GYMNÁZIUM JÍROVCOVA 8

MATURITNÍ PRÁCE

Rozpoznávání obličeje pomocí strojového učení

Lukáš Lacina
vedoucí práce: Dr. rer. nat. Michal Kočer

V Českých Budějovicích 14. února 2025 **2024/2025**

Prohlášení

Prohlašuji, že jsem tuto práci vypracoval samostatně s vyznačením všech použitých pramenů.

V Českých Budějovicích dne ..... podpis .....

Lukáš Lacina

Abstrakt

Klíčová slova

Poděkování

Obsah

Úvod

Část I

Rozpoznávání obličeje a strojové učení

1 Úvod do rozpoznávání obličeje

1.1 Definice technologie rozpoznávání obličeje

Technologie rozpoznávání obličeje je technologií, která dokáže detekovat a extrahovat lidskou tvář z digitálního obrazu a poté porovnat tuto tvář s databází předem identifikovaných tváří. Tato technologie se obecně rozděluje na 3 formy dle její funkce: [1, 25]

- **One-to-One Identification**: Tato forma rozpoznávání obličeje porovnává obličej jedné osoby s předem identifikovanou tváří v databázi. Nejčastěji se využívá pro autentizaci uživatelů, například při odemykání mobilních zařízení. [25]
- **One-to-Many Identification**: Tato forma umožňuje technologie identifikovat konkrétní osobu mezi mnoha dalšími tím, že porovná obličej s rozsáhlou databází identit. Používá se zejména k masovému sledování. [25]
- **Emotion and Demographic Recognition**: Tato forma zpracování obličeje je nejpokročilejší a zaměřuje se na analýzu rysů obličeje tak, aby následovně odhadla demografické charakteristiky, jako je věk, pohlaví nebo emoční stav konkrétní osoby. [25]

1.2 Způsob fungování

Rozpoznávání obličeje je složitý proces, který zahrnuje několik klíčových kroků, které lze shrnout následovně: [1, 18]

1. **Shromažďování dat:** Prvním krokem je shromáždění dat z digitálních snímků. Tyto snímky mohou pocházet z různých zdrojů, jako jsou fotografie, videa nebo bezpečnostní kamery. [1, 18]

2. **Detekce obličeje:** V této fázi dochází za pomoci složitých algoritmů k identifikaci a lokalizaci obličeje v obraze. Tento krok zahrnuje analýzu obrazu za účelem určení oblastí, kde se nacházejí obličeje. [1, 18]

3. **Extrakce rysů:** Jakmile je obličej detekován, následuje extrakce klíčových rysů obličeje, jako jsou vzdálenost mezi očima, tvar lícních kostí a délka čelisti. Tyto rysy se dále převádějí do matematické reprezentace, která umožňuje jejich snadnější analýzu a porovnání. [1, 18]

4. **Porovnání:** Extrahované rysy se poté porovnávají s uloženými daty v databázi obličejů. Tato fáze zahrnuje hodnocení podobnosti mezi jednotlivými obličeji, což je klíčové pro určení identity osob. [1, 18]

1.3 Význam a využití

Význam technologie rozpoznávání obličeje je v dnešní době větší, než si většina lidí uvědomuje. Nejčastěji se využívá **v bezpečnosti**, kdy tato technologie umí ověřovat identitu jednotlivců, například při přihlašování k různým službám či zařízením. Také však hledá pohřešované osoby nebo odhaluje osoby podezřelé. Newyorská policie například uvedla, že díky technologii rozpoznání obličeje dokázala chytit pachatele do 24 hodin od útoku. Bezpečnost je hlavní důvod, proč se technologie rozpoznávání obličeje stává populární, avšak poskytuje mnohem více přínosů. [22]

Rozpoznávání obličeje se používá **v marketingu** pro zefektivnění různých procesů. Například v obchodě Amazon Go stačí zákazníkovi k nákupu pouze projít obchodem, vzít si zboží a odejít. Dále se používá **v sociálních médií**, kde jsou platformy schopny automaticky označit uživatele na fotografiích. V posledních pár letech se však tato technologie objevuje i **ve zdravotnictví**, kde je schopna sledovat zdravotní stav pacientů a tím zvýšit efektivitu poskytování zdravotní péče. [22]

1.4 Zneužití technologie rozpoznávání obličeje

Ačkoli má technologie rozpoznávání obličeje mnoho přínosů, může být také zneužita. Jeden z nejzásadnějších problémů je **nelegální sledování osob**. Například některé autoritářské

režimy využívají tuto technologii k monitorování a potlačování opozičních skupin, což je narušení základních lidských práv a svobod. [10]

Dalším velkým problémem jsou **kybernetické útoky**, kdy hackeři využívají data o obličeji, aby mohli provádět různé podvody, například odemykání zařízení nebo bankovních účtů. Jelikož změnit biometrické údaje je v podstatě nemožné, jejich odcizení může mít závažné následky. Tyto útoky často vedou k ztrátě finančních prostředků nebo ke ztrátě soukromí. [10]

Z těchto důvodů jsou technologie rozpoznávání obličeje regulovány tak, aby se co nejvíce zamezilo zmíněným problémům. Zavádí se právní rámce a etické standardy pro jejich používání. Avšak bez těchto regulací může tato technologie způsobit více škody než užitku. [22]

2 Historie a vývoj rozpoznávání obličeje

2.1 Počátky technologie rozpoznávání obličeje

Snaha o vytvoření technologie rozpoznávání obličeje sahá již do roku **1964**, kdy americký vědec **Woodrow Wilson Bledsoe** přišel s myšlenkou vytvořit stroj, který bude rozpoznávat lidské tváře. A tak Bledsoe společně s Helen Chan Wolfovou a Charlesem Bissonem **vytvořil poloautomatickou metodu na rozpoznávání obličeje**. Bledsoeova metoda rozpoznávala rysy manuálně, vědci museli identifikovat celkově 20 různých měřítek, jako je vzdálenost mezi očima, délka nosu nebo šířka rtů. Tyto parametry byly poté využity k vytvoření vektorové reprezentace obličeje, kterou počítač porovnával. V roce **1977** byl systém vylepšen o dalších 21 parametrů pro zlepšení přesnosti. [8, 29]

Bledsoeova metoda byla sice průlomová, avšak měla spousty výrazných omezení jako její **nízkou rychlost či častou nepřesnost**. Nepřesnost byla způsobována variabilitou osvětlení, úhly pohledu a individuálními rysy obličejů, což omezovalo její praktické užití. [8, 29]

2.2 Automatizace a První algoritmy

2.2.1 Metoda Eigenfaces a PCA

V roce **1988** se začala uplatňovat statistická metoda **PCA (Principal component analysis)** v oblasti počítačového vidění. Tato metoda byla schopna redukovat potřebné množství dat a extrahovat klíčové rysy v obličeji, které do této doby určovali vědci manuálně. [1, 8]

Díky metodě PCA vyvinul **Matthew Turk a Alex Pentland** v roce **1991** novou revoluční metodu rozpoznávání obličeje - **metoda Eigenfaces**. Tato metoda byla schopna detekovat obličeje na snímcích, což vedlo k **první automatické technologii rozpoznávání obličeje**. Tato metoda byla jednoduchá a měla relativně vysokou přesnost, což vedlo

k jejímu rychlému rozšíření. Nicméně tomuto průlomu bránily technologické a environmentální faktory, jako například omezený výpočetní výkon a kvalita databází, ze kterých se algoritmus učil. [2, 8]

2.2.2 Projekt FERET

V roce **1993** přišla agentura **DARPA** (Defense Advanced Research Projects Agency) s programem **FERET** (Facial Recognition Technology), který měl za cíl vyvinout **rozsáhlou standardizovanou databázi obličejů**. Tento projekt vznikl jako reakce na rostoucí zájem o technologie rozpoznávání obličeje a jejich aplikace v bezpečnosti a identifikaci. [7, 8, 30]

Databáze FERET obsahovala přes **14 000 snímků obličejů**, které byly foceny **v rozdílných podmínkách**, včetně osvětlení, úhlů a výrazů. Tato rozmanitá databáze umožnila testování a trénování různých algoritmů rozpoznávání obličejů na kvalitním datasetu, který je klíčový pro spolehlivou funkčnost algoritmu. Projekt FERET také přinesl **standardizované metodologie pro hodnocení výkonu algoritmů**, což umožnilo porovnávat efektivitu různých přístupů. [7, 30]

Díky tomuto projektu se zlepšila efektivita a přesnost technologií rozpoznávání obličeje a FERET se stal základem pro budoucí aplikace. [7]

2.3 Strojové učení a jeho vliv v technologii rozpoznávání obličeje

Strojové učení je podmnožina umělé inteligence, která má schopnost se učit nebo predikovat žádané stavy. Strojové učení se začíná v oblasti rozpoznávání obličejů uplatňovat **již v 90. letech** - poprvé se objevilo **v metodě Eigenfaces**. Strojové učení umožňovalo algoritmům učit se z velkého množství dat, což způsobovalo výrazné zlepšení přesnosti a spolehlivosti rozpoznávání obličeje. [11, 14]

Avšak k významnému pokroku došlo až **na začátku 21. století**, kdy se začíná aplikovat technika **hlubokého učení**. [11, 14]

2.3.1 Hluboké učení a konvoluční neuronové sítě

Hluboké učení je technika strojového učení, která používá **konvoluční neuronové sítě (CNN)**. CNN je umělá neuronová síť, která je navržena tak, aby efektivně zpracovávala a analyzovala obrazová data. Tato technika se začala v rozpoznávání obličejů používat **v roce 2012**, kdy **Alex Krizhevsky a jeho tým vytvořili síť AlexNet**. AlexNet vyvolal velký příliv zájmu a investic do metod hlubokého učení a díky tomu vznikají během následujících 4 let nové revoluční algoritmy - **DeepFace, VGGFace, DeepID a FaceNet**, které posunuly technologii rozpoznávání obličeje na novou úroveň. [11, 14]

2.3.2 CNN modely

Tyto modely byly trénovány na několika milionech snímků, měly vysokou rychlost zpracování a také vysokou úspěšnost. Například model **DeepFace** od společnosti Meta měl úspěšnost správného rozpoznání obličeje **přes 97 %**, což je stejná přesnost, jakou má člověk. [14, 28]

2.4 iPhone X a Face ID

Technologie rozpoznávání obličeje se začala používáním hlubokého učení rychle zdokonalovat a **12. září 2017 představila společnost Apple iPhone X, první iPhone s funkcí Face ID**. Zařízení bylo uvedeno na trh 3. listopadu téhož roku. **Face ID** je biometrická metoda, která pomocí **3D skenování** obličeje umožňuje odemknout iPhone a nabízí tak vysokou úroveň bezpečnosti. Poprvé v historii přineslo Face ID pokročilou technologii 3D rozpoznávání obličeje přímo do spotřebitelského zařízení, což nastavilo nový standard v oblasti zabezpečení. Před Face ID bylo rozpoznávání obličeje u spotřebitelských zařízení převážně pouze 2D, a tedy ne moc bezpečné, jelikož ho bylo možné oklamat například fotografií. [8, 23]

V dnešní době je Face ID tak dokonalé, že rozpozná obličej přes čepici, šátky, optické brýle, kontaktní čočky, roušku či růst vousů. **Úspěšnost Face ID je 99,99 %** a kromě odemykání zařízení umí také ověřovat platby, přihlašovat uživatele do aplikací nebo vytvářet animované emoji (Animoji a Memoji) na základě pohybů obličeje. [23]

2.5 Technologie rozpoznávání obličeje dnes

Dnes se technologie rozpoznávání obličeje stále více opírá o pokroky v hlubokém učení a konvolučních neuronových sítích. Moderní algoritmy, jako jsou DeepFace, FaceNet a Face ID, dosahují vysoké přesnosti rozpoznávání obličeje i v extrémních podmínkách, což posunulo technologii na novou úroveň. Tyto modely jsou schopny v některých aspektech překonávat lidské schopnosti, avšak stále narážejí na problémy, jako je variabilita osvětlení, úhly pohledu a odlišnosti v obličejových výrazech. [25]

3 Umělá inteligence a strojové učení)

Předtím, než se podrobně zaměříme na technologii rozpoznávání obličeje, je klíčové si vysvětlit základní principy umělé inteligence a strojového učení, na kterých tato technologie staví a které jí umožňují dosahovat vysoké přesnosti a spolehlivosti při identifikaci obličejů.

3.1 Umělá inteligence

Umělá inteligence (AI, artificial intelligence) je oblast informatiky, která se zabývá vytvářením systémů, které mají napodobovat lidský mozek a být tedy schopné vykonávat úkoly, které vyžadují lidskou inteligenci. Tyto úkoly zahrnují rozpoznávání a používání řeči, plánování tras, řešení problémů, rozpoznávání obrazů atd. Umělá inteligence se dělí na několik podkategorií, z nichž strojové učení je jednou z nejvýznamnějších. [11]

Hlavním cílem umělé inteligence je vyvinout systémy, které se dokážou samostatně učit a adaptovat na nové situace. To zahrnuje schopnost analyzovat velké množství dat, vyvozovat z nich závěry a zrychlovat se. V kontextu rozpoznávání obličeje se umělá inteligence využívá k trénování modelů, které dokážou efektivně rozpoznávat vzory v obličejích. [11]

3.2 Strojové učení

Strojové učení (ML, machine learning) je podmnožina umělé inteligence, která se zabývá návrhem metod a algoritmů, které umožňují systémům se učit z předchozích zkušeností, tedy z dat. Data jsou základem správného fungování strojového učení a je nutné, aby byla správně zvolená a předzpracovaná, aby algoritmy mohly efektivně fungovat. Například při analýze počasí je potřeba mít dostatek dat, která obsahují minimální a maximální teploty, rychlost větru, objem srážek atd. Na základě takto zpracovaných dat algoritmus identifikuje vzorce a vztahy, podle kterých predikuje výsledky. [11]

Hlavní výhodou strojového učení je, že algoritmy mohou postupně zlepšovat svou přesnost

bez toho, aby bylo nutné je ručně upravovat. Tento proces zlepšování je možný díky adaptivnímu učení, kdy model zohledňuje své chyby v dalších predikcích. S rostoucím množstvím dat a trénováním modelu tedy strojové učení dosahuje stále vyšší přesnosti. [11]

Strojové učení je díky svému autonomnímu zlepšování zásadní v mnoha moderních aplikacích, jako je rozpoznávání obrazu, zpracování přirozeného jazyka, automatizované rozhodovací systémy, diagnostika onemocnění nebo personalizace obsahu na internetu. [11]

Strojové učení lze rozdělit na tři základní typy, z nichž každý má své specifické fungování a algoritmy:

3.2.1 Učení s učitelem (Supervised learning)

Učení s učitelem je metoda, při které je algoritmus trénován na datech obsahujících jak vstupy, tak i odpovídající výstupy. Během trénování model porovnává své predikce s reálnými výstupy a na základě rozdílu upravuje své parametry tak, aby se zvýšila přesnost predikce. Tento typ učení se využívá v úlohách **regrese** (kdy je cílová proměnná spojitého typu, například při odhadování tržní ceny domu) a **klasifikace** (kdy je cílová proměnná kategoriálního typu, například přiřazení obličeje z fotografie ke konkrétní osobě). Tento typ učení se využívá právě u technologie rozpoznávání obličeje, kde se často používají konvoluční neuronové sítě (CNN). Podrobnější informace o CNN se nachází v kapitole 5.3. [11]

3.2.2 Učení bez učitele (Unsupervised learning)

Učení bez učitele, oproti učení s učitelem, neobsahuje žádné informace o struktuře dat. Cílem těchto metod je právě strukturu v datech identifikovat. Struktura se identifikuje především pomocí **shlukování** (kdy se dělí data podle jejich podobnosti) nebo **redukce dimenzionality** (kdy se data přetransformují tak, aby byla zachována jejich nejdůležitější struktura a vzory, což zjednodušuje jejich reprezentaci). Algoritmy učené bez učitele se uplatňují tam, kde nejsou dostupná označená data, například analýza chování zákazníků nebo v doporučovacích systémech, kde se identifikují skupiny s podobnými preferencemi. [11]

3.2.3 Učení posilováním (Reinforcement learning)

Učení posilováním je metoda, při které se algoritmus učí **pomocí zpětné vazby (odměny)**. Algoritmus přijímá pouze dvě formy zpětné vazby - kladné odměny a záporné odměny. Učí

se průběžně a upravuje své parametry, díky využívání zpětných vazeb z okolí, a postupně se optimalizuje jeho chování. Algoritmus využívá **Q-učení** (vyhledává optimální predikci tím, že maximalizuje očekávanou hodnotu odměny v dalších krocích) a **TD-učení** (stroj se učí tak, že zkouší určité akce a upravuje predikce podle dat z minulosti). Učení posilováním se využívá v situacích, kde je klíčová schopnost rozhodování, například hraní her (šachy nebo Go) či autonomní ovládání systémů. [11]

4 Předzpracování dat a datasetu

Aby algoritmus pro rozpoznávání obličeje správně fungoval a dosahoval vysoké přesnosti, je nezbytné správně předzpracovat jak vstupní data, která slouží přímo k rozpoznávání obličejů, tak dataset, na kterém se model trénuje. Tato kapitola se tedy zaměří na procesy předzpracování dat, včetně úprav vstupních dat a přípravy datasetu pro efektivní trénink modelu.

4.1 Předzpracování dat pro modely strojového učení

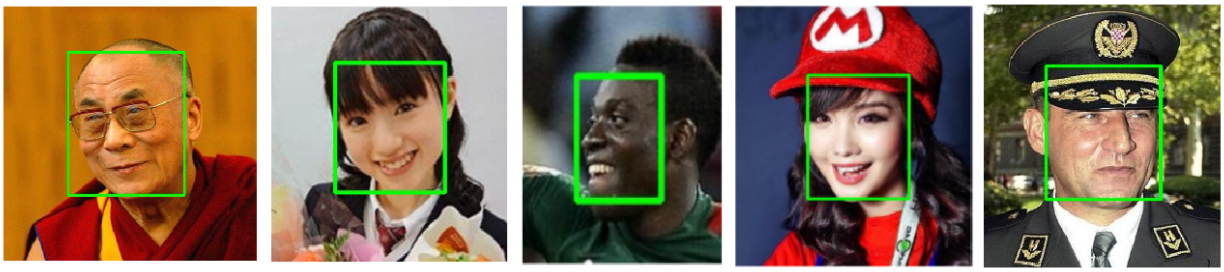
Předzpracování dat je klíčové pro dosažení vysoké přesnosti při zachování co nejkratší doby zpracování. Tento proces zajišťuje, že data jsou konzistentní, správně strukturovaná a optimálně připravená, což minimalizuje riziko chyb a zvyšuje efektivitu modelu.

4.1.1 Detekce obličeje

Prvním krokem předzpracování dat je identifikace oblastí, kde se vyskytuje obličej. Pro tento proces se používají různé metody detekce, které se liší svým přístupem a složitostí. Mezi nejpoužívanější metody patří:

- Haarovy kaskády (Haar Cascades)
- Histogram orientovaných gradientů (HOG)
- DNN (Deep Neutral Networks)
- YOLO (You Only LOOk Once)
- RetinaFace?
- MTCNN

Každá z těchto metod má své výhody a nevýhody a její použití závisí na konkrétních požadavcích aplikace, jako je její rychlost, přesnost nebo odolnost vůči okolnímu šumu.



Obrázek 4.1: Detekované obličeje z obrázků z datasetu WGGFace2

4.1.2 Změna velikosti a normalizace dat

Po detekci oblastí, kde se obličeje nacházejí, je dalším krokem jejich oříznutí, zmenšení a normalizace. Pro zajištění konzistence dat je nutné zajistit, aby každý vyříznutý obličej měl stejný poměr stran a rozlišení. Nejprve se obraz ořízne tak, aby obsahoval pouze oblast s detekovaným obličejem. Přičemž je potřeba zachovat konzistentní poměr stran u všech oříznutých snímků. Oříznutím se zajistí, že se model soustředí výhradně na relevantní část obrazu.

Následně se změní velikost obrázku na jednotnou velikost, například 224×224 pixelů. Tento rozměr je používán jako standardní vstup pro mnoho modelů strojového učení, protože umožňuje optimalizované zpracování a jeho trénink. Také se však stará o kompatibilitu se standardními architekturami hlubokých neuronových sítí, jako jsou ResNet, VGG nebo MobileNet.

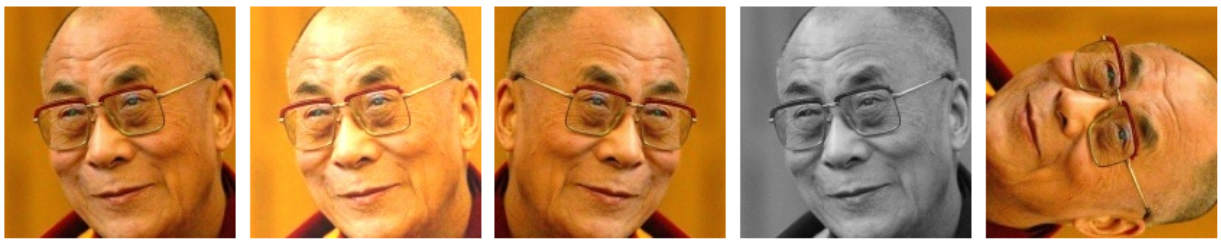
Posledním důležitým krokem je normalizace hodnot pixelů. Většina obrázků je reprezentována hodnotami pixelů v rozsahu 0 až 255, což odpovídá intenzitě jasu jednotlivých bodů. Normalizace tyto hodnoty převede do rozmezí 0 až 1. Tato úprava dat snižuje vliv rozdílů v jasu a kontrastu mezi jednotlivými snímky, což opět napomáhá modelu pracovat s větší přesností.

Tato kombinace kroků – oříznutí, změna velikosti a normalizace – zajišťuje, že data vstupující do modelu jsou konzistentní a připravená pro optimální tréninkový proces a rozpoznávání dat.

Obrázek 4.2: Předzpracované obličeje z obrázků z datasetu WGGFace2

4.1.3 Augmentace dat

Augmentace je technika, která se využívá pro zvýšení rozmanitosti trénovacího datasetu. Tento proces zahrnuje užití různých transformací na původní snímky, čímž se vytvoří nové variace vstupních dat. Augmentace zlepšuje robustnost modelu a jeho schopnost generalizace na nové příklady. Mezi typické metody augmentace dat patří rotace, změna jasu, kontrastu a sytosti, převrácení, geometrická deformace atd.



Obrázek 4.3: Augmentace obličeje z obrázku z datasetu WGGFace2

4.2 Vstupní data pro inferenci

Vstupní data jsou data, která vstupují do již přetrénovaného modelu s účelem rozpoznání osob. Tato data musí vstupovat do modelu již předzpracovaná tak, jak bylo zmíněno: detekce obličeje, jeho oříznutí, změna velikosti a normalizace. Vstupní data lze získat dvěma způsoby, vložením obrázku nebo snímáním obrazu z živé kamery.

4.3 Dataset

Dataset je soubor dat obsahující obrázky obličejů a informace o nich. Účelem datasetu je trénovat a optimalizovat algoritmy rozpoznávání obličejů. Dataset je obvykle organizován do tříd, kdy každé třídě většinou odpovídá identita osoby, avšak mohou jí odpovídat i jiné

charakteristické vlastnosti, jako je věk, pohlaví či emoce. Každá třída poté obsahuje různé variace obličejů, například různé úhly pohledu, osvětlení, výrazy v obličeji nebo zakrytí části obličeje.

Pro efektivní využití datasetu při trénování modelu je nutné jej opět předzpracovat. Tento proces zahrnuje detekci obličeje, jeho oříznutí, změnu velikosti, normalizaci a augmentaci dat. Takto připravený dataset je poté využit k trénování a testování modelu rozpoznávání obličejů.

4.3.1 Datasety pro identifikaci obličejů

Pro trénování modelů zaměřených na identifikaci obličejů se používají různé datasety. Mezi nejznámější datasety patří:

- **LFW (Labeled Faces in the Wild)** - Tento dataset obsahuje přes 13 000 obrázků obličejů 5 749 různých osob. Dataset je vhodný pro testování základních modelů a pro porovnání jejich výkonu v jednoduchých podmínkách, avšak svou velikostí je na trénování robustních modelů pro praktické užití nedostatečný. [13]

- **VGGFace2** - Tento dataset má přes 3 miliony obrázků obličejů 9 131 osob a je velmi populární pro svou velkou variabilitu podmínek, jako jsou různé výrazy a úhly pohledu. Je to jeden z nejrozsáhlejších datasetů pro identifikaci obličejů. [19]

- **Casia-WebFace** - Obsahuje více než 500 000 obrázků obličejů 10 575 osob. Je známý svou rozmanitostí, ale obrázky jsou převážně pořízeny z internetových zdrojů, což může znamenat nižší kvalitu a nedostatečně rozmanité variace dat. [5]

- **MS-Celeb-1M** - Tento dataset je největší svého druhu s více než 10 miliony obrázků obličejů 100 000 sob. Obsahuje obrázky veřejně známých osob a to ve velmi různorodých podmínkách. Díky své velikosti je vhodný pro trénování robustních modelů. [15]

4.3.2 Rozdělení na sady

Dataset pro rozpoznávání obličejů se zpravidla dělí na dvě hlavní sady: **trénovací a testovací**. Trénovací sada (80 % dat) slouží k natrénování modelu, zatímco testovací sada (20 % dat) se používá k hodnocení jeho výkonu na neznámých datech. V některých případech

může být trénovací sada dále rozdělena na **validační sadu**, která se používá k ladění hyperparametrů a prevenci přetrénování. V praxi se však často používá pouze dělení na trénovací a testovací sadu, pokud není potřeba specifické ladění modelu.

5 Algoritmy pro rozpoznávání obličeje

Algoritmus pro rozpoznávání obličeje je jádrem celého rozpoznávacího procesu. Jeho hlavním úkolem je převést vstupní obraz na strukturovanou reprezentaci rysů, nazývanou otisk obličeje, kterou lze snadno porovnávat s reprezentací rysů jiných osob a samotnou klasifikaci získaných dat s databází známých osob. Tato kapitola se zaměřuje na fungování algoritmů, na moderní algoritmy, zejména na konvoluční neuronové sítě, a také na hodnocení výkonu algoritmu.

5.1 Fungování algoritmu

Zásadní proces, který algoritmus provádí, je **extrakce rysů**. Z fotografie se extrahují jedinečné charakteristiky, které odlišují konkrétní obličej od ostatních. Existuje mnoho přístupů k extrakci rysů, avšak v moderních aplikacích dominují metody založené na hlubokém učení. Tyto metody využívají konvoluční neuronové sítě, které se samy učí rozpoznávat důležité rysy obličeje na základě analýzy velkého množství dat. [24]

Výsledkem extrakce rysů je **vektor čísel**, který reprezentuje obličej. Tento vektor je následně porovnáván s vektory ostatních obličejů uložených v databázi, kde se hledá shoda. Čím více se vektory shodují, tím vyšší je pravděpodobnost, že se jedná o stejnou osobu. [24]

5.2 Algoritmy strojového učení

V oblasti rozpoznávání obličejů existuje mnoho různých typů algoritmů, avšak v současné době dominují metody založené na strojovém učení. Tyto algoritmy jsou schopny efektivně analyzovat a porovnávat složité vizuální informace díky své schopnosti učit se z dat. [11, 24]

Zejména metody využívající hluboké učení, jako jsou konvoluční neuronové sítě (CNN, Convolutional Neural Networks), se staly standardem v moderních aplikacích. CNN mají schopnost automaticky se učit extrahovat klíčové rysy obličeje přímo z obrazových dat, bez

nutnosti manuální definice těchto rysů. Díky své architektuře jsou CNN schopny zpracovávat velké množství obrazových dat, identifikovat složité vzory a generalizovat své učení na nové vstupy. [11, 24]

5.3 Rozpoznávání pomocí konvolučních neuronových sítí (CNN)

Konvoluční neuronové sítě jsou jedním z nejvýznamnějších pokroků v oblasti hlubokého učení, zejména v oblasti rozpoznávání obrazu. Konvoluční neuronová síť byla navržena speciálně pro analýzu obrazových dat, kde se klade důraz na klasifikaci objektů. [11]

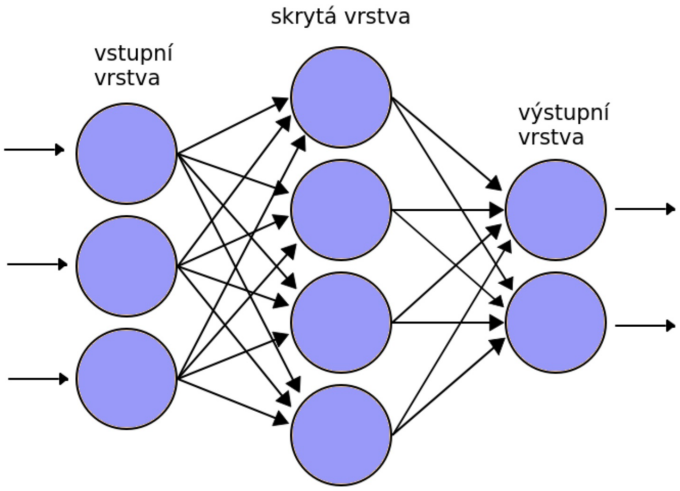
5.3.1 Základy neuronových sítí

Konvoluční neuronové sítě vycházejí z architektury klasických **neuronových sítí (ANN, Artificial Neural Networks)**. ANN jsou složeny z propojených vrstev neuronů, které napodobují činnost biologických neuronů. Díky tomu jsou schopny se přizpůsobit nové situaci. [11]

ANN mají podobnou strukturu a činnost jako lidský mozek. Každá ANN zpracovává daný vstup, například obrázek nebo text, což vede k určitému výstupu, například rozpoznání objektu na obrázku. Podobně jako v mozku jsou ANN tvořeny propojenými umělými neurony, označovanými jako uzly. Tyto uzly spolu komunikují prostřednictvím spojení, která se nazývají hrany sítě. [11]

Jedním z klíčových aspektů fungování neuronových sítí jsou váhy, které reprezentují vztahy mezi jednotlivými uzly. Váhy určují, jak silný vliv má daný vstup na výstup neuronů. Během procesu trénování jsou hodnoty těchto vah upravovány tak, aby model co nejlépe zachytil vztahy ve vstupních datech. Tím se síť postupně přizpůsobuje a zlepšuje svou schopnost správně rozpoznávat vzory a klasifikovat objekty. [11]

V CNN jsou uzly uspořádány do vrstev: do vstupní vrstvy, která přijímá data, jedné nebo více skrytých vrstev, které zpracovávají data, a výstupní vrstvy, která poskytuje výsledek, například určí, co se nachází na obrázku. [11]



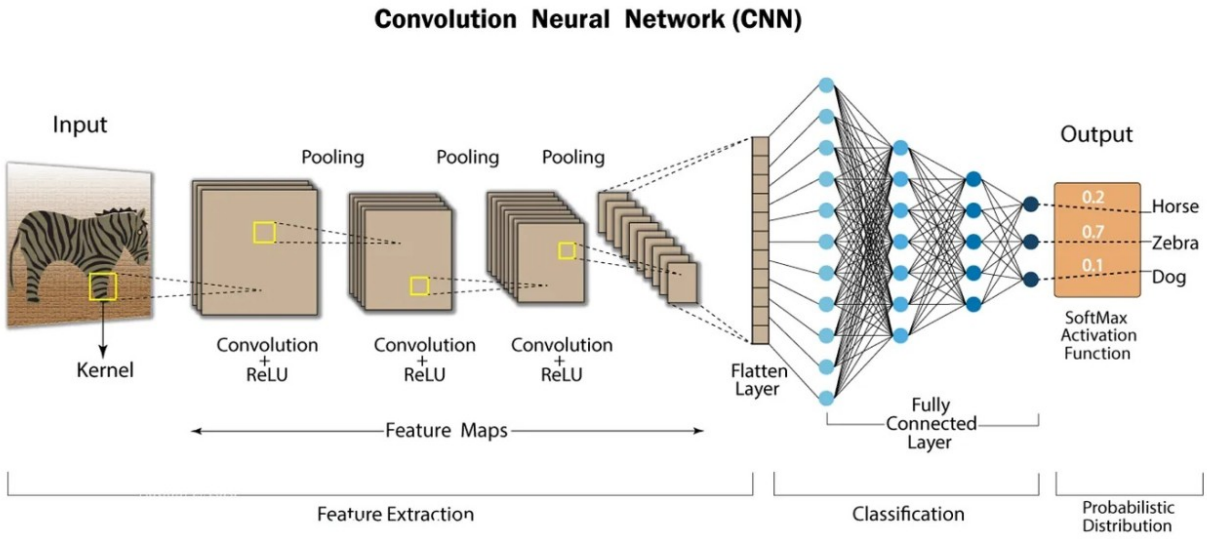
Obrázek 5.1: Struktura umělé neuronové sítě

Ačkoliv ANN jsou úspěšné ve zpracování strukturovaných dat, jejich nevýhodou při práci s obrazovými daty je neschopnost zachytit prostorové vztahy mezi jednotlivými pixely. To vedlo k vývoji CNN, které tuto omezenost překonávají.

5.3.2 Konvoluční neuronová síť

Konvoluční neuronová síť je speciální typ umělé neuronové sítě, která využívá **konvoluční a pooling vrstvy** k efektivnímu zpracování vstupních dat, jako jsou obrázky. Oproti klasickým umělým neuronovým sítím je CNN schopna pracovat s daty velkých rozměrů za použití menšího množství parametrů, díky čemuž je naučení takové sítě správnému fungování mnohem jednodušší. [16, 26, 31]

Kromě konvolučních a pooling vrstev obsahují CNN také **plně propojené vrstvy**, které jsou zodpovědné za finální rozhodnutí modelu, například klasifikaci objektů na obrázcích. Kombinace těchto vrstev umožňuje síti efektivně extrahovat rysy z dat a následně je použít k přesnému rozpoznání a klasifikaci objektů. [16, 26]



Obrázek 5.2: Struktura konvoluční neuronové sítě

5.3.3 Konvoluční a pooling vrstva

Data mohou procházet konvoluční a pooling vrstvou opakovaně v několika vrstvách. S každým dalším opakováním se síť zaměřuje na stále složitější a abstraktnější rysy obrazu, čímž roste její schopnost rozpoznávat konkrétní objekty. [21, 16, 26]

Konvoluční vrstva

Konvoluční vrstva provádí operaci konvoluce, což je matematická operace, jejímž cílem je převést obraz na číselné hodnoty, které může neuronová síť dále analyzovat. [21, 16, 26]

Konvoluce zahrnuje použití filtrů (nebo jader), které jsou menší než celý vstupní obrázek a postupně se posouvají po obraze. Každý filtr detekuje specifické rysy, jako jsou okraje nebo textury, tím, že vynásobí hodnoty pixelů váhami. Součet těchto hodnot je následně zpracován aktivační funkcí. Výstupem této operace je aktivační mapa, která zvýrazňuje detekované rysy. První vrstva obvykle detekuje základní prvky, jako jsou okraje, zatímco hlubší vrstvy se zaměřují na složitější vzory a objekty. Po extrakci těchto rysů jsou výsledky předávány pooling vrstvě. [21, 16, 26]

Pooling vrstva

Pooling vrstva se využívá k redukci rozměrů aktivačních map a tím snižuje výpočetní náročnost sítě. Tento proces zachovává klíčové informace a zároveň zjednodušuje data pro

další zpracování. Nejběžnějšími operacemi v pooling vrstvách jsou max pooling, kdy je vybírána maximální hodnota v určité oblasti, a average pooling, kdy se počítá průměr hodnot v oblasti. [12, 17]

Aktivační funkce

Aktivační vrstva se používá v konvoluční vrstvě, kde upravuje výstupní informace. Jejím úkolem je eliminovat negativní hodnoty z aktivační mapy tím, že je nastaví na nulu, aniž by ovlivnila samotnou konvoluční vrstvu. Tím se snižuje linearita výstupních informací a síť se stává schopná modelovat složitější vzory. V současnosti je nejčastěji používanou funkcí **ReLU (Rectified Linear Unit)**. ReLU nastavuje všechny negativní hodnoty na nulu a zrychluje trénování sítě, i když neřeší chování pro negativní vstupy. [17]

5.3.4 Plně propojené vrstvy

Plně propojené vrstvy (Fully Connected Layers) spojují všechny neurony s předchozími vrstvami a slouží k integraci extrahovaných rysů. Umožňují klasifikaci a rozhodování na základě kombinace informací z celého obrazu. Výstupy této vrstvy určují finální výsledek, například rozpoznání objektu. [21, 16]

5.3.5 Konvoluční neuronové sítě a rozpoznávání obličeje

Konvoluční neuronové sítě se v oblasti rozpoznávání obličeje ukázaly jako vysoce efektivní nástroj díky své schopnosti extrahovat rysy z obrazových dat a následně je použít pro klasifikaci. Pro úspěšnou detekci obličeje je klíčové, jak CNN zachytí různé rysy obličeje, jako jsou oči, nos, ústa a jejich vzájemné vztahy. Tyto informace jsou následně zpracovávány v několika vrstvách, přičemž první vrstvy zachytávají základní rysy (např. okraje), zatímco hlubší vrstvy kombinují složitější informace pro komplexní rozpoznání obličeje.

V dnešní době se pro rozpoznávání obličeje téměř vždy používají algoritmy na bázi CNN. Nejčastěji se používají modely jako VGG-Face, ResNet a Inception. Tyto modely jsou navrženy tak, aby efektivně identifikovaly obličeje v různých podmínkách a dosahovaly vysoké přesnosti při rozpoznávání.

5.4 Hodnocení výkonu algoritmu

Hodnocení výkonu algoritmu pro rozpoznávání obličeje se používá k ověření jeho přesnosti a spolehlivosti. K vyhodnocení výkonu modelu se používají různé metriky jako **přesnost, F1 skóre, či ROC křivka**. Na základě těchto výsledků se poté optimalizují **hyperparametry** modelu, které slouží k nastavení klíčových parametrů ovlivňujících trénování a výkon modelu, jako je rychlost učení nebo počet vrstev. [4]

5.4.1 Metriky pro hodnocení výkonu

Jednou z nejpoužívanějších metrik pro hodnocení výkonnosti modelu je **přesnost (Accuracy)**, která měří podíl správně klasifikovaných vzorků na celkovém počtu vzorků. Tato metrika je jednoduchá a lze snadno pochopit a interpretovat, avšak je potřeba, aby byl počet vzorků v jednotlivých datech vyvážený, aby nebyla tato metrika zavádějící. [4]

F1 skóre představuje **harmonický průměr přesnosti (Precision) a úplnosti (Recall)**. Tato metrika je užitečná zejména v situacích, kdy jsou třídy dat nerovnoměrně zastoupeny, protože kombinuje schopnost modelu správně klasifikovat pozitivní příklady i minimalizovat falešné poplachy. [4]

Dalším nástrojem pro hodnocení je **ROC křivka (Receiver Operating Characteristic)**, která ilustruje vztah mezi pravými pozitivními a falešnými pozitivními případy při různých prahových hodnotách. Klíčovou hodnotou je **AUC (Area Under the Curve)**, která udává schopnost modelu rozlišovat mezi pozitivními a negativními případy. Vyšší hodnota AUC znamená lepší výkon modelu. [4]

5.4.2 Optimalizace hyperparametrů

Hyperparametry jsou parametry, které musí být nastaveny před zahájením trénování modelu, a jejich správná volba může zlepšit přesnost a efektivitu modelu. Mezi běžně laděné hyperparametry patří míra učení, velikost dávky, počet epoch, počet vrstev a neuronů nebo koeficienty regularizace. [6]

Optimalizace hyperparametrů zahrnuje hledání nejlepší kombinace těchto hodnot s cílem dosáhnout co nejlepšího výkonu modelu. Nejčastější metody zahrnují **vyhledávání v mřížce**, kde jsou testovány všechny možné kombinace předdefinovaných hodnot, a náhodné vyhledávání, které vybírá náhodné kombinace parametrů. Pokročilejší metody, jako je **baye-**

sovská optimalizace, využívají pravděpodobnostní modely pro efektivní výběr nejlepších hyperparametrů. [6]

Vhodně nastavené hyperparametry jsou nezbytné pro dosažení vysoké přesnosti a dobré schopnosti modelu generalizovat na nová data. [6]

6 Problematika a etické otázky

Technologie rozpoznávání obličeje je dnes již běžnou součástí života a její přítomnost nelze přehlédnout. Avšak s rozšiřující se a zdokonalující se technologií rozpoznávání obličeje roste i potřeba uvědomit si rizika a nebezpečí spojená právě s touto technologií, která mohou mít zásadní dopad na společnost i jednotlivce. A proto je potřeba si říci něco o ochraně soukromí, možné zaujatosti algoritmů a etických dopadech spojených s technologií rozpoznávání obličeje.

6.1 Ochrana soukromí

6.2 Diskriminace a zaujatost

6.3 Etické otázky...

Část II

Implementace programu na rozpoznávání obličeje

7 Cíl a záměr práce

8 Představa a použité nástroje

9 Struktura programu

10 Vlastní program

10.1 Předzpracování dat

10.2 Vlastní algoritmus

10.3 Přetrénování algoritmu

10.4 Uživatelské rozhraní

Závěr

Bibliografie

1. ADJABI, Insaf; OUAHABI, Abdeldjalil; BENZAOUI, Amir; TALEB-AHMED, Abdelmalik. Past, Present, and Future of Face Recognition: A Review. *Electronics*. 2020, roč. 9, č. 8. Dostupné také z: https://www.mdpi.com/2079-9292/9/8/1188.
2. ANTONIADIS, Panagiotis. *How Do Eigenfaces Work?* 2024. Dostupné také z: https://www.baeldung.com/cs/cv-eigenfaces.
3. AUTORŮ, kolekriv. *JEDNODUŠE: Umělá inteligence*. Universum, 2023.
4. BLOG, Deepchecks Community. *Understanding F1 Score, Accuracy, ROC-AUC, and PR-AUC Metrics for Models*. 2024. Dostupné také z: https://www.deepchecks.com/f1-score-accuracy-roc-auc-and-pr-auc-metrics-for-models/?utm_source=chatgpt.com.
5. *CASIA-WebFace*. [B.r.]. Dostupné také z: https://www.kaggle.com/datasets/debarghamitraroy/casia-webface.
6. *Co znamená ladění hyperparametrů?* 2024. Dostupné také z: https://cs.eitca.org/artificial-intelligence/eitc-ai-gcml-google-cloud-machine-learning/introduction/what-is-machine-learning/what-does-hyperparameter-tuning-mean/.
7. CONTIBUTORS, NIST. *Face Recognition Technology (FERET)*. 2011. Dostupné také z: https://www.nist.gov/programs-projects/face-recognition-technology-feret.
8. CONTRIBUTORS, NEC. *A brief history of Facial Recognition*. 2022. Dostupné také z: https://www.nec.co.nz/market-leadership/publications-media/a-brief-history-of-facial-recognition/.
9. FRANÇOIS, CHOLLET. *Deep learning v jazyku Python: knihovny Keras, Tensorflow. Knihovna programátora*. Grada, 2019.

10. GARGARO, David. *The pros and cons of facial recognition technology*. 2024. Dostupné také z: https://www.itpro.com/security/privacy/356882/the-pros-and-cons-of-facial-recognition-technology.

11. HENDL, Jan. *Big data - Věda o datech, základy a aplikace*. GRADA, 2021.

12. KIŠŠ, Martin. *KONVOLUČNÍ NEURONOVÉ SÍTĚ PRO BEZPEČNOSTNÍ APLIKACE*. 2016. Dostupné také z: https://www.vut.cz/www_base/zav_prace_soubor_verejne.php?file_id=132324. Bakalářská práce. VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ, FAKULTA INFORMAČNÍCH TECHNOLOGIÍ.

13. *Labeled Faces in the Wild*. [B.r.]. Dostupné také z: https://vis-www.cs.umass.edu/lfw/.

14. LYTVYNENKO, Oleksandr. *Machine Learning Algorithms for Face Recognition*. 2022. Dostupné také z: https://codeit.us/blog/machine-learning-face-recognition#what-is-face-recognition.

15. *MS-Celeb-1M (MS1M)*. [B.r.]. Dostupné také z: https://exposing.ai/msceleb/.

16. NELSON, Daniel. *Co jsou to CNN (konvoluční neuronové sítě)?* 2020. Dostupné také z: https://www.unite.ai/cs/what-are-convolutional-neural-networks/.

17. PETROVIČOVÁ, Klára. *Aplikace konvolučních neuronových sítí*. 2019. Dostupné také z: https://nlp.fi.muni.cz/uui/referaty2019/petrovicova_klara/referat.pdf.

18. QINJUN, Li; CUI TIANWEI, Zhao Yan; YUYING, Wu. Facial Recognition Technology: A Comprehensive Overview. *Academic Journal of Computing & Information Science*. 2023.

19. QIONG CAO, LI SHEN, WEIDI XIE, OMKAR PARKHI, ANDREW ZISSERMAN. *VGGFace2 Dataset*. [B.r.].

20. RUDOLF, Pecinovský. *Python - Kompletní příručka jazyka pro verzi 3.11*. Grada, 2022.

21. SERHII, Bondarenko. *Jak funguje VGG16 – neuronová síť pro extrakci obrazových prvků*. [B.r.]. Dostupné také z: https://robotdreams.cz/blog/317-jak-funguje-vgg16-neuronova-sit-pro-extrakci-obrazovych-prvku.

22. SOUKUPOVÁ, Jana. *Možnosti využití technologie rozpoznávání obličejů v kontextu ochrany osobních údajů v EU*. Praha, 2022. Rigorózní práce. Univerzita Karlova, Právnická fakulta, Katedra evropského práva.

23. SUPPORT, Apple. *Informace o vyspělé technologii Face ID*. 2023. Dostupné také z: https://support.apple.com/cs-cz/102381.

24. *Technologie rozpoznávání obličeje podle fotky: Budoucnost, nebo hrozba?* 2024. Dostupné také z: https://prekon.cz/technologie-rozpoznavani-obliceje-podle-fotky-budoucnost-nebo-hrozba/.

25. *The Cambridge Handbook of Facial Recognition in the Modern State*. Cambridge University Press, 2024. Cambridge Law Handbooks.

26. TOXIN. *Konvoluční neuronové sítě pomáhají detekovat a blokovat nevhodný obsah sociálních sítí*. 2021. Dostupné také z: https://www.toxin.cz/blog/blog/konvolu%C4%8Dn%C3%AD-neuronov%C3%A9-s%C3%ADt%C4%9B-pom%C3%A1haj%C3%AD-detekovat-a-blokovat-nevhodn%C3%BD-obsah-soci%C3%A1ln%C3%ADch-s%C3%ADt%C3%AD.

27. VALLA, Tomáš. *Průvodce labyrintem algoritmů*. CZ.NIC, 2022.

28. WIKIPEDIA CONTRIBUTORS. *DeepFace — Wikipedia, The Free Encyclopedia*. 2024. Dostupné také z: https://en.wikipedia.org/w/index.php?title=DeepFace&oldid=1240075590. [Online; accessed 29-October-2024].

29. WIKIPEDIA CONTRIBUTORS. *Facial recognition system — Wikipedia, The Free Encyclopedia*. 2024. Dostupné také z: https://en.wikipedia.org/w/index.php?title=Facial_recognition_system&oldid=1250429409. [Online; accessed 28-October-2024].

30. WIKIPEDIA CONTRIBUTORS. *FERET (facial recognition technology) — Wikipedia, The Free Encyclopedia*. 2024. Dostupné také z: https://en.wikipedia.org/w/index.php?title=FERET_(facial_recognition_technology)&oldid=1232099360.

31. ZACHA, Jiří. *Konvoluční neuronové sítě pro klasifikaci objektů z LiDARových dat*. 2019. Bakalářská práce. České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra kybernetiky.

Zkratky

atd. a tak dále. 10, 15

Seznam obrázků

Seznam tabulek

Přílohy

A Fotky z pokusů

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

B Příloha další