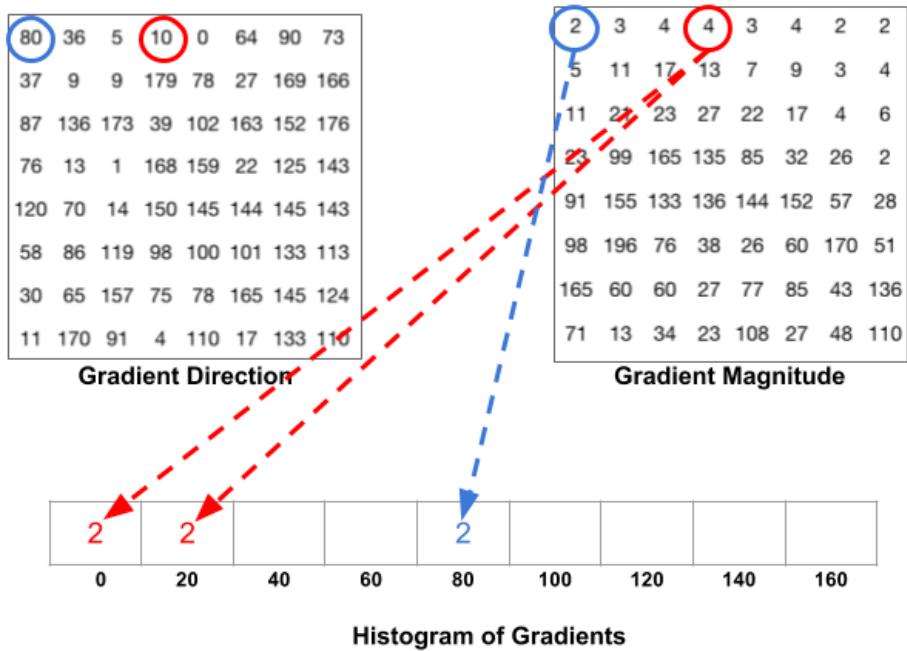
$$\theta = \arctan \frac{g_y}{g_x} \quad (3)$$

Výsledný obrázek gradientu odstranil mnoho nepotřebných informací jako konstantní barva pozadí, ale zvýraznil obrysy. Jinými slovy se lze podívat na obrázek a stále je možné v něm rozpoznat objekt. V každém pixelu má pak gradient velikost a směr. Dalším krokem je sestrojení histogramu gradientů. V tomto kroku je obrázek rozdělen na buňky ve formátu 8x8 a pro každou je následně vypočítán histogram. Důvodem téhoto rozměru je nejenom jejich kompaktnost, ale hlavně také robustnost vůči potenciálnímu šumu. Histogram je pak tvořen 9 intervaly, kdy každý reprezentuje úhly směru – 0, 20, 40 atd. až 160. Při pohledu do Obrázku 29 je vidět výsledek této operace. Uprostřed obrázku lze vidět jednu buňku překrytou výslednými gradienty, šipka ukazuje směr gradientu a délka odpovídá jeho velikosti. Lze pozorovat, že šipky ukazují ve směru změny intenzity a délky odpovídají tomu, jak velká změna je mezi pixely. Vpravo na obrázku pak lze vidět čísla reprezentující jednotlivé gradienty v buňce. Zde je vhodné poznamenat, že směr označuje úhel, nicméně pouze do 180°.[37, 44]



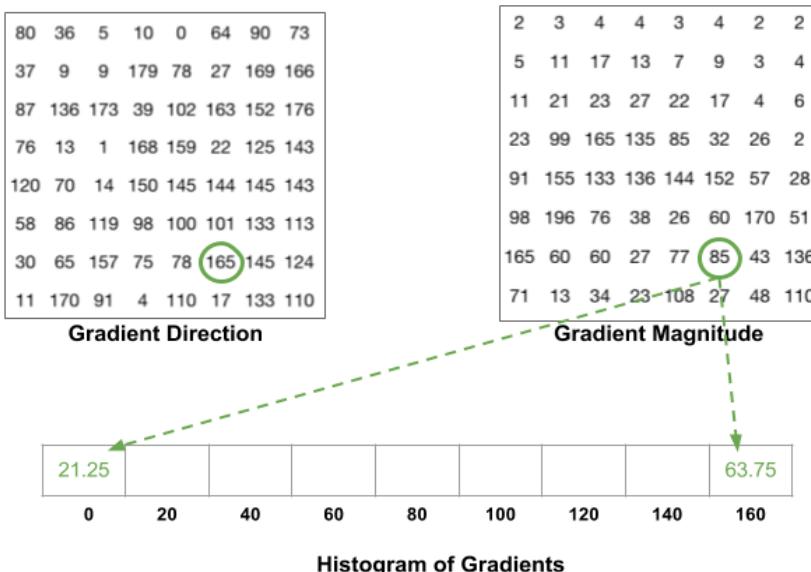
Obrázek 29 Příklad možného rozdělení gradientů a jejich hodnot – vpravo rozvržený obrázek na buňky, uprostřed gradienty v buňce, vlevo hodnoty gradientu. V překladu: Gradient magnitude – Velikost gradientu; Gradient direction – Směr gradientu [37]

Dalším krokem je vytvoření histogramu těchto gradientů. Histogram bude disponovat 9 sloupců s hodnotami, které již byly dříve zmíněny. Způsob, jakým jsou jednotlivé gradienty členěny do sloupců je znázorněn v Obrázku 30. Při pohledu na hodnotu zakroužkovanou modrou barvou lze vidět, že hodnota úhlu je 80, bude tedy zařazena do sloupce s hodnotou 80. Číslo, respektive velikost čísla, které tam bude zapsané je reprezentováno velikostí samotného gradientu, v tomto případě 2. Trochu jiný případ je pak gradient označený červenou barvou. Hodnota jeho směru je 10 a tudíž leží na pomezí mezi prvním a druhým sloupcem. Jeho velikost bude tedy poměrně rozdělena do obou sloupců s respektem k jeho směru. Čtyři bude tedy rozděleno na 2 a 2, jelikož je směr přesně uprostřed hodnot sloupců.[37, 44]



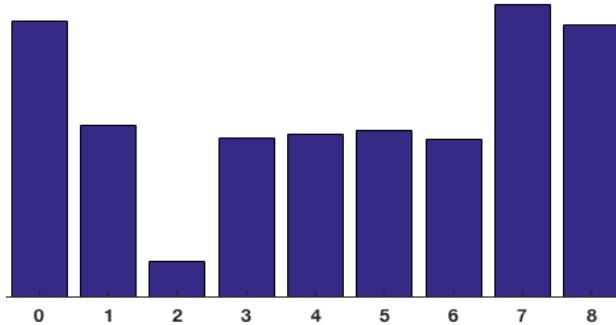
Obrázek 30 Způsob rozčlenění gradientů do histogramu [37]

Zvláštní případ nastane, pokud hodnota směru bude větší než 160. Jak je vidět v Obrázku 31, tak gradient s hodnotou směru 165 stupňů přispěje proporcionalně jak do sloupce 0 stupňů, tak do sloupce pro 160 stupňů.[37, 44]



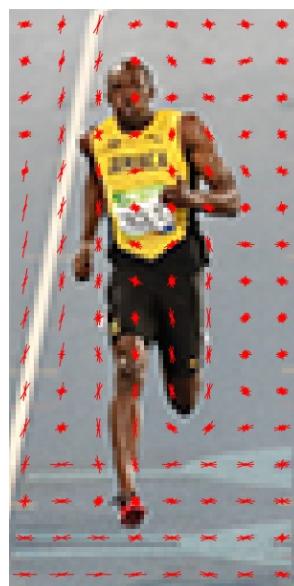
Obrázek 31 Speciální případ rozdělení hodnot gradientu [37]

Po rozdělení všech pixelů v buňce lze vytvořit histogram s devíti sloupcí. Příkladem je Obrázek 32.



Obrázek 32 Histogram gradientů [37]

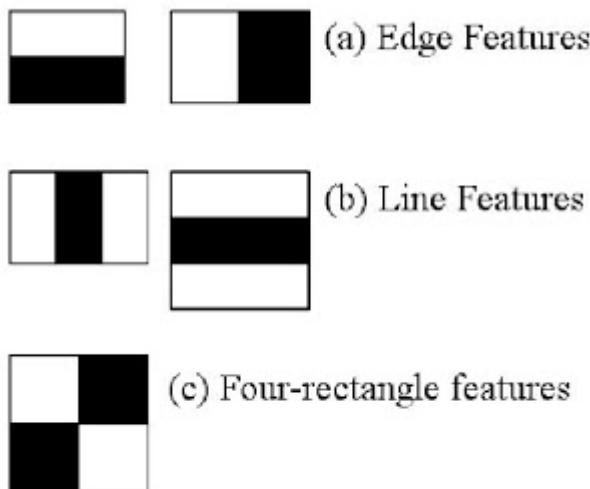
V předešlých krocích byl vytvořen histogram z gradientů části obrázku. Gradienty jsou nicméně citlivé na celkové osvětlení obrázku. Pokud bude obrázek ztmaven při vydělení všech pixelů dvěma, velikost gradientů se taktéž zmenší o polovinu a tím pádem i hodnoty v histogramu. Ideálním případem je ten, kdy bude deskriptor nezávislý na osvětlení. Jinými slovy dalším krokem je normalizace histogramu do vektorů pro zajistění této schopnosti. Po normalizaci všech hodnot je posledním krokem zřetězení získaných vektorů do jednoho obrovského vektoru. HOG deskriptor pak bývá většinou vizualizován vnesením normalizovaných histogramů do jednotlivých buněk, jak lze vidět v Obrázku 33. Je vhodné si zde povšimnout směru, který histogramy ukazují okolo postavy – doslova kopírují jeho tvar a nejvíce zřetelně v oblasti nohou.[37, 44]



Obrázek 33 Vizualizace HOG deskriptoru [37]

3.2.4 Haar kaskádující klasifikátor

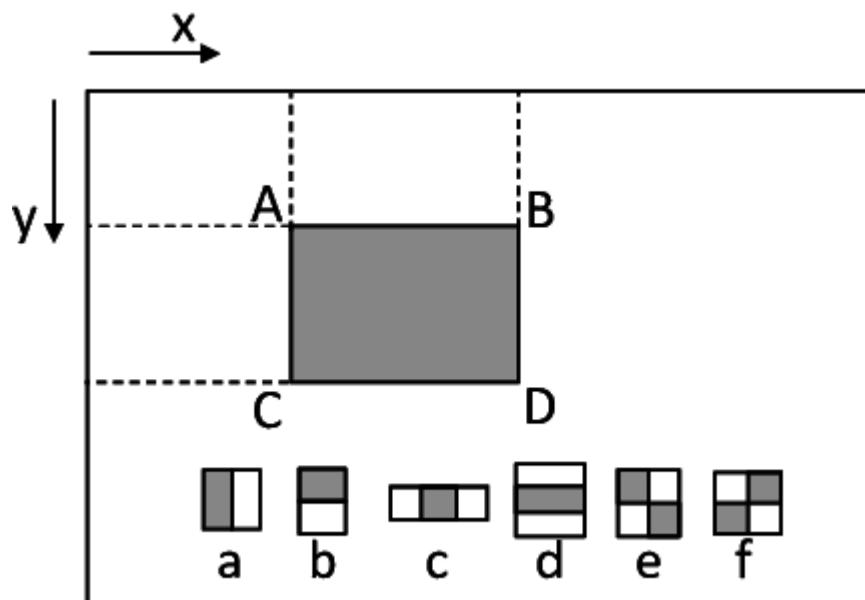
Haar klasifikátor, nebo také Haar kaskádující klasifikátor, je způsob objektové detekce založené na strojovém učení. Je to klasifikátor, který byl jako první na světě využit pro detekci obličeje v reálném čase. V této části kapitoly bude představeno fungování tohoto algoritmu, které se skládá ze tří částí: výpočet Haar vlastností, vytvoření integrálního obrazu a využití Adaboost. Je vhodné poznamenat, že tento algoritmus vyžaduje značné množství pozitivních obrázků obličejů, stejně jako negativních obrázků, na kterých nejsou obličeje, pro jeho korektní fungování. Nejdříve je tedy nutné vypočítat Haar vlastnosti. V esenci se jedná o kalkulace, které jsou prováděny na sousedních obdélníkových regionech na specifické lokaci v detekčním okně. Při kalkulaci se sumarizují intenzity jednotlivých pixelů v každém region. Také je nutné vypočítat rozdíly mezi těmito sumami. V Obrázku 34 lze vidět některé příklady těchto Haar vlastností. Problémem ale je, že tyto vlastnosti se těžko detekují na větších obrázcích.[38, 45]



Obrázek 34 Příklady možných Haar vlastností. V překladu: Edge features – Okrajové vlastnosti; Line Features – Linkové vlastnosti; Four-rectangle features – Vlastnosti o čtyřech obdélnících [38]

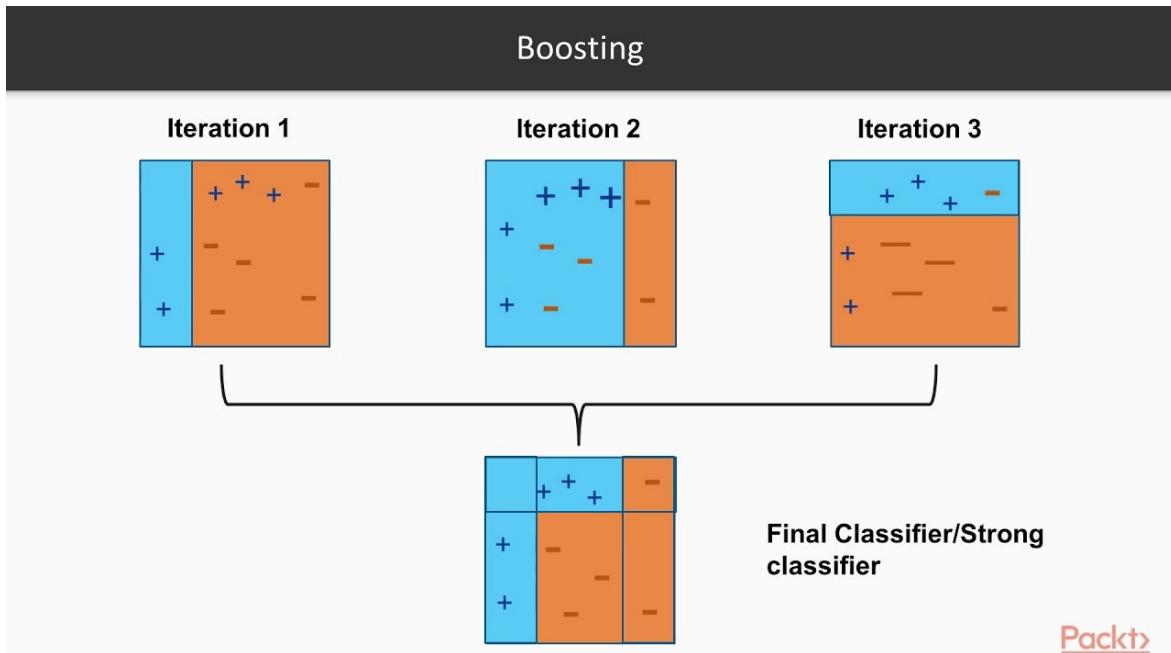
Řešením je právě integrální obraz, kde je množství operací za jeho využití sníženo. Integrální obrazce v podstatě zrychlí výpočty tak, že místo výpočtu na každém pixelu vytvoří určité „podobdélníky“ a vytvoří zároveň pole referencí pro každý tento „podobdélník“. Ty jsou pak použity pro výpočet Haar vlastností obrazu, jak je ilustrováno v Obrázku 35. Je poměrně důležité zmínit, že téměř všechny Haar vlastnosti se stanou irelevantní, pokud provádíme objektovou detekci, jelikož důležitou jsou pouze vlastnosti daného objektu. Zde opět nastává

problém, jak určit ty nejlepší vlastnosti pro reprezentaci daného objektu. Řešením pro tuto situaci je využití Adaboost.[38, 45]



Obrázek 35 Ilustrace fungování integrálních obrazců [38]

Adaboost v podstatě najde nejlepší vlastnosti a naučí je klasifikátor používat. Využívá kombinace tzv. „slabých klasifikátorů“ pro vytvoření „silného klasifikátoru“, který pak algoritmus využije pro detekci. Slabý student je vytvořen tak, že přeneseme okno nad vstupní obrázek a vypočítáme Haar vlastnosti pro každou subsekci daného obrázku. Rozdíl je pak porovnán s naučenou hranicí, které separuje ne-objekty od objektů. Jelikož se jedná o slabé klasifikátory, je nutné shromáždění velkého množství Haar vlastností pro zaručení přesnosti pozdějšího silného klasifikátoru. Posledním krokem je pak kombinace těchto studentů do jednoho silného využitím kaskádujících klasifikátorů (viz Obrázek 36).[38, 45]



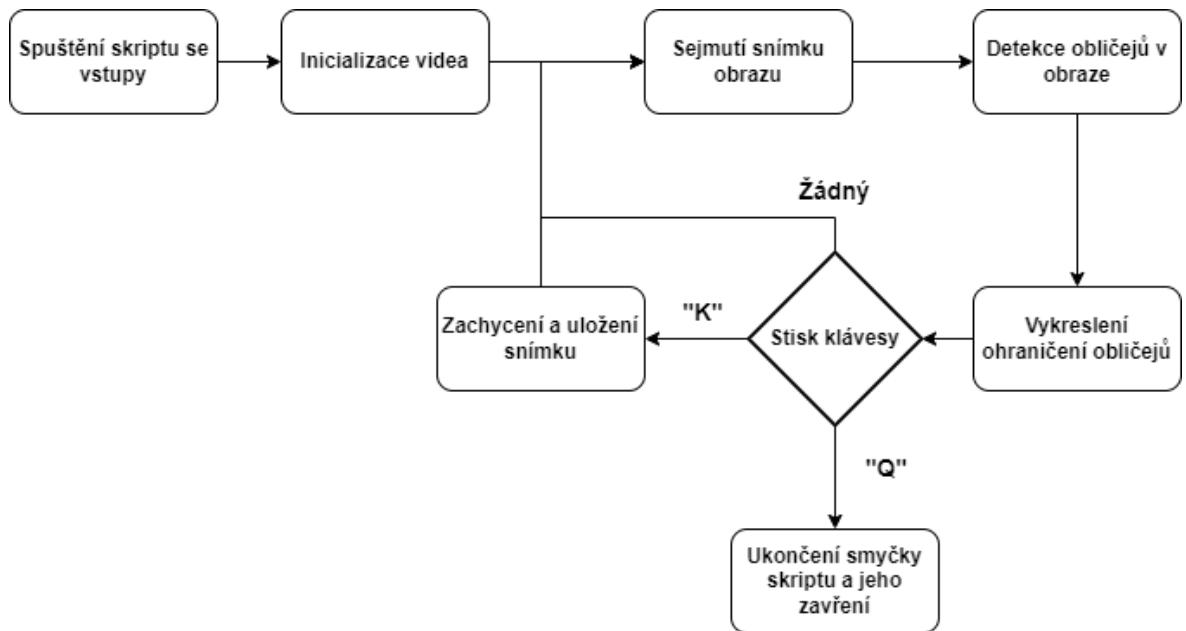
Obrázek 36 Reprezentace fungování Adaboostu [38]

3.2.5 Použité skripty

V této kapitole budou popsány jednotlivé skripty použité pro tuto práci. Budou rozebrány z hlediska jejich koncepce a úkolů, jež mají plnit. Veškeré zdrojové kódy lze nalézt na Github repozitáři (<https://github.com/BarbTheOnlyOne/masters-thesis>). Repozitář je strukturován do několika složek a hlavního skriptu využívaného pro detekci. Také se zde nachází pomocný modul pro posílání emailu.

První důležitou částí je složka s názvem DatasetBuilding. Obsahuje skripty využívané pro stavění datového souboru, který je následně využit při rekognici. První skript, který je potřebný k této stavbě je *build_face_dataset.py*. Při jeho spuštění dojde k inicializaci videa z kamery, která je připojena k počítači. U tohoto skriptu se počítá s jeho využitím na laptopu. Skript pak čte jednotlivé snímky a pomocí Haar klasifikátoru v nich hledá obličeje, okolo kterých vykreslí ohrazení a zobrazí jej uživateli. Dále také odposlouchává stisky jednotlivých kláves, při stisku klávesy „k“ zachytí aktuální snímek a uloží jej do složky, která byla určena při jeho spuštění. V případě stisku klávesy „q“ se skript ukončí se správným odstraněním všech inicializovaných objektů a vypíše na konzoli počet zachycených obrázků. Tento skript tedy slouží k tomu, aby byl daný uživatel schopen vytvořit fotografie svého obličeje pro následující stavbu datového souboru. Pro jednodušší a intuitivnější obsluhu slouží právě vykreslení ohrazení obličeje. V tom okamžiku je uživatel obeznámen s tím, že detektor našel obličej,

a je tedy vhodné uložit tento snímek pomocí dříve zmíněné klávesy. Průběh skriptu `build_face_dataset.py` lze vidět v Obrázku 37.



Obrázek 37 Průběh skriptu `build_face_dataset.py` [autor]

Druhým skriptem v této složce je `face_encoding.py`. Tento skript slouží ke zpracování sesbíraných obrázků jednotlivých osob do jednoho souboru s názvem `encodings`. Při spuštění skriptu je nutné nejdříve předat umístění fotografií daného datového souboru a také zadat, jaká detekční metoda má být použita pro nalezení obličejů. Skript nejdříve načte veškeré obrázky z předané cesty a vytvoří soubor pro zápis jednotlivých „embeddings“, což jsou 128dimenzionální vektory, které jsou výstupem zpracování jednoho obličeje. Po načtení je následně procházena jedna složka osoby za druhou, kdy se v každém obrázku najde obličej podle metody, která byla určena na vstupu. Následně je tento obličej předán do CNN, která jej zpracuje na výše zmíněný vektor a později jej zapíše do souboru s koncovkou `pickle`. Takto jsou postupně iterovány veškeré složky s obrázky až se nakonec skript ukončí. Na konci je soubor uložen a v průběhu je uživatel informován o stavu zpracování, tedy kolikátý obrázek z celku je právě zpracováván.

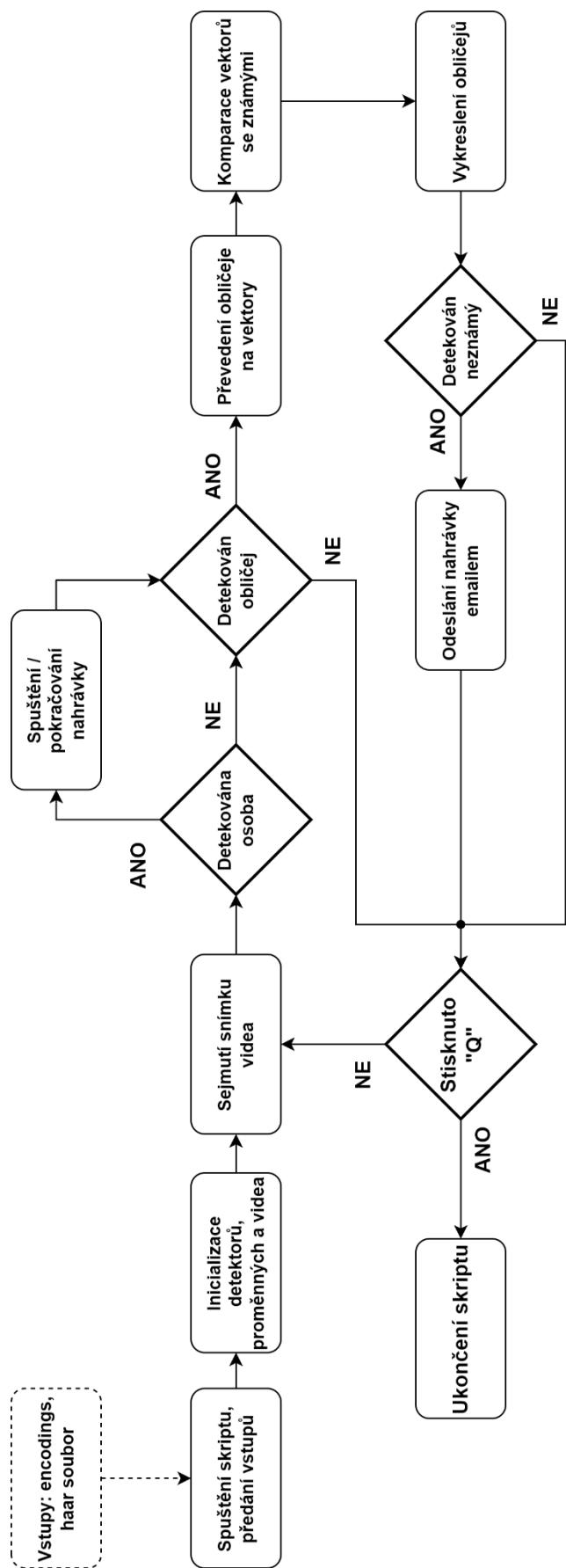
Složka `ProofOfConcept` obsahuje prvotní skripty, které byly použity pro demonstraci a investigaci potenciálního provedení ideje této práce. Je zde skript pro komunikaci s vyhledávačem Bing, který byl použit k sestavení datových souborů známých osobností

z internetu. Skript *recognize_faces.py* byl využit pro otestování detekce ze statického obrázku a *recognize_faces_videostream.py* byl pak první test detekce z živého přenosu obrazu. Detailní fungování těchto skriptů není příliš podstatné v této práci a jejich jádro je vesměs promítnuto v hlavním souboru, který bude později diskutován. Další složka s názvem *DistanceMeasurement* obsahuje dva jednoduché skripty, které byly použity v rámci přípravných testů, kdy *face_distance.py* byl využit pro test vzdálenosti, na kterou je kamera schopna detektovat obličej, *person_distance.py* potom na detekci osoby jako objektu.

Poslední pomocné skripty se nacházejí ve složce s názvem *DataProcessing*. Jelikož počet záznamů měření dosahuje čísla 720 a každé jednotlivé měření obsahuje více sledovaných veličin, je na místě vytvoření pomocných skriptů pro zpracování tohoto množství dat a jeho následnou vizualizaci. Soubor *confusion_matrix.py* slouží k načtení CSV souborů se záznamy a zjištění hodnot potřebných k vytvoření matice záměn (matice je detailněji rozebrána v pozdějších kapitolách). Výstupem tohoto skriptu jsou zobrazené informace na konzoli. Druhý soubor ve složce je *histrogram.py*. Ten má za cíl vizualizaci předaných dat do histogramu četnosti. V tomto případě nebyl vytvořen předávač argumentů, ale je zde přímo proměnná na pevné uvedení cesty k datům. Ve skriptu je proveden výpočet průměru společně se směrodatnou odchylkou, dále vytvořen seznam jednotlivých mezí v histogramu a následně je histogram vytvořen. Graf je vytvářen pomocí knihovny Seaborn. Jak průměr, tak i pravidlo tří sigma je vyneseno do histogramu. Posledním pomocným skriptem je *roc_space.py*. Jedná se o skript, jehož pomocí lze vytvářit graf prostoru receiver operating characteristics (ROC) - více o něm v pozdější kapitole. Stejně jako předešlý, ani tento skript nemá předávač argumentů, a tedy hodnoty pro vytvoření grafu jsou pevně vepsané ve skriptu. Graf je opět vytvořen pomocí knihovny Seaborn. V rámci skriptu jsou obsaženy dvě sady hodnot: jedna pro exemplární ukázku grafu v následující kapitole, druhá pak s aktuálními výsledky měření.

Posledním, a asi i nejdůležitějším, skriptem v repozitáři je soubor s názvem *faces_videostream_raspberry.py*. Jedná se o hlavní skript, ve kterém je vykonáváno snímání obrazu a následná detekce osob, obličejů a také rekognice nalezených obličejů. Skript je poměrně obsáhlý, a proto bude postupně rozebrán a také bude popsán způsob jeho fungování. Při spuštění je nutné předat několik vstupních hodnot. První je soubor pro Haar detektor, který se využívá k detekci obličejů v obraze. Druhý je soubor datového souboru ve formátu pickle, kde jsou zaznamenány veškeré známé obličeje. V praxi se jedná o databázi známých obličejů, vůči které se budou nalezené obličeje porovnávat, a kterou lze vytvořit pomocí již dříve

diskutovaného *face_encodings.py*. Třetím vstupem je zadání cesty pro ukládání zaznamenaných videozáznamů. Poslední je argument, zdali je požadováno zobrazit výstup na obrazovku. Po předání argumentů se inicializují jednotlivé detektory (HOG a Haar) za pomocí OpenCV, dále se načítají data předaná ve formě pickle souboru a také kodeky pro zápis videa. Následně se nastartuje videopřenos a inicializují se určité potřebné proměnné, jako počítadla nebo podmínky a skript vstupuje do smyčky. Ve smyčce vždy dojde k přečtení snímku z kamery a jeho následné zpracování. Je provedena detekce osob pomocí HOG detektoru. Pokud je v obraze zaznamenána přítomnost osoby, je spuštěno nahrávání obrazu. Následně se v obraze detekují možné obličeje pomocí Haar detektoru. Pokud je obličej nalezen, je spuštěno počítadlo snímku a provádí se převedení tohoto obličeje (případně obličejů, vyskytuje-li se na obraze více osob) na 128dimenzionální vektor pomocí CNN a následně je komparován se známou databází. V případě shody je označen příslušným názvem, který je uložen v databázi. V případě neshody je potom označen jako neznámý. V esenci je toto porovnání klasifikátorem typu k-NN, neboli k-nejbližších sousedů. Tento algoritmus klasifikuje neznámé datové body tím, že se pokouší nalézt tu nejběžnější třídu mezi k-nejbližšími příklady. Každý takový bod pak dostává hlas a kategorie s nejvíce hlasy je pak zvolena. Po první identifikaci je, jak již bylo dříve zmíněno, spuštěno počítadlo, které je pak každý snímek zvýšeno o jedna. To slouží k tomu, aby se identifikace prováděla každý určitý počet snímků a mohlo se tak ušetřit výpočetní síly zařízení. Tato hranice byla stanovena na každý desátý snímek. Po dosažení této hodnoty se počítadlo resetuje na hodnotu nula a proces probíhá od začátku. Potom, co byly identifikovány nalezené obličeje, je okolo nich vykresleno označení společně s jejich jménem, případně označením „neznámý“. V poslední části skriptu se pak zapíše sejmuty snímek za předpokladu, že byla detekována osoba ve snímaném obrazu. Pokud se zde již osoba nenachází, ukládání obrazu pokračuje dalších deset sekund pro zajištění adekvátní délky záznamu. Jak tomu bylo i u předešlých skriptů, pokud uživatel stlačí klávesu „Q“, je skript ukončen. Orientační schéma fungování tohoto skriptu lze vidět v Obrázku 38.



Obrázek 38 Schéma průběhu fungování skriptu `faces_videostream_raspberry.py` [autor]

3.3 Selekce detekční metody

Tato kapitola je věnována možnostem detekce obličeje a jejich náročnosti na výpočetní sílu Raspberry Pi, na kterém budou testy prováděny. Cílem této kapitoly, respektive testů provedených v rámci ní, je stanovení nejméně náročné detekční metody pro její následné využití při hledání obličejů v obraze za účelem co nejplynulejšího toku obrazu. Testovanými metodami budou: HOG, CNN a Haar kaskádový klasifikátor. Testovací skript je v esenci skript používaný pro vytvoření *face_encodings.py*, ve kterém byla vždy zvolena příslušná testovaná metoda na vstupu. Tímto testem se tedy, mimo jiné, dá zjistit i časová náročnost vytváření datové sady, respektive databáze známých obrázků.

Pro každou metodu bude proveden test se stanoveným počtem vstupních obrázků. Každý test má tedy za cíl vytvoření datového souboru pro rozpoznání osoby v poskytnutých obrázcích. Zároveň bude pro každou metodu proveden test se stejnou sadou obrázků. V první fázi testu bude deset obrázků, v druhé dvacet, třetí padesát a v poslední sto. Výsledky provedených testů lze pak vidět v Tabulce 3.

Tabulka 3 Výsledky testů detekčních metod [autor]

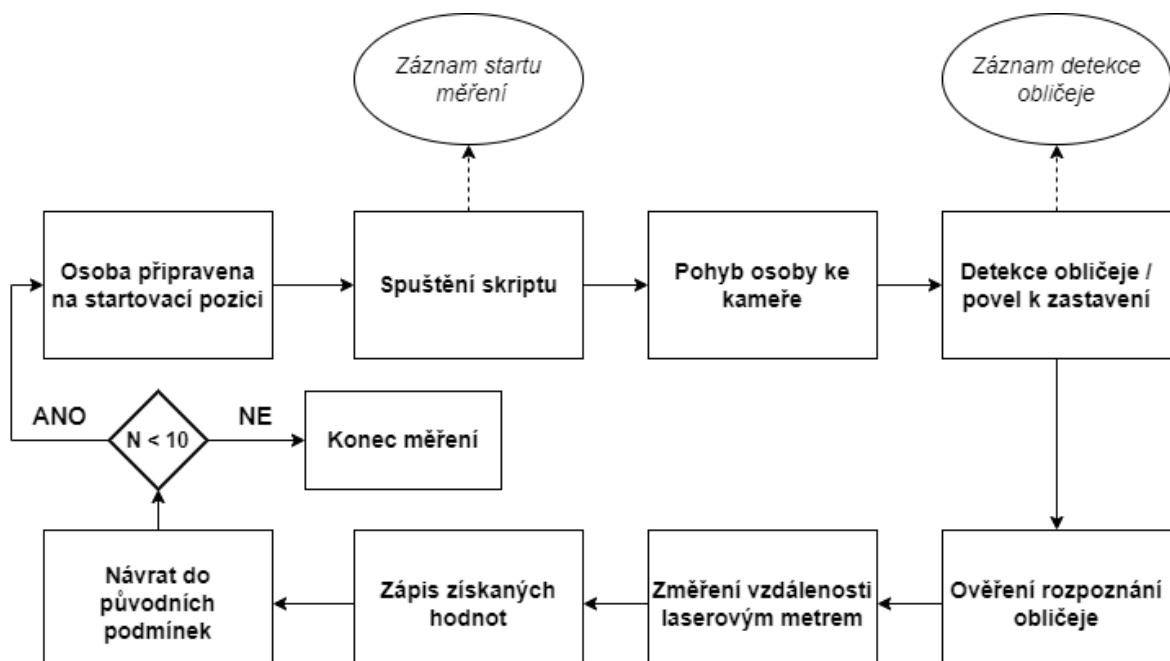
Počet zpracovaných obrázků	Detekční metody		
	CNN	HOG	Haar Cascade
10	178.3731 s	12.3345 s	4.5166 s
20	342.4471 s	23.7945 s	7.6595 s
50	814.6064 s	56.1358 s	19.5668 s
100	1521.4105 s	108.2199 s	38.915 s

Při pohledu na výsledky lze vidět, že jednoznačně nejméně náročná metoda detekce je pomocí Haar kaskádujícího klasifikátoru. Rychlosť zpracování obrázků je v kontextu práce poměrně kritická věc, jelikož je požadavek na detekci obličejů v reálném čase. Proto byla i tato metoda následně zvolena pro užití na detekci obličeje ve skriptu *faces_videostream_raspberry.py*.

4 Metodika měření

Tato kapitola obsahuje deskripci způsobu, jak bylo měření provedeno v obecné rovině, stanovení laboratorních podmínek, uvedení sledovaných veličin a následně grafická ukázka průběhu měření pro jeden testovaný subjekt.

Měření bylo provedeno na vzorku deseti osob. V prvotní fázi bylo nutné sesbírat datový soubor referenčních obrázků pro každou jednotlivou osobu, která se měla účastnit měření. Soubor byl pro osobu vytvořen pomocí skriptu *build_face_dataset.py*, kdy testovaný subjekt byl schopen sám vytvořit vhodný datový soubor. Počet referenčních obrázků se pohyboval okolo 50 až 60 na osobu, nikdy ale ne méně než 50. Po získání všech souborů z nich byl vytvořen jeden celistvý referenční soubor pro klasifikátor pomocí *face_encoding.py*. Dalším krokem bylo již samotné měření osob, jehož průběh lze vidět na Obrázku 39.



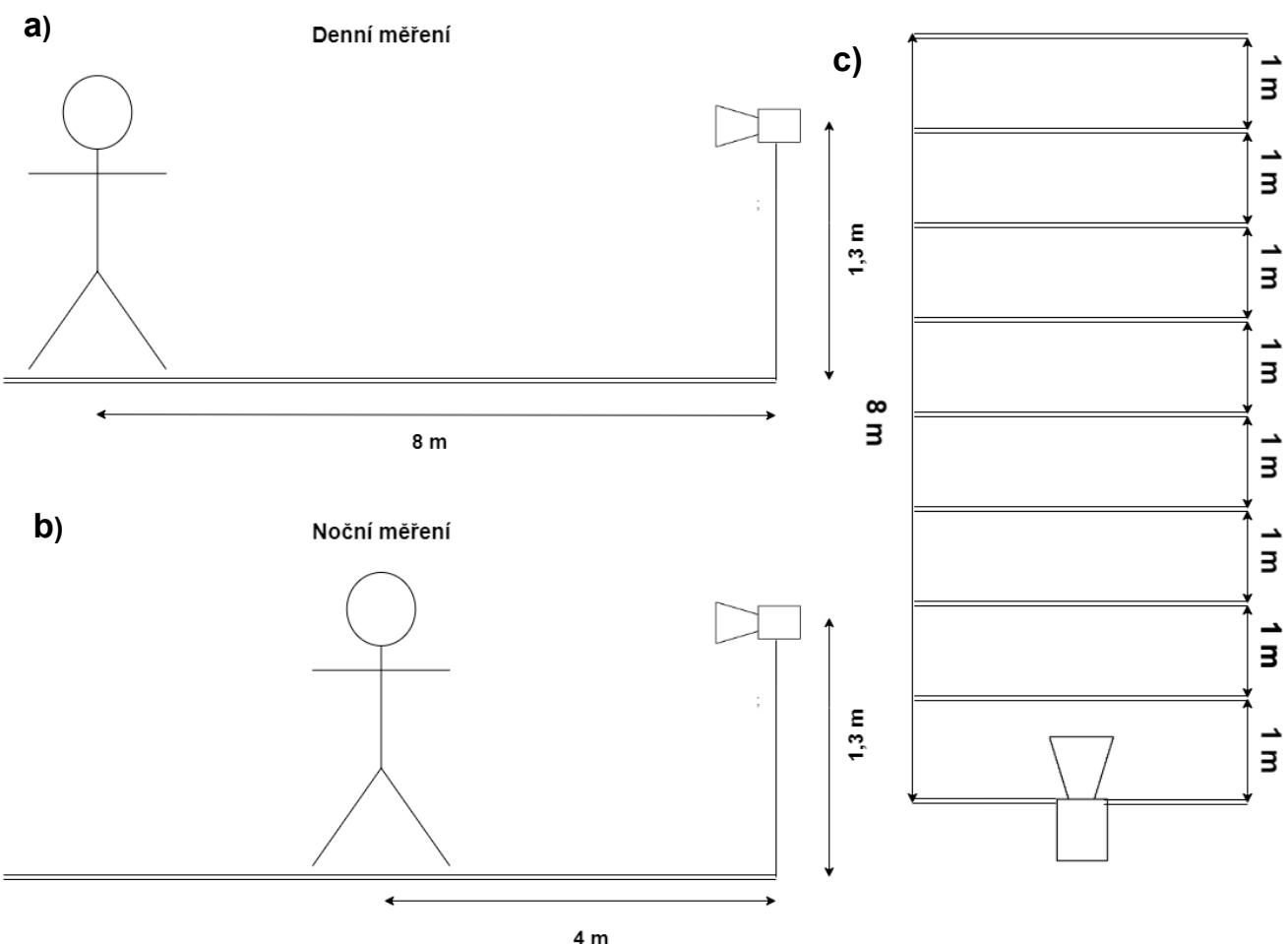
N = počet měření

Obrázek 39 Flowchart průběhu měření jedné osoby [autor]

Měření jednotlivých subjektů pak probíhalo ve dvou fázích – denní a noční. Každá fáze byla rozdělena na tři části, které byly pro denní i noční režim stejné. Více k jednotlivým částem v kapitole 4.2.

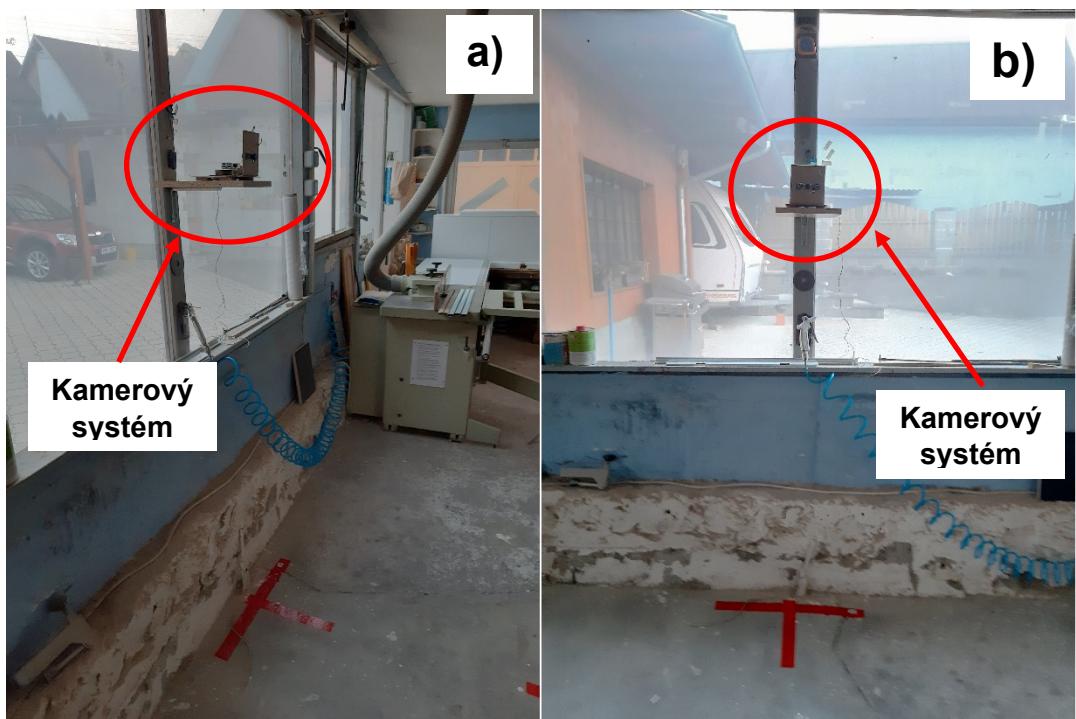
4.1 Laboratorní podmínky

Pro testování byl vytyčen dostatečný prostor v budově s vhodnými rozměry, ve kterém se testované osoby mají pohybovat. Tato plocha byla zároveň segmentována na metrové úseky pro jednoduchou a rychlou kontrolu vzdáleností a orientaci (viz Obrázek 40 část c). Kamera byla umístěna ve výšce 1,3 metru. Tato výška byla zvolena na základě výpočtu optického záběru kamery (50°) tak, aby byl v co nejmenší vzdálenosti stále ještě v záběru obličej a zároveň bylo viditelně rozpoznatelné značení vzdálenosti na podlaze. Startovní pozice byla umístěna 8 metrů od kamerového systému pro denní měření, 4 metry pak pro noční z důvodu dosvitu IR diod, viz kapitola 3.1.2. Nákres prostředí se vzdálenostmi lze vidět v Obrázku 40.



Obrázek 40 Nákres vzdáleností měření – a) rozložení denního měření, b) rozložení nočního měření, c) segmentace prostoru při pohledu shora [autor]

Fotografie umístění kamery pak lze vidět na Obrázku 41. Rozčleněná plocha použitá k testování, kterou lze vidět na nákresu c) v Obrázku 40, je zachycena na fotografii v Obrázku 42.



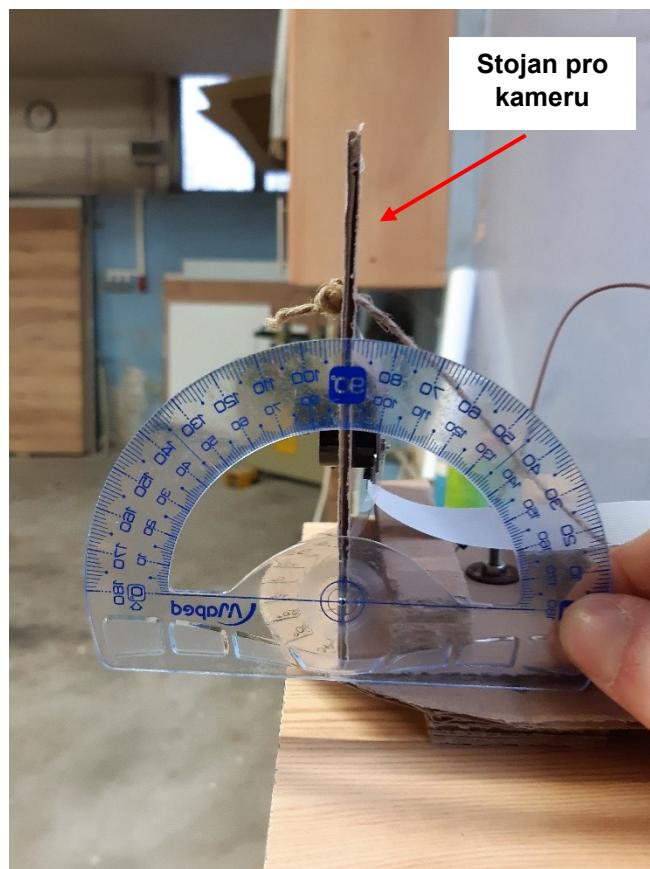
Obrázek 41 Fotografie umístění kamerového systému – a) pohled z boku, b) čelní pohled [autor]



Obrázek 42 Fotografie označené plochy pro měření [autor]

4.2 Představení sledovaných veličin a poloh měření

Měření bylo v jednotlivých fázích členěno na tři části, přičemž v každé části měla kamera jinou polohu, viz předcházející kapitoly. Smyslem různých poloh bylo otestovat rozpoznávací algoritmus v krajních polohách či při nepřímých úhlech kamery. Jako první a výchozí poloha bylo přímé namíření kamery na testovanou osobu. Kamera je v pravém úhlu s podložkou a není nijak vytočena do stran, viz Obrázek 43. V průběhu práce je tato poloha označována jako poloha A.



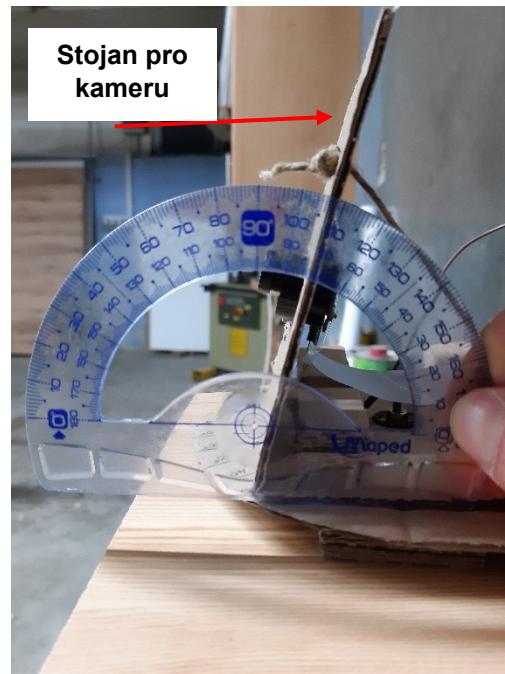
Obrázek 43 Fotografie kolmé přímé polohy A [autor]

Druhou polohou, která byla využita při měření, bylo vytočení kamery do strany pod úhlem dvacet stupňů doleva při čelním pohledu na kameru. Důvodem této polohy je otestování schopnosti algoritmu v případě pohybu osob na okrajové části záběru kamery. V průběhu práce je tato poloha označována jako poloha B a lze ji vidět v Obrázku 44.



Obrázek 44 Fotografie vytočení kamery do strany, poloha B [autor]

Poslední, a to třetí polohou, je poloha C. Kamera je v přímém směru vůči testovanému subjektu a v záklonu dvaceti stupňů. Tato poloha má také za cíl otestování rozpoznání obličeje v krajních situacích. Záklon lze vidět na Obrázku 45.



Obrázek 45 Fotografie polohy v záklonu, polohy C [autor]

Jako první možný parametr, který lze sledovat, se jeví korektní rozpoznání obličeje testovaného subjektu. Výsledek je binárního charakteru, to znamená, že záznamem bude pouze ano či ne. V případě autorizovaných osob bude správné rozpoznání ohodnoceno výsledkem „ANO“, v případě špatného rozpoznání nebo dokonce přiřazení označení jako „Neznámý“ je pak ohodnocen výsledek jako „NE“. U neautorizovaných osob je pak předpokládáno, že algoritmus by tyto osoby neměl vůbec rozpoznávat, a tedy by měly být označeny jako neznámé a v tomto případě tedy bude záznam „ANO“, jako správné rozpoznání neznámé osoby. V případě, kdy bude neautorizovaná osoba rozpoznána jako autorizovaná, je výsledek hodnocen jako „NE“. Při měření bylo z důvodu principu fungování skriptu na rozpoznání obličeje nutno vyčkat na alespoň tři za sebou jdoucí správná rozpoznání obličeje, aby bylo možné pokus vyhodnotit jako „ANO“. Pokud se během této série vyskytlo vyhodnocení obličeje jako jiné či neznámé osoby, bylo výsledkem „NE“.

Druhou sledovanou veličinou je vzdálenost detekce obličeje od kamerového systému. Jedná se o vzdálenost, kdy kamera poprvé detekovala obličej v záběru. Testovaná osoba se na této vzdálenosti zastavila a následně byla vzdálenost meziní a kamerou změřena pomocí laserového metru pro zajištění co největší přesnosti. Vzdálenost je měřena od objektivu kamery až po osobu.

Mimo tyto dvě dříve zmíněné veličiny, bylo v rámci testování také zaznamenáváno několik časových parametrů pro zajištění věrohodnosti měření a případnou budoucí investigaci možných dalších problémů či rozšíření. Při každém začátku měření byl zaznamenán čas startu experimentu, dále pak čas detekce obličeje ve snímaném obrazu. Oba tyto časy byly vypisovány automaticky skriptem při spuštění a následné detekci, aby byla zajištěna přesnost časových informací.

V rámci měření byl navrhován ještě jeden parametr pro sledování, a to vzdálenost detekce samotné osoby ve snímku. Nicméně v přípravných testech bylo zjištěno, že daný algoritmus pro detekci osob v obrazu je natolik efektivní v nalezení subjektů, že by jeho zahrnutí v testech nemělo vypovídající hodnotu z toho důvodu, protože osoba byla detekována prakticky ihned po spuštění testovacího skriptu, a tudíž efektivní vzdálenost i čas detekce sledované osoby se rovnal začátku startovací pozice. Jak již bylo řečeno v přípravných testech, detekce osoby probíhala na mnohem větší vzdálenosti, než bylo možné zajistit v rámci testovací plochy.

4.3 Metody vyhodnocení výsledků

V této kapitole budou představeny metody, jež budou použity k interpretaci a vizualizaci výsledků a jejich následnému hodnocení včetně komentáře. Představena bude také matice záměn společně s grafem ROC pro evaluaci správného rozhodování algoritmu. Jako další nástroj pak bude uveden histogram, který znázorní četnosti v měření vzdálenosti detekce obličeje při experimentech.

4.3.1 Matice záměn

Matice záměn, kterou lze vidět na Obrázku 46, bývá také často nazývána jako konfuzní matice a je hojně využívaným nástrojem na poli strojového učení. Jedná se o konkrétní rozložení tabulky, která umožňuje vizualizaci výkonnosti daného algoritmu – nejčastěji za přítomnosti supervizora. Každý řádek pak představuje instanci v dané třídě, naopak sloupce ukazují predikovanou instanci. Sloupce s řádky jsou samozřejmě zaměnitelné bez vlivu na fungování matice. Název této matice pramení z faktu, že je poměrně jednoduché určit, zdali zkoumaný systém zaměňuje jednu instanci za druhou.[49]

		Předpovídáný stav	
		Celkové obsazení: P + N	Positivní predikce (PP)
Skutečný stav	Positivní skutečnost (P)	Pravý pozitiv (TP)	Falešný negativ (FN)
	Negativní skutečnost (N)	Falešný pozitiv (FP)	Pravý negativ (TN)

Obrázek 46 Exemplární matice záměn [autor]

Po formulaci matice lze pomocí jednoduchých výpočtů získat ukazatele, které pomohou při hodnocení efektivnosti zkoumaného algoritmu. Jako základní dva ukazatele lze zvolit citlivost (sensitivity), také nazývanou jako True Positive Rate (TPR), a specifičnost (specificity), také nazývanou jako True Negative Rate (TNR).[49]

Citlivost nám ukazuje, jaké množství osob ze skutečně pozitivní skupiny bylo označeno klasifikátorem jako pozitivní. Lze ji tedy vnímat jako pravděpodobnost, s jakou bude zkoumaný

subjekt zařazen do pozitivní sekce za předpokladu, že tam opravdu náleží. Vyšší hodnota tohoto ukazatele je tedy žádoucí. Lze ji vypočítat dle rovnice (4):[49]

$$TPR = \frac{TP}{TP + FN} \quad (4)$$

Specifičnost lze chápat podobně, jako ukazatel citlivosti, ale vztažený na negativní subjekty. Určuje, jaké množství osob ze skupiny skutečně negativních bylo klasifikátorem rozlišeno jako negativní. Srostoucím číslem specifičnosti je méně lidí označeno jako autentizováno i přes fakt, že by do této skupiny spadat neměli. Tudíž opět vyšší hodnota je preferována. Specifičnost lze určit následně pomocí rovnice (5):[49]

$$TNR = \frac{TN}{TN + FP} \quad (5)$$

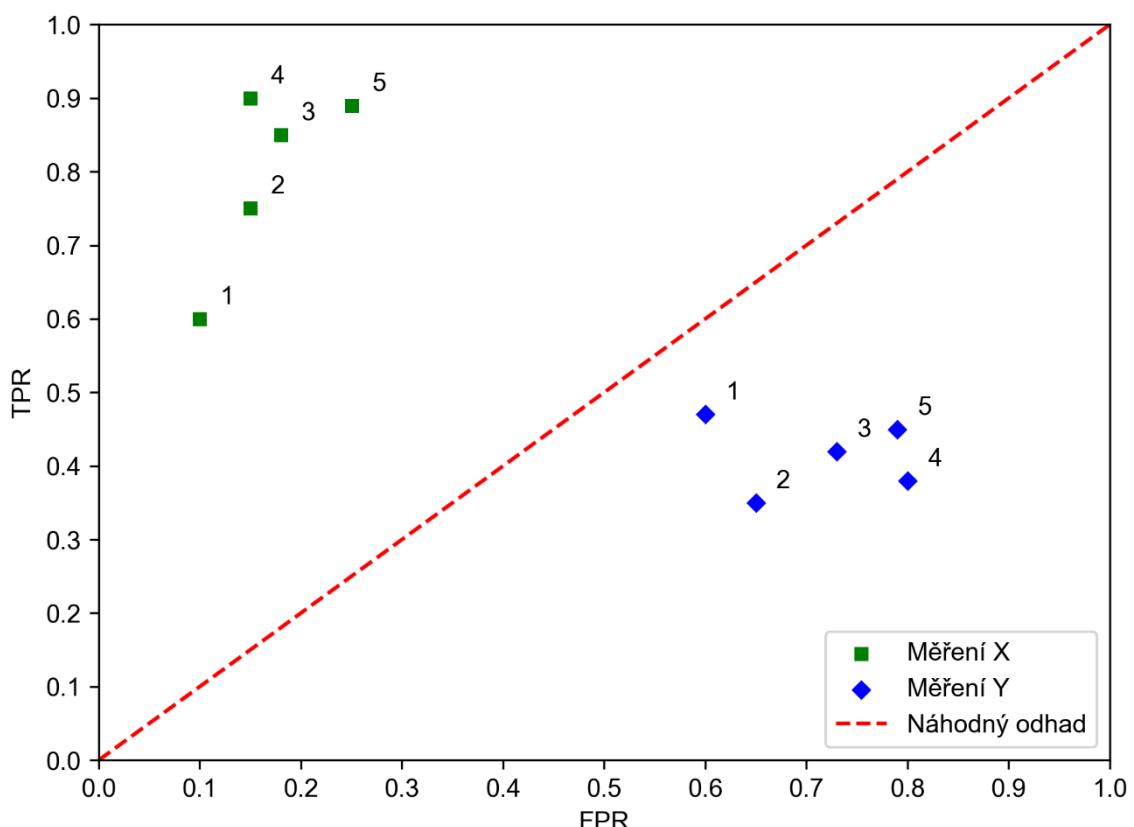
Jako poslední možný ukazatel lze uvést přesnost (accuracy, též označovanou ACC), která obecně vypovídá o tom, jak blízko či daleko jsou zkoumané hodnoty od jejich reálných hodnot. Takže ji lze uvést jako obecný ukazatel výkonu algoritmu pro dané podmínky měření. Výpočet vychází z rovnice (6):[49]

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

4.3.2 ROC prostor

ROC je grafický způsob ilustrace, jak zjistit schopnosti binárního klasifikátoru. Graf je vytvořen pomocí dvou veličin, True positive Rate a False Positive Rate, kdy dřívější je známo také jako citlivost a pozdější jako pravděpodobnost falešného poplachu, někdy označováno jako $1 - \text{specifičnost}$ (stejně tak se i provádí výpočet této veličiny). Je to efektivní způsob, jak zobrazit výkonnost strojového učení a přehledné interpretace výsledků. Na ose X bývá vynesena FPR a na ose Y pak TPR. Výsledek každé matice záměn pak představuje jeden bod na grafu. V grafu bývá taktéž vynesena linka vedoucí z bodu (0,0) do bodu (1,1), která reprezentuje výsledek, jenž by odpovídal náhodnému odhadu (tedy šance 50:50 při odhadování objektu). Výsledky se následně hodnotí dle toho, v jaké vzdálenosti a na které straně jsou od této linky náhodného odhadu. Pokud je bod nad touto linkou, značí efektivnější řešení než právě zobrazený náhodný odhad. Je-li pod ní, značí, že klasifikátor dělá odhadu přesně opačně, než je požadovaný výsledek, a tedy je kontraproduktivní a horší než náhodný odhad.[48, 49]

Ukázkový případ lze vidět v Obrázku 47. Na grafu jsou vyneseny dva druhy měření, jedním je měření X a druhým Y. Jak již bylo dříve zmíněno, body vykreslené nad čárou náhodného odhadu jsou pro nás přínosným výsledkem, stejně tomu je i zde, u měření X. Druhé měření Y má naopak horší výsledky jak náhodný odhad. Lze o něm říci, že sledované veličiny tento klasifikátor vyhodnocuje právě v náš neprospěch, a tudíž není vhodný k reálnému použití nebo pro hodnocení daných parametrů.



Obrázek 47 Exemplární graf ROC prostoru [autor]

4.3.3 Histogram

Pro efektivní vyhodnocení sledované veličiny vzdálenosti detekce obličeje od kamerového systému se jeví použití histogramu. Jedná se o grafické znázornění distribuce dat za pomocí sloupcového grafu, kdy každý sloupec má stejnou šířku a vyjadřuje rozpětí intervalů, též nazývaných tříd. Výška sloupce následně ukáže četnost veličiny sledované v daném intervalu, kterou sloupec zároveň reprezentuje. Abychom zajistili adekvátní vypovídající hodnotu histogramu, je nutné zajistit vhodné zvolení tříd, tedy šířky intervalů. Pro účely této práce byly zvoleny meze intervalů autorem práce. Nejnižší mez začíná na hodnotě 1,25 m a nejvyšší mez je 4,00 m. Jednotlivé kroky jsou pak o hodnotě 0,25 m. Nejnižší mez byla vybrána na základě nejnižší nalezené hodnoty v měření. Stejně platí pro nejvyšší mez.[14, 46, 47]

Abychom mohli výsledky měření korektně posoudit, je vhodné na daný soubor dat aplikovat tzv. Pravidlo tří sigma. Toto pravidlo nám říká, že pokud máme přibližně normálně rozdelený statistický soubor, tak téměř všechny jeho relevantní hodnoty by se měly nacházet do tří směrodatných odchylek (které se značí σ – odtud název pravidla) od aritmetického průměru. V okolí jedné směrodatné odchylky od průměru uvažujeme zhruba 68,27 % hodnot rozdělení, u dvou již 95,45 % a u tří až 99,73 % hodnot. Lze tedy obecně říci, že čím vyšší je násobek směrodatné odchylky, tím více tolerujeme extrémní naměřené hodnoty.[14, 46, 47]

Většina měření bývá zatížena chybami, které mohou střední hodnotu zkreslit a zhoršit tak její vypovídající hodnotu. Pro přiblížení ke skutečné střední hodnotě tedy zjistíme aritmetický průměr souboru se všemi hodnotami (i s těmi, co mohou být zatížené hrubou chybou). Následně zjistíme standardní odchylku hodnot a použijeme pravidlo tří sigma, a to tak, že ze souboru vyřadíme ty hodnoty, které jsou vzdálené více než 3 směrodatné odchylky od průměru, spočítaného na začátku. Z tohoto očištěného souboru dat následně zjistíme nový průměr, který již bude mnohem blíže skutečné střední hodnotě.[46, 47]

5 Výsledky měření

Pátá kapitola je věnována přehledu a hodnocení výsledků, které byly získány předešlými experimenty. Kapitola je rozdělena na výsledky denního a nočního měření. Závěr kapitoly je věnován celkovému zhodnocení výsledků. Ty jsou jsou tvořeny ze souboru deseti známých osob a dvou neznámých osob.

5.1 Výsledky denního měření

Výsledky budou strukturovány do tří skupin dle pozic kamerového aparátu – přímá pozice, rotace o 20° do strany a v záklonu 20° .

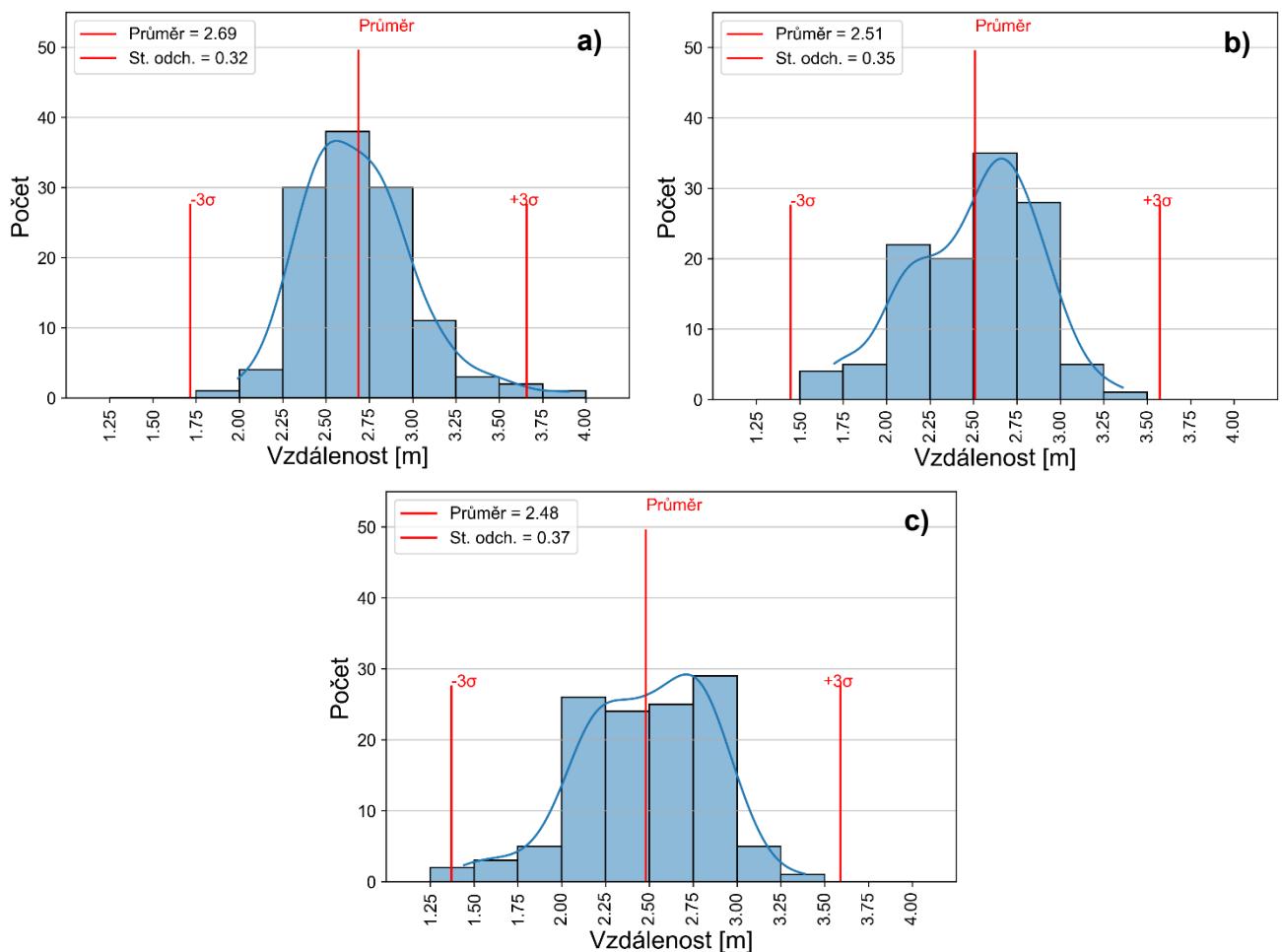
Na Obrázku 48 lze jednoduše opticky porovnat výsledky z denních měření. Srovnáme-li přesnost (accuracy) jednotlivých tabulek, vidíme, že algoritmus měl nejpřesnější výsledky v tabulce označené písmenem b), tedy v měření při rotaci do strany o 20° . Citlivost dosáhla hodnoty 0.94. To nám ukazuje jednoznačně vysokou míru komfortu při užívání systému v těchto podmínkách, jelikož pouhých 6 pokusů ze 100 bylo vyhodnoceno jako falešně negativní. Na druhé straně stojí tabulka a), kdy citlivost má hodnotu 0.85, a tedy není na stejně úrovni komfortu, jako v ostatních případech. Dále pak tabulka s označením c) stojí přesně mezi již zmíněnými tabulkami, má hodnotu hodnotu citlivosti 0,90 a deset falešně negativních pokusů.

Druhý hodnocený ukazatel je specifičnost, která je pro nás důležitějším parametrem, jelikož se jedná o možnost, kdy systém neoprávněnou osobu identifikuje jako oprávněnou, a tudíž hrozí riziko zneužití možným vetřelcem. Nejlepšími výsledky opět disponuje tabulka b), ve které specifičnost dosáhla hodnoty 0,95 a tudíž lze v tomto případě označit systém jako spolehlivý. Zbylé dvě tabulky dosahují stejných výsledků specifičnosti a to 0,90. Lze tedy prohlásit, že spolehlivost je v obou případech na stejně úrovni, alespoň co se bezpečnostní stránky týče.

		Předpovídáný stav			
		Positivní predikce (PP)	Negativní predikce (PN)		
Skutečný stav	Celkové obsazení: P + N				
	Positivní skutečnost (P)	85	15		
	Negativní skutečnost (N)	2	18		
	Accuracy =	0.8583		Skutečný stav	
TPR =	0.85		Positivní skutečnost (P)	90	
TNR =	0.90		Negativní skutečnost (N)	10	
b)	Předpovídáný stav				
Skutečný stav	Celkové obsazení: P + N				
	Positivní skutečnost (P)	94	6		
	Negativní skutečnost (N)	1	19		
	Accuracy =	0.9416		Skutečný stav	
TPR =	0.94		Positivní skutečnost (P)	90	
TNR =	0.95		Negativní skutečnost (N)	18	

Obrázek 48 Matice záměn pro denní měření – a) přímá pozice, b) rotace do strany, c) záklon [autor]

Při pohledu do Obrázku 49 lze vidět, že grafy a) i b) mají oba určitou šikmost. V grafu a) je šikmost zřejmá nejvíce, jedná se o pozitivní šikmost a má tzv. pravý ocas. Graf b) má patrně menší šikmost distribuce dat, ale i zde se vyskytuje negativní šikmost. Graf c) je ze všech nejblíže symetrickému rozdělení, nicméně i v tomto grafu lze pozorovat lehkou negativní šikmost. V případě pozitivní šiknosti lze předpokládat výskyt mediánu v nižší oblasti, tedy nalevo od průměru. Pro negativní platí opak, tedy medián bude mít větší hodnotu než průměr. Tyto fakta jsou pozorovatelná ve všech třech grafech. Takto minimální asymetričnost nemá žádný význam na vyhodnocení výsledkové množiny a reprezentativní hodnotu pozorovaného průměru.



Obrázek 49 Histogram četností pro denní měření – a) přímá pozice, b) rotace do strany, c) záklon [autor]

Je vhodné taktéž poznamenat, že pouze v grafu a) je zaznamenána odlehlá hodnota, která nespadá do intervalu stanoveného pravidlem 3 sigma. Tato hodnota se nachází jak na kladné straně histogramu, tak i na záporné, ačkoliv opravdu ve velmi malém množství. Odlehlé hodnoty můžeme také nalézt v grafu s označením c), kde je tento trend opačný vůči grafu a). Vidíme zde určité množství odlehlých hodnot na záporné straně osy (na straně kladné pak už jen minimální). Při porovnání průměrů lze vidět, že největší průměrná hodnota vzdálenosti je u polohy přímé, tedy a). Je zde také patrná koncentrace hodnot v rozmezí 2,25 do 3,0 metrů, kdy ostatní hodnoty se vyskytují již jen zřídka. Tento trend je vidět u grafu c) s lehkou anomalií kdy středové hodnoty okolo průměru jsou v počtu nižší než jejich sousední sloupce, a tedy vytváří ve středu grafu pomyslnou prohlubeň. V grafu b) lze vidět středový nejpočetnější sloupec, od kterého se hodnoty následně postupně snižují.

5.2 Výsledky nočního měření

Stejně jako v předešlé kapitole pro denní měření, i zde jsou výsledky strukturovány ve stejném pořadí: přímá pozice, rotace o 20° do strany a v záklonu 20° .

Jako vhodným obecným ukazatelem pro rychlé porovnání jednotlivých pozic může být použita přesnost (accuracy). V Obrázku 50 má nejvyšší hodnotu poloha b), a to 0,575. Zbylé dvě pozice nejsou příliš daleko. Obecně mají hodnotu nižší v řádech setin, kdy stav c) je druhý nejpřesnější s hodnotou 0,5333 a jako poslední stav a) dosáhl přesnosti 0,5. Citlivost u všech tří poloh není příliš odlišná, nejvyšší hodnotu má opět poloha b) a to 0,51. V poloze c) je citlivost 0,49 a poslední je poloha a) s výsledkem 0,45. Obecně lze říci, že to nejsou příliš vysoké hodnoty, a tedy komfort používání systému za těchto podmínek je značně snížený.

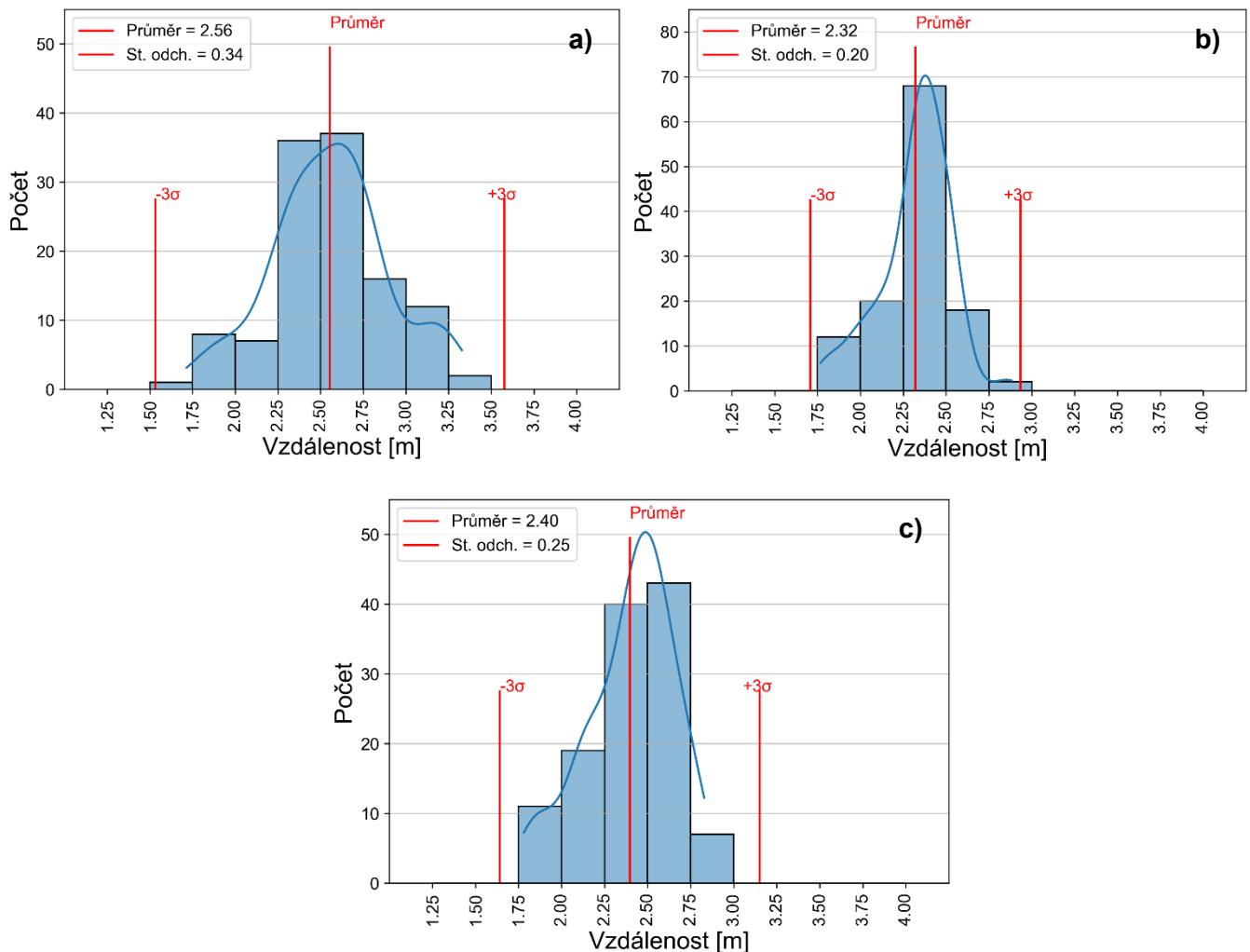
		Předpovídaný stav	
		Positivní predikce (PP)	Negativní predikce (PN)
Skutečný stav	Celkové obsazení: P + N		
	Positivní skutečnost (P)	45	55
Skutečný stav	Negativní skutečnost (N)	5	15
Accuracy =	0.5000		
TPR =	0.45		
TNR =	0.75		

		Předpovídaný stav	
		Positivní predikce (PP)	Negativní predikce (PN)
Skutečný stav	Celkové obsazení: P + N		
	Positivní skutečnost (P)	51	49
Skutečný stav	Negativní skutečnost (N)	2	18
Accuracy =	0.5750		
TPR =	0.51		
TNR =	0.90		

		Předpovídaný stav	
		Positivní predikce (PP)	Negativní predikce (PN)
Skutečný stav	Celkové obsazení: P + N		
	Positivní skutečnost (P)	49	51
Skutečný stav	Negativní skutečnost (N)	5	15
Accuracy =	0.5333		
TPR =	0.49		
TNR =	0.75		

Obrázek 50 Matice záměn pro noční měření – a) přímá pozice, b) rotace do strany, c) záklon [autor]

Nejvyšší specifičnost je v poloze rotace do strany, tedy b), kde dosahuje hodnoty 0,90. Lze říci, že se jedná o vysokou hodnotu, schopnou konkurovat hodnotám při denním měření. Zbylé dvě polohy mají obě hodnotu specifičnosti 0,75, tedy ne příliš vysokou v porovnání s hodnotou ze stavu b).



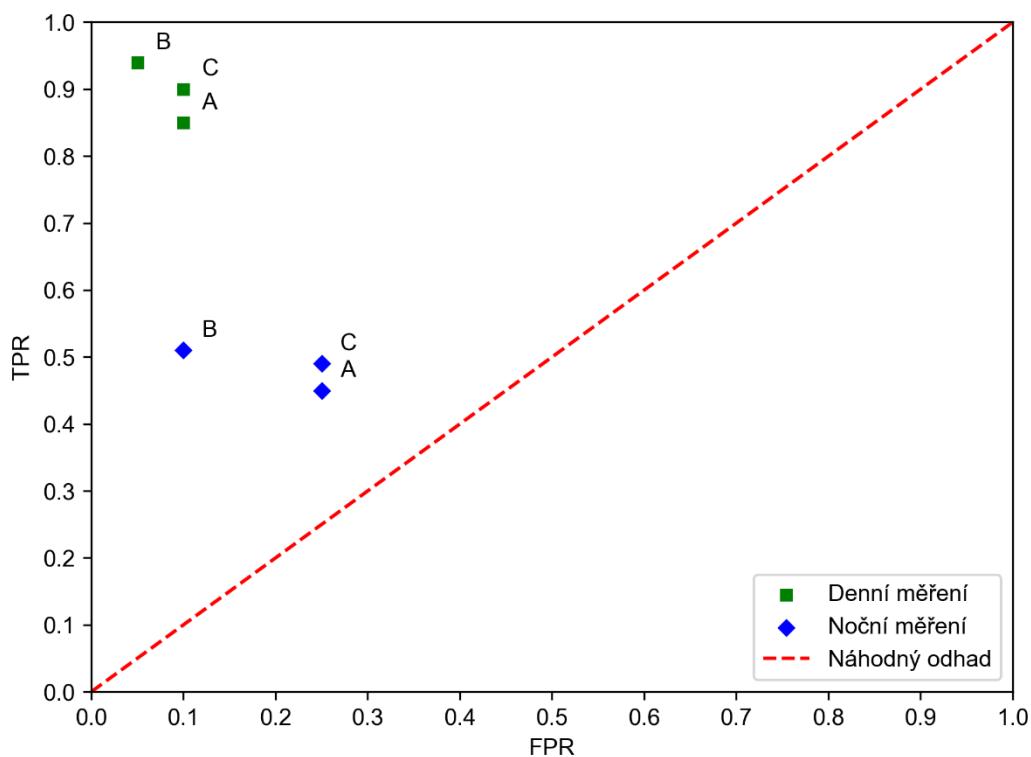
Obrázek 51 Histogram četností pro noční měření – a) přímá pozice, b) rotace do strany, c) záklon [autor]

Histogramy četnosti jsou vyobrazeny v Obrázku 51. První histogram a) ukazuje poměrně symetrické rozdělení hodnot bez šikmosti. Průměr leží v nejčetnější množině a jednotlivé četnosti postupně klesají na obě strany. Podobný stav vidíme i v grafu b), kde je ale středový sloupec obzvláště četný, dosahující okolo hodnoty 70 měření, což je v porovnání s ostatními případy neobvyklé. Histogram také disponuje lehkým zakřivením, které je negativní, má tedy levý ocas. Kuriozním úkazem v tomto případě je průměr, který nespadá do nejčetnějšího sloupce. Tento jev je předpověditelný, vezmeme-li v potaz zakřivení grafu.

Nicméně graf c) disponuje dobrými výsledky, co se odlehlých hodnot týče, jelikož všechny spadají do stanoveného intervalu. V grafu a) i b) tomu tak již není, protože dřívější graf má část odlehlých hodnot v záporné straně histogramu, pozdější pak v té kladné části. Pokud bychom porovnali hodnoty průměrů ve všech třech případech, opět nejlepších výsledků dosáhla poloha a), tedy přímá poloha s hodnotou 2,56 metrů. Druhý nejlepší výsledek je v poloze c), tedy v záklonu se vzdáleností 2,40 metrů a nejhůře dopadla poloha c) s hodnotou 2,32 metrů.

5.3 Vyhodnocení a diskuze

Při pohledu na Obrázku 52 můžeme jednoduše posoudit, v které situaci algoritmus nejlépe vyhodnotil rozpoznání obličejů. Obecně lze říci, že měření bylo úspěšnější při denních měřeních, kdy bylo mnohem více skutečně pozitivních záznamů než v nočních měřeních. Počet skutečně negativních vyhodnocení byl v obou případech značně podobný. Důvodem tolika chybných vyhodnocení může být fakt, že použitý algoritmus byl trénován na obličejích v denním světle, a tudíž při noční detekci není schopen korektně nalézt obličejové markanty a převést je na vektorový formát. Další možnou přispívající proměnnou může být nedostačující rozlišovací schopnost kamery při nočním přísvitu, a tedy by k potenciální zlepšení pomohlo



Obrázek 52 Graf ROC prostoru s výsledky nočního a denního měření [autor]

kvalitnější rozlišení snímacího zařízení. Neočekávaným výsledkem je větší úspěšnost rozpoznání při natočení do strany oproti přímé poloze, kdy je záběr obličeje ve středu obrazu, nikoliv na okraji, a nepodléhá tedy případné deformaci či zkreslení.

Vzdálenosti, při kterých byl detekován obličej v porovnání denních a nočních měření, jsou si poměrně blízké, nicméně při nočních měření dosahuje průměr ve všech polohách nižších hodnot než za denního světla. Rozdíly se pohybují v rozmezí deseti až dvaceti centimetrů pro jednotlivé polohy. Příčinou toho je zjevně slabší infračervený přísvit, který je integrovaný v samotné kameře. Jednoduchým řešením by bylo přidání výkonnější IR diody při užívání nočního režimu kamery.

Další nepřesnosti v měření vzdálenosti může také způsobovat fakt, že při měření musely testované osoby reagovat na povel zastavení při detekci obličeje. To má za následek opožděné reakce, kdy nelze počítat s okamžitým zastavením na daném místě. Tato skutečnost se může následně promítnout do měřených hodnot, a to jejich zkreslením vlivem tohoto provedení situace. Možným způsobem, jak získat přesnější hodnoty, by bylo pořízení laserového metru, u kterého by byla možnost připojení na jeden z pinů, kterými je Raspberry osazeno. Tento metr by pak dostal signál z běžícího skriptu při detekci obličeje a téměř okamžitě a mnohem efektivněji by změřil vzdálenost.

V závěru lze tedy zhodnotit denní výkonnost jako znatelně lepší, a tudíž lze uvažovat o použití v systému fungující za denního světla. Dále je vhodné zmínit spolehlivost systému, co se týče rozpoznání neautorizované osoby (specifickost), což je poměrně kritický parametr jakéhokoliv bezpečnostního systému. Ta byla na poměrně vysoké hodnotě, na rozdíl od citlivosti, která v nočním režimu klesala na nízké hodnoty. Nicméně citlivost udává pouze komfort používání, takže není příliš důležitá, pokud hodnotíme systém ze security hlediska. Jako řešením tohoto problému by se jevilo možné přetrénování klasifikátoru, aby byla zvýšena jeho schopnost rozpoznávání v nočním režimu.

6 Finanční vyhodnocení

Tato poslední kapitola je věnována finančnímu vyhodnocení celého projektu. Jedná se o výčet nákladů na pořízení jednotlivých komponentů, následuje představení potenciálního komerčního konkurenta a v poslední části jejich srovnání. Smyslem tohoto porovnání je uvedení finanční náročnosti v kontextu komerčních výrobků s cílem ukázat možnosti sestrojení domácího zařízení jako alternativy.

6.1 Raspberry pi

Do nákladů pro realizaci tohoto projektu bude začleněn samotný jednodeskový počítač Raspberry Pi 4 Model B ve verzi 8 GB RAM, kamera značky Waveshare RPi IR-CUT, oficiální Raspberry Pi USB-C 5,1 V=3 A napájecí zdroj. Dále je nutné započítat SD adaptér s kartou Raspberry Pi 64 GB microSDXC třídy 10 a posledním komponentem je chladící sada Zonepi galaxy, jež obsahuje pasivní chladič, větrák a držák na něj. Veškeré komponenty byly zakoupeny od oficiálního dodavatele Raspberry Pi pro Českou republiku RPishop.cz. Uvedená cena je v korunách českých.[19]

Tabulka 4 Souhrn finančních nákladů komponentů pro sestavení kamerového systému [autor]

Název komponentu	Cena komponentu
Raspberry Pi 4 Model B - 8 GB RAM	2 249,00 Kč
Waveshare RPi IR-CUT kamera	699,00 Kč
Raspberry Pi USB-C 5,1 V=3 A napájecí zdroj	247,00 Kč
Raspberry Pi 64 GB microSDXC třída 10 + SD adaptér	403,00 Kč
Zonepi chladicí sada	279,00 Kč
Souhrn	3877,00 Kč

V Tabulce 4 lze vidět souhrn nákladů jednotlivých komponentů, stejně jako pak jejich individuální ceny. Konečná cena kamerové sestavy tedy činí 3 877 korun českých. Co se týče nákladů na softwarové straně práce, tak jsou prakticky neexistující. Veškerý vývoj byl prováděn

na open-source platformách a stejně tak i programovací jazyk Python je open-source, a tudíž zdarma pro jakékoliv využití. Totéž platí i pro využité knihovny jako Dlib či OpenCV.

6.2 Komerční konkurence

Jako potenciální ekvivalenty v komerční sféře byly zvoleny tři modely chytrých kamer. Modely byly vybírány na základě podobných cenových relací a také ekvivalentního vybavení s podobnou mírou výkonu. První model je od firmy Google s názvem Nest Cam IQ. Jedná se o chytrou kameru s rozměry 73,66 mm v šířce a 124,46 mm ve výšce. Zorný úhel se pohybuje okolo 130°. Rozlišení kamery je pak na úrovni Full HD (1080p). Noční dosvit výrobce neuvádí, nicméně rozpoznání obličejů je uveden v nočním režimu do tří metrů. Pro rozlišení známých a neznámých osob kamera používá software Nest Awarem který umožňuje ukládání všech pořízených záznamů z kamery na cloudové úložiště. Problém této služby je ten, že funguje na měsíční subskripci a je tedy nutné ji kupovat pro její využívání. Kamera je také schopna se spojit s chytrým telefonem a disponuje jak Bluetooth spojením, tak WiFi připojením k lokální síti. Kameru pak lze vidět na Obrázku 53.[39]



Obrázek 53 Kamera Google Nest Cam IQ [39]

Druhý model je také od firmy Google. Jedná se o chytrý video zvonek s názvem Google Nest Hello. Jde o kameru o rozměrech 43 mm na šířku a 117 mm na výšku. Zorný úhel objektivu je uváděn na 130° a rozlišení obrazu je pouze HD. Dosvit infračerveného záření, ani vzdálenost rozpoznání obličejů výrobce neuvádí. Stejně jako v předešlém případě i tato kamera používá Nest Aware službu, která realizuje detekci obličejů v obraze. Obraz je pak možné přenášet v živém přenose či se podívat na záznam až tři hodiny nazpět. Možná konektivita je realizována pomocí WiFi sítě. Velkou nevýhodou tohoto zařízení je právě napájení bateriemi. Kameru lze vidět na Obrázku 54.[50]



Obrázek 54 Kamera Google Nest Hello [50]

Třetím zvoleným modelem je kamera od firmy Netatmo s názvem Netatmo Welcome. Kamera je válcovitého charakteru, kdy podesta má průměr 45 mm a kamera je vysoká 155 mm. Netatmo Welcome disponuje kamerou se 4 MPx, mající zorný úhel 130° a poskytující obraz ve FullHD. Zařízení disponuje infračerveným přísvitem, nicméně opět není výrobcem uveden jeho dosvit, nebo na jakou vzdálenost je kamera schopna rozpoznat obličeje. Výhodou této kamery oproti předešlým je fakt, že software rozpoznávání obličejů je zahrnut v ceně kamery, a tudíž není třeba vytvářet subskripci k takové službě, jak to bylo řešeno u Googlu. Kamera pak podporuje ukládání na lokální úložiště ve formě SD karty až do 32 GB. Mimo tuto formu úložiště také podporuje i cloudové řešení. Záznam je prováděn pouze tehdy, když je detekován

pohyb a tím tedy předchází prázdným záběrům. Kamera nabízí i živý přenos. Konektivita Netatmo Welcome je zajištěna pomocí ethernetového kabelu nebo WiFi připojení a též disponuje aplikací pro její správu. Napájení kamery je řešeno pomocí sítě přes USB kabel. Tento model kamery lze vidět na Obrázku 55.[51]



Obrázek 55 Kamera Netatmo Welcome [51]

6.3 Komparace

Jednotlivé kamery byly vybrány v podobných cenových relacích pro zajištění relevantní komparace parametrů produktů. Srovnání specifických vlastností kamer lze vidět v Tabulce 5. Při pohledu do tabulky lze jednoduše vidět, že model vytvořený v rámci práce je nejlevnější variantou ze všech prezentovaných modelů, dosahující částky 3 677 Kč. Kamera na bázi Raspberry Pi je schopna konkurovat v kvalitě videa, úložišti, či živém přenosu videa a nočního vidění. Stejně tak je kamerový systém (vytvořený autorem práce) schopný nabídnout vzdálený přístup, podobně jako konkurenti a spolehlivější napájení z elektrické sítě. Nezanedbatelná je také konektivita, která je rozmanitější než u konkurentů. Značnou výhodou je také platební model, kdy po zakoupení komponentů není již třeba žádné subskripce a je možné nahrávky bezpečněji ukládat na lokální úložiště, a nikoliv na cloudovou službu. Jediné kategorie, ve kterých není tento kamerový systém ve vedení, je zorný úhel a fakt, že není příliš vhodný do vnějšího prostředí.

Tabulka 5 Srovnání parametrů kamerových systémů [autor]

Parametry	Raspberry Pi	Nest Cam IQ	Nest Hello	Netatmo Welcome
Cena	3 877 Kč	9 420 Kč	7 031 Kč	5 299 Kč
Zorný úhel	50°	130°	130°	130°
Kvalita videa	FullHD	FullHD	HD	FullHD
Lokální úložiště	SD karta (64 GB)	Není	Není	SD karta (32 GB)
Cloudové úložiště	Ano	Ano	Ano	Ano
Noční vidění	Ano (až 3 m)	Ano	Ano	Ano
Živý přenos	Ano	Ano	Ano	Ano
Vzdálený přístup	Ano	Ano	Ano	Ano
Platební model	Jednorázová koupě	Koupě + předplatné	Koupě + předplatné	Jednorázová koupě
Konektivita	WiFi, Ethernet, Bluetooth	WiFi, Bluetooth	WiFi	Ethernet, WiFi
Napájení	Elektrická síť	Elektrická síť	Baterie	Elektrická síť
Prostředí užití	Vnitřní	Vnitřní	Vnější	Vnitřní

Pokud bychom chtěli srovnání v kontextu normativních požadavků, které byly rozebrány v první kapitole, stojí srovnávané modely na podobné úrovni. U kamerového systému vytvořeného v rámci práce lze těžko stanovit jednotlivé třídy: Moselo by dojít k rozsáhlým testům; nicméně lze s jistotou konstatovat, že by kamera mohla spadat do vnitřní třídy prostředí, bez větších komplikací i obecné kategorie. U kamer vyráběných Googlem není

striktně určena třída prostředí, nicméně výrobce uvádí vnitřní využití pro Nest Cam IQ a vnější využití pro Nest Hello, která má být schopna odolat vnějším vlivům počasí, jelikož je zamýšleno její umístění na vchodové dveře. Netatmo Welcome je také kamerou určenou pro domácí vnitřní využití. Lze tedy předpokládat, že všechny kamery, s výjimkou Nest Hello, budou na podobné úrovni tříd prostředí. Ve srovnání přístupových úrovní lze jednoznačně označit za vítěze systém na bázi Raspberry Pi, poněvadž se jedná o plnohodnotný počítač s operačním systémem, tudíž je snadné vytvořit jednotlivé úrovně přístupu, které mohou sahat až na úroveň výrobce, a tedy lze snadno modifikovat celý systém dle potřeby. Oproti tomu ostatní tři kamery nabízí pouze přístup na úrovni administrátora (3 úroveň) a povolují jen některé konfigurace systému, kterými nelze systém příliš upravit dle potřeby. Další požadavky z normy jsou požadavky na ukládání. Při náhledu do Tabulka 1 a pak do Tabulka 5 lze vidět, že kamery Netatmo Welcome a Raspberry Pi by bez problémů splňovaly podmínky pro druhou bezpečnostní třídu. Při lehkých modifikacích Raspberry Pi jej bez problému zařadit i do požadavků kladených na bezpečnostní třídu tři. Zařazení kamer od Googlu do této třídy je diskutabilní, protože neposkytují spolehlivé lokální úložiště a je u těchto modelů třeba využít clouдовých služeb.

Závěr práce

Diplomová práce se zabývala tématem využitelnosti jednodeskového počítače Raspberry Pi jako základem pro bezpečnostní kamerový systém s využitím biometrické technologie rozpoznání obličeje. Cílem práce bylo jak hardwarové, tak softwarové řešení detekčního systému za využití jednodeskového počítače, následné zhodnocení spolehlivosti, efektivity a také finančního zatížení v porovnání s komerčními produkty.

Teoretická část práce je uvedena literární rešerší, kde jsou představeny stěžejní publikace v kontextu práce, a také je doplněna normativními požadavky na kamerové systémy. Dále pokračuje základními pojmy a definicemi, kde autor práce představil Raspberry Pi a uvedl stručně problematiku strojového učení, na jehož základech je postavena softwarová část práce.

V praktické části jsou nejdříve uvedeny použité prostředky, jež byly nezbytné pro sestrojení finální podoby kamerového systému. Jsou zde popsány jednotlivé komponenty, využité detektory a použitý jazyk, společně s popisem veškerých nezbytných skriptů pro dosažení cíle. Následně je představena metodika experimentální části práce, kde autor uvádí jednotlivé sledované veličiny společně s laboratorními podmínkami a způsoby vyhodnocení výsledků. Dále jsou představeny výsledky jak nočních, tak denních měření a práci pak uzavírá finanční vyhodnocení. Cíle práce, jež byly dříve stanoveny, byly tímto naplněny.

Výsledky práce ukazují, že použité detekční metody vykazují spolehlivé výkony v rámci denního režimu, a tím zde vzniká značný potenciál využití v bezpečnostní praxi. Toto provedení má však dvě úskalí. Prvním je fungování v nočním režimu, kde testované algoritmy dosahují ne příliš pozitivních výsledků. Druhým je omezená výpočetní síla počítače, které měla za cíl lehkou kompromitaci hladkého průběhu fungování detekčních skriptů. Obě tyto skutečnosti pak otevírají možnosti rozšíření tématu práce ve smyslu optimalizace systému. Diplomová práce také může sloužit jako stručný úvod do problematiky detekce obličejů za pomoci strojového učení, případně také jako průkaz proveditelnosti pro nasazení do bezpečnostní praxe.

Použitá literatura

- [1] HATTERSLEY, Lucy. *The Official Raspberry Pi Handbook 2022*. Londýn: Seymour Distribution, 2022. ISBN 978-1-912047-78-9.
- [2] MONK, Simon. *Raspberry Pi Cookbook*. Sebastopol: O'Reilly Media, 2014. ISBN 978-1-449-36522-6.
- [3] BROWNLEE, Jason. *Deep Learning for Computer Vision: Image Classification, Object Detection and Face Recognition in Python* [online]. V1.4. San Juan: Machine Learning Mastery, 2019 [cit. 2022-04-12]. Dostupné z: <https://machinelearningmastery.com/deep-learning-with-python/>
- [4] MANGANELLO, Fabio. *Computer Vision with Maker Tech: Detecting People With a Raspberry Pi, a Thermal Camera, and Machine Learning*. Berkeley, CA: Apress, 2021. ISBN 978-1484268209.
- [5] POOLE, Matthew. *Building a Home Security System with Raspberry Pi*. Birmingham, UK: Packt Publishing, 2015. ISBN 978-1-78217-527-8.
- [6] NADAF, Raju A, S.M. HATTURE, Vasudha M BONAL a Susen P NAIK. Home Security against Human Intrusion using Raspberry Pi. *Procedia Computer Science* [online]. 2020, (167), 1811-1820 [cit. 2022-04-12]. ISSN 1877-0509. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S1877050920306657>
- [7] SAJJAD, Muhammad, Mansoor NASIR, Fath U MIN ULLAH, Khan MUHAMMAD, Arun Kumar SANGAIAH a Sung Wook BAIK. Raspberry Pi assisted facial expression recognition framework for smart security in law-enforcement services. *Information Sciences* [online]. 2019, (479), 416-431 [cit. 2022-04-12]. ISSN 0020-0255. Dostupné z: <https://www.sciencedirect.com/science/article/abs/pii/S0020025518305425?via%3Dihub>
- [8] MCBRIDE, Wesley J. a Jason R. COURTER. Using Raspberry Pi microcomputers to remotely monitor birds and collect environmental data. *Ecological Informatics* [online]. 2019, (54) [cit. 2022-04-12]. ISSN 1574-9541. Dostupné z: <https://www.sciencedirect.com/science/article/abs/pii/S1574954119303279?via%3Dihub>

- [9] BERLE, Ian. *Face Recognition Technology: Compulsory Visibility and Its Impact on Privacy and the Confidentiality of Personal Identifiable Images* [online]. Cham, CHE: Springer, 2020 [cit. 2022-04-12]. ISBN 978-3-030-36887-6. Dostupné z: <https://link.springer.com/book/10.1007/978-3-030-36887-6>
- [10] SMITH, Marcus a Seumas MILLER. *Biometric Identification, Law and Ethics* [online]. Cham, CHE: Springer, 2021 [cit. 2022-04-12]. ISBN 978-3-030-90256-8. Dostupné z: <https://link.springer.com/book/10.1007/978-3-030-90256-8>
- [11] ŠČUREK, Radomír. *Speciální bezpečnostní technologie na ochranu osob a majetku* [online]. Ostrava: Vysoká škola báňská – Technická univerzita Ostrava, 2015 [cit. 2022-04-12]. Dostupné z: https://www.fbi.vsb.cz/export/sites/fbi/060/.content/galerie-souboru/studijni-materialy/Specialni_bezpecnostni_technologie.pdf
- [12] ŠČUREK, Radomír. *Biometrické technologie – technické prostředky bezpečnostních služeb* [online]. Ostrava: Vysoká škola báňská – Technická univerzita Ostrava, 2015 [cit. 2022-04-12]. ISBN 978-80-248-3786-4. Dostupné z: <https://www.fbi.vsb.cz/export/sites/fbi/060/.content/galerie-souboru/studijni-materialy/BiometrickeTechnologie.pdf>
- [13] AMR, Tarek. *Hands-On Machine Learning with scikit-learn and Scientific Python Toolkits: A practical guide to implementing supervised and unsupervised machine learning algorithms in Python*. Birmingham, UK: Packt Publishing, 2020. ISBN 978-1838826048.
- [14] DUDEK, Martin. Histogram. *Kvalita jednoduše* [online]. 11.4.2016 [cit. 2022-04-12]. Dostupné z: <http://kvalita-jednoduse.cz/histogram/>
- [15] Raspberry Pi trademark rules and brand guidelines. In: *Raspberry Pi* [online]. Raspberry Pi Foundation [cit. 2022-04-12]. Dostupné z: <https://www.raspberrypi.com/trademark-rules/>
- [16] Raspberry Pi 400. *RPishop* [online]. Roudné: RPishop.cz [cit. 2022-04-12]. Dostupné z: https://rpishop.cz/400/3152-645-raspberry-pi-400.html#/178-rozlozeni_klavesnice-uk
- [17] Raspberry Pi. *Opensource* [online]. Red Hat, 2022 [cit. 2022-04-12]. Dostupné z: <https://opensource.com/resources/raspberry-pi>

- [18] Raspberry Pi OS. *Raspberry Pi* [online]. Raspberry Pi Foundation [cit. 2022-04-12]. Dostupné z: <https://www.raspberrypi.com/software/>
- [19] Raspberry Pi 4 Model B – 8GB RAM. *RPishop* [online]. Roudné: RPishop.cz [cit. 2022-04-12]. Dostupné z: <https://rpishop.cz/raspberry-pi-4b/2611-raspberry-pi-4-model-b-8gb-ram-0765756931199.html>
- [20] Blink a LED. *Microsoft technical documentation* [online]. Microsoft, 2022 [cit. 2022-04-12]. Dostupné z: <https://docs.microsoft.com/en-us/dotnet/iot/tutorials/blink-led>
- [21] AI vs. Machine Learning vs. Deep Learning. *7wData* [online]. Diest, BEL, 2015, 19.5.2019 [cit. 2022-04-12]. Dostupné z: <https://7wdata.be/big-data/ai-vs-machine-learning-vs-deep-learning/>
- [22] Machine Learning: What it is and why it matters. *SAS* [online]. Cary (North Carolina): SAS Institute, 2022 [cit. 2022-04-12]. Dostupné z: https://www.sas.com/en_us/insights/analytics/machine-learning.html#
- [23] WAKEFIELD, Katrina. A guide to machine learning algorithms and their applications: Understanding the types of machine learning algorithms and when to use them. *SAS* [online]. Cary (North Carolina): SAS Institute, 2022 [cit. 2022-04-12]. Dostupné z: https://www.sas.com/en_us/insights/articles/analytics/machine-learning-algorithms-guide.html
- [24] Machine Learning In a Nutshell: When, Why and How. *8allocate* [online]. Tallinn: 8allocate, 16.1.2019 [cit. 2022-04-12]. Dostupné z: <https://8allocate.com/article/machine-learning-in-a-nutshell-when-why-and-how/>
- [25] BABU, Alan Davis. Artificial Intelligence vs Machine Learning vs Deep Learning (AI vs ML vs DL). *Medium* [online]. Medium, 5.11.2019 [cit. 2022-04-12]. Dostupné z: https://medium.com/@alanb_73111/artificial-intelligence-vs-machine-learning-vs-deep-learning-ai-vs-ml-vs-dl-e6afb7177436
- [26] Neural Networks In a Nutshell. *Medium* [online]. Medium, 14.2.2021 [cit. 2022-04-12]. Dostupné z: <https://medium.com/analytics-vidhya/neural-networks-in-a-nutshell-bb013f40197d>
- [27] ELMAHMUDI, Ali a Hassan UGAIL. A framework for facial age progression and regression using exemplar face templates. *The Visual Computer* [online]. 2021, (37)

- [cit. 2022-04-12]. Dostupné z: https://www.researchgate.net/publication/343699139_A_framework_for_facial_age_progression_and_regression_using_exemplar_face_templates
- [28] What is Facial Recognition Technology and How Does it Work?. *Medium* [online]. Medium, 29.9.2020 [cit. 2022-04-12]. Dostupné z: <https://medium.com/@mygreatlearning/what-is-facial-recognition-technology-and-how-does-it-work-9fdefd335c3b>
- [29] Waveshare RPi IR-CUT Camera. *RPishop* [online]. Roudné: RPishop.cz, 2022 [cit. 2022-04-12]. Dostupné z: <https://rpishop.cz/mipi-kamerove-moduly/1366-waveshare-rpi-ir-cut-camera.html>
- [30] What Is Python Used For? A Beginner's Guide. *Coursera* [online]. Coursera, 2022, 21.3.2022 [cit. 2022-04-12]. Dostupné z: <https://www.coursera.org/articles/what-is-python-used-for-a-beginners-guide-to-using-python>
- [31] *Python* [online]. 2001 [cit. 2022-04-12]. Dostupné z: <https://www.python.org/>
- [32] KING, Davis E. Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research* [online]. 2009, (10), 1755-1758 [cit. 2022-04-12]. Dostupné z: <https://jmlr.csail.mit.edu/papers/volume10/king09a/king09a.pdf>
- [33] High Quality Face Recognition with Deep Metric Learning. *Dlib C++ Library* [online]. 12.2.2017 [cit. 2022-04-12]. Dostupné z: <http://blog.dlib.net/2017/02/high-quality-face-recognition-with-deep.html>
- [34] GEITGEY, Adam. Face Recognition. *Face Recognition Documentation* [online]. Read the Docs, 2017 [cit. 2022-04-12]. Dostupné z: <https://face-recognition.readthedocs.io/en/latest/readme.html#>
- [35] OpenCV About. *OpenCV* [online]. OpenCV team, 2022 [cit. 2022-04-12]. Dostupné z: <https://opencv.org/about/>
- [36] SAHA, Sumit. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. *Towards Data Science* [online]. 15.12.2018 [cit. 2022-04-12]. Dostupné z: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

- [37] MALLICK, Satya. Histogram of Oriented Gradients explained using OpenCV. *LearnOpenCV* [online]. BIG VISION, 2022, 6.12.2016 [cit. 2022-04-12]. Dostupné z: <https://learnopencv.com/histogram-of-oriented-gradients/>
- [38] MITTAL, Aditya. Haar Cascades, Explained. *Medium* [online]. Medium, 21.12.2020 [cit. 2022-04-12]. Dostupné z: <https://medium.com/analytics-vidhya/haar-cascades-explained-38210e57970d>
- [39] Google Nest Cam IQ. *Alza* [online]. Praha 7 – Holešovice: Alza.cz, 1994 [cit. 2022-04-12]. Dostupné z: <https://www.alza.cz/google-nest-cam-iq-d4983436.htm#>
- [40] ČSN EN 62676-1-1. *Dohledové videosystémy pro použití v bezpečnostních aplikacích – Část 1-1: Systémové požadavky – Obecně*. Praha: Úřad pro technickou normalizaci, metrologii, a státní zkušebnictví, 2014, 48 s.
- [41] KOKOL, Peter. *Introduction to Data Mining and Knowledge Discovery*. 3rd ed. Potomac: Two Crows Corporation, 1999. ISBN 1-892095-02-5.
- [42] GOODFELLOW, Ian, Yoshua BENGIO a Aaron COURVILLE. *Deep learning*. Cambridge: MIT Press, 2016. Adaptive computation and machine learning (MIT Press). ISBN 978-026-2035-613.
- [43] LE CALLET, P., C. VIARD-GAUDIN a D. BARBA. A Convolutional Neural Network Approach for Objective Video Quality Assessment. In: *IEEE Transactions on Neural Networks* [online]. 2006, s. 1316-1327 [cit. 2022-04-12]. ISSN 1045-9227. Dostupné z: doi:10.1109/TNN.2006.879766
- [44] DALAL, N. a B. TRIGGS. Histograms of Oriented Gradients for Human Detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* [online]. IEEE, 2005, 886-893 [cit. 2022-04-12]. ISBN 0-7695-2372-2. Dostupné z: doi:10.1109/CVPR.2005.177
- [45] VIOLA, Paul a Michael J. JONES. Robust Real-Time Face Detection. *International Journal of Computer Vision* [online]. 2004, 57(2), 137-154 [cit. 2022-04-12]. ISSN 0920-5691. Dostupné z: doi:10.1023/B:VISI.0000013087.49260.fb
- [46] PEARSON, Karl. X. Contributions to the mathematical theory of evolution.—II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of*

London. (A.) [online]. 1895, 186, 343-414 [cit. 2022-04-12]. ISSN 0264-3820.
Dostupné z: doi:10.1098/rsta.1895.0010

- [47] UPTON, Graham J. G. a Ian COOK. *A dictionary of statistics. Reprinted with corrections* 2011. Oxford: Oxford University Press, 2011. ISBN 978-019-9541-454.
- [48] FAN, Jerome, Suneel UPADHYE a Andrew WORSTER. Understanding receiver operating characteristic (ROC) curves. *CJEM* [online]. 2006, 8(01), 19-20 [cit. 2022-04-12]. ISSN 1481-8035. Dostupné z: doi:10.1017/S1481803500013336
- [49] POWERS, David. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. *Mach. Learn. Technol.* [online]. 2008, (2) [cit. 2022-04-12]. Dostupné z: https://www.researchgate.net/publication/228529307_Evaluation_From_Precision_Recall_and_F-Factor_to_ROC_Informedness_Markedness_Correlation
- [50] Google Nest Hello chytrý video zvonek. *Ab-com* [online]. Hradec Králové: AB COM CZECH, 2003 [cit. 2022-04-12]. Dostupné z: <https://www.ab-com.cz/google-nest-hello-chytry-video-zvonek/>
- [51] NETATMO WELCOME. *Electro World* [online]. Electro World [cit. 2022-04-12]. Dostupné z: <https://www.electroworld.cz/netatmo-nsc01-eu-bezpecnostni-kamera>

Seznam obrázků

Obrázek 1 Schéma rozdělení stupňů zabezpečení	11
Obrázek 2 Logo Raspberry Pi	15
Obrázek 3 Raspberry Pi 400	16
Obrázek 4 Raspberry Pi 4 Model B – pohled shora	16
Obrázek 5 Rozložení GPIO pinů u nejnovější generace modelů	17
Obrázek 6 Schéma oboru umělé inteligence	19
Obrázek 7 Schéma průběhu učení s učitelem	19
Obrázek 8 Schéma průběhu učení bez učitele	21
Obrázek 9 Schéma učení posilováním	22
Obrázek 11 Fungování iterací upravování vah	24
Obrázek 10 Příkladná struktura umělé neuronové sítě	24
Obrázek 12 Schéma všech veličin pro výpočet výstupu	25
Obrázek 13 Schéma exemplárního reálného vstupu do funkce	26
Obrázek 14 Příklad z knihovny dlib - (a) označené obličeiové markanty, (b) rozmístění a očíslování jednotlivých bodů	28
Obrázek 15 Fotografie Raspberry Pi 4 s pasivním chladičem	29
Obrázek 16 Aktivní chladící zařízení pro Raspberry Pi	30
Obrázek 17 Fotografie zapojení jednotlivých komponentů	31
Obrázek 18 Schéma zapojení kamerového systému včetně legendy nalevo	31
Obrázek 19 Kamera Waveshare RPi IR-CUT	32
Obrázek 20 Výsledky testů vzdálenosti detekce obličeje v noci	34
Obrázek 21 Stojan pro kameru – a) zadní pohled, b) čelní pohled	35
Obrázek 22 Fotografie základny (a) a držáku na kameru (a)	36
Obrázek 23 Plně osazený stojan na kameru	36
Obrázek 24 Struktura nejdůležitějších knihoven při stavbě skriptu	39
Obrázek 25 Příklad rozboru obrázku na jeho barevné vrstvy	41
Obrázek 26 Konvoluce obrázku	42
Obrázek 27 Exemplární výsledky sdružovacích metod	43
Obrázek 28 Filtrační jádra pro gradienty [37]	44
Obrázek 29 Příklad možného rozdělení gradientů a jejich hodnot – vpravo rozvržený obrázek na buňky, uprostřed gradienty v buňce, vlevo hodnoty gradientu	45

Obrázek 30 Způsob rozčlenění gradientů do histogramu	46
Obrázek 31 Speciální případ rozdelení hodnot gradientu	46
Obrázek 32 Histogram gradientů.....	47
Obrázek 33 Vizualizace HOG deskriptoru.....	47
Obrázek 34 Příklady možných Haar vlastností	48
Obrázek 35 Ilustrace fungování integrálních obrazců	49
Obrázek 36 Reprezentace fungování Adaboostu.....	50
Obrázek 37 Průběh skriptu build_face_dataset.py	51
Obrázek 38 Schéma průběhu fungování skriptu faces_videostream_raspberry.py.....	54
Obrázek 39 Flowchart průběhu měření jedné osoby	56
Obrázek 40 Nákres vzdáleností měření – a) rozložení denního měření, b) rozložení nočního měření, c) segmentace prostoru při pohledu shora	57
Obrázek 41 Fotografie umístění kamerového systému – a) pohled z boku, b) čelní pohled....	58
Obrázek 42 Fotografie označené plochy pro měření.....	58
Obrázek 43 Fotografie kolmě přímé polohy A.....	59
Obrázek 44 Fotografie vytočení kamery do strany, poloha B	60
Obrázek 45 Fotografie polohy v záklonu, polohy C	60
Obrázek 46 Exemplární matice záměn	62
Obrázek 47 Exemplární graf ROC prostoru	64
Obrázek 48 Matice záměn pro denní měření – a) přímá pozice, b) rotace do strany, c) záklon	67
Obrázek 49 Histogram četností pro denní měření – a) přímá pozice, b) rotace do strany, c) záklon	68
Obrázek 50 Matice záměn pro noční měření – a) přímá pozice, b) rotace do strany, c) záklon	69
Obrázek 51 Histogram četností pro noční měření – a) přímá pozice, b) rotace do strany, c) záklon	70
Obrázek 52 Graf ROC prostoru s výsledky nočního a denního měření	71
Obrázek 53 Kamera Google Nest Cam IQ	74
Obrázek 54 Kamera Google Nest Hello	75
Obrázek 55 Kamera Netatmo Welcome	76

Seznam tabulek

Tabulka 1 Požadavky na ukládání	12
Tabulka 2 Přístup k funkcím systému na základě přístupové úrovně	12
Tabulka 3 Výsledky testů detekčních metod	55
Tabulka 4 Souhrn finančních nákladů komponentů pro sestavení kamerového systému ...	73
Tabulka 5 Srovnání parametrů kamerových systémů	77