

**SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE  
FAKULTA ELEKTROTECHNIKY A INFORMATIKY**

Evidenčné číslo: FEI-5384-72758

**KLASIFIKÁCIA TVÁRÍ POMOCOU  
KONVOLUČNÝCH SIETÍ PRE ANDROID OS  
DIPLOMOVÁ PRÁCA**

**2018**

**Bc. Ján Polaček**

**SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE  
FAKULTA ELEKTROTECHNIKY A INFORMATIKY**

Evidenčné číslo: FEI-5384-72758

**KLASIFIKÁCIA TVÁRÍ POMOCOU  
KONVOLUČNÝCH SIETÍ PRE ANDROID OS  
DIPLOMOVÁ PRÁCA**

Študijný program: Aplikovaná informatika  
Číslo študijného odboru: 2511  
Názov študijného odboru: 9.2.9 Aplikovaná informatika  
Školiace pracovisko: Ústav informatiky a matematiky  
Vedúci záverečnej práce: Ing. Dominik Sopiak

**Bratislava 2018**

**Bc. Ján Polaček**



## ZADANIE DIPLOMOVEJ PRÁCE

Autor práce: Bc. Ján Polaček  
Študijný program: aplikovaná informatika  
Študijný odbor: 9.2.9. aplikovaná informatika  
Evidenčné číslo: FEI-5384-72758  
ID študenta: 72758

Vedúci práce: Ing. Dominik Sopiak

Miesto vypracovania: Ústav informatiky a matematiky FEI STU

Názov práce: **Klasifikácia tvárí pomocou konvolučných sietí pre Android OS**

Jazyk, v ktorom sa práca vypracuje: slovenský jazyk

Špecifikácia zadania: Cieľom tejto práce je natréňovanie konvolučnej siete a jej aplikácie na problém klasifikácie ľudskej tváre v neriadených podmienkach. Ako prvé je dôležité naštudovať si problematiku konvolučných sietí. Následne je potrebné tieto znalosti aplikovať pri tréňovaní a aplikácie konvolučnej siete na zariadení Android OS.

### Úlohy:

1. Naštudovať teóriu konvolučnej siete.
2. Oboznámiť sa s problematikou klasifikácie tvárí.
3. Navrhnuť štruktúru konvolučnej siete.
4. Vytvoriť navrhnutú konvolučnú sieť.
5. Diskusia získaných výsledkov.

### Literatúra:

- JAIN, A. *Handbook of Biometrics*. London: Springer-Verlag, 2008. ISBN 978-0-387-71040-2.
- GOODFELLOW, M. -- WILLIAMS, S. -- MORDAESKI, M. *Actinomycetes in Biotechnology*. San Diego : Academic Press, 1988. 501 p.
- Vincent Dumoulin, Francesco Visin: "A guide to convolution arithmetic for deep learning", Cornell University, 23. Mar 2016

Dátum zadania: 18. 09. 2017

Dátum odovzdania: 11. 05. 2018

# SÚHRN

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE  
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Študijný program:	Aplikovaná informatika
Autor:	Bc. Ján Polaček
Diplomová práca:	Klasifikácia
	tvárí pomocou
	konvolučných sietí
	pre Android OS
Vedúci záverečnej práce:	Ing. Dominik Sopiak
Miesto a rok predloženia práce:	Bratislava 2018

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean et est a dui semper facilisis. Pellentesque placerat elit a nunc. Nullam tortor odio, rutrum quis, egestas ut, posuere sed, felis. Vestibulum placerat feugiat nisl. Suspendisse lacinia, odio non feugiat vestibulum, sem erat blandit metus, ac nonummy magna odio pharetra felis. Vivamus vehicula velit non metus faucibus auctor. Nam sed augue. Donec orci. Cras eget diam et dolor dapibus sollicitudin. In lacinia, tellus vitae laoreet ultrices, lectus ligula dictum dui, eget condimentum velit dui vitae ante. Nulla nonummy augue nec pede. Pellentesque ut nulla. Donec at libero. Pellentesque at nisl ac nisi fermentum viverra. Praesent odio. Phasellus tincidunt diam ut ipsum. Donec eget est. A skúška mäččňov a dlžňov.

Kľúčové slová: Android

# ABSTRACT

SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA

FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY

Study Programme:	Applied Informatics
Author:	Bc. Ján Polaček
Master's thesis:	Face classification using CNN for Android OS
Supervisor:	Ing. Dominik Sopiak
Place and year of submission:	Bratislava 2018

On the other hand, we denounce with righteous indignation and dislike men who are so beguiled and demoralized by the charms of pleasure of the moment, so blinded by desire, that they cannot foresee the pain and trouble that are bound to ensue; and equal blame belongs to those who fail in their duty through weakness of will, which is the same as saying through shrinking from toil and pain. These cases are perfectly simple and easy to distinguish. In a free hour, when our power of choice is untrammelled and when nothing prevents our being able to do what we like best, every pleasure is to be welcomed and every pain avoided. But in certain circumstances and owing to the claims of duty or the obligations of business it will frequently occur that pleasures have to be repudiated and annoyances accepted. The wise man therefore always holds in these matters to this principle of selection: he rejects pleasures to secure other greater pleasures, or else he endures pains to avoid worse pains.

Keywords: Android

# Pod'akovanie

I would like to express a gratitude to my thesis supervisor.

# Obsah

<b>Úvod</b>	<b>1</b>
<b>1 Rozpoznávanie tváre</b>	<b>2</b>
1.1 Rozdelenie . . . . .	3
1.2 Proces identifikácie . . . . .	4
1.3 Techniky rozpoznávania tvári . . . . .	5
1.3.1 Eigenfaces . . . . .	6
1.3.2 Linear Discriminant Analysis, Fisherfaces . . . . .	6
1.3.3 Independent Component Analysis . . . . .	7
1.3.4 Local Feature Analysis . . . . .	7
1.3.5 Neurónové siete a Support Vector Machines . . . . .	8
1.4 Detekcia tváre . . . . .	8
1.5 Hľadania bodov tváre . . . . .	8
1.6 Meranie úspešnosti . . . . .	8
1.7 Súčasný stav . . . . .	8
<b>2 Neurónové siete</b>	<b>9</b>
2.1 Neurón . . . . .	9
2.2 Aktivačné funkcie . . . . .	11
2.3 Trénovanie neuróvej siete . . . . .	13
2.4 Gradient descent . . . . .	13
<b>Záver</b>	<b>15</b>
<b>Zoznam použitej literatúry</b>	<b>16</b>
<b>Prílohy</b>	<b>I</b>

# Zoznam obrázkov a tabuliek

Obrázok 1	Proces rozpoznávania tváre . . . . .	4
Obrázok 2	Prvých 6 bazových vektorov Eigenfaces, prebraté z[10, p. 45] . .	6
Obrázok 3	Prvých 6 bazových vektorov Fisherfaces, prebraté z[10, p. 46] . .	7
Obrázok 4	Zjednodušený popis fungovania biologického neurónu . . . . .	9
Obrázok 5	Schéma fungovania umelého neurónu . . . . .	10
Obrázok 6	Príklad doprednej neurónovej vrstvy skladajúcej sa z troch vrstiev a tromi neurónmi vo vrstve. . . . .	11
Obrázok 7	Priebeh sigmoidnej funkcie . . . . .	12
Obrázok 8	Priebeh funkcie tanh . . . . .	12
Obrázok 9	Priebeh ReLU funkcie . . . . .	13





# Zoznam algoritmov

# Zoznam výpisov

# Úvod

Tu bude krásny úvod s diakritikou atd.

A možno aj viac riadkový úvod.

# 1 Rozpoznávanie tváre

Rozpoznávanie tváre je jedným z úkonov, ktoré človek robí pravidelne a bez námahy v každodennom živote. Široká dostupnosť silných počítačov za nízku cenu vytvára obrovský záujem o automatické spracovanie digitálneho obrazu v širokom spektre aplikácií, ako sú napríklad biometrická autentifikácia, monitorovanie osôb, interakcia s počítačom či spravovanie multimédií. Výskum a vývoj v oblasti rozpoznávania tváří nasleduje tento trend automaticky.

Hlavnými výhodami využívania rozpoznávania tváří voči iným biometrickým metódam ako napríklad odtlačky prstov či dúhovka, je ich prirodzené a nerušivé používanie, ale najmä možnosť použitia na väčšiu vzdialenosť. Zo šiestich biometrických metód (tvár, odtlačky prstov, dlaň, hlas, dúhovka, podpis) je podľa The International Civil Aviation Organization(ICA0)[1] rozpoznávanie tváre primárnou metódou pri kontrole identity na letisku.

Prvým automatickým systémom na rozpoznávanie tváří bol podľa[2], systém navrhnutý Takeo Kanadem v jeho práci[3] z roku 1974. Za ním nasledovalo obdobie bez výraznejšieho pokroku v automatickom rozpoznávaní tváre, až do roku 1990, kedy Sirovich a Kirby zverejnili článok[4], v ktorom popisujú využitie nízko dimenzionálnych reprezentácií tváre odvodených z Karhunen-Loevovej transformácie alebo Principal Component Analysis PCA. Podľa Jaina v[2] bol ďalším veľkým míľnikom práca[5] od Turka a Pentlanda na vlastných vektoroch tváre (eigenface), ktorá znovu naštartovala výskum v oblasti rozpoznávania tváre. Medzi ďalšie míľniky tiež Jain zaraďuje prácu na Fisherovej metóde[6], ktorá aplikuje Linear Discriminant Analysis(LDA) po aplikovaní PCA k dosiahnutiu vyššej presnosti, výskum na lokálnych Gaborových filtroch[7] k dosiahnutiu efektívnejších príznakov tváre a prístup AdaBoost učenia, založeného na architektúra kaskádneho klasifikátora pre detekciu v reálnom čase[8].

Od obdobia kedy bola navrhnutá Eigenface metóda nastal veľký pokrok v oblasti rozpoznávania tváre. V kontrolovaných podmienkach, kde je možné ovládať svetelnosť, postoj osoby či výraz tváre, prekonáva automatické rozpoznávanie tváre ľudí, a to najmä pokiaľ databáza obsahuje veľké množstvo snímok tváre. Napriek tomu rozpoznávanie tváre stále čelí mnohým výzvam, najmä problematike rozpoznávania tváre v neriadených podmienkach.

## 1.1 Rozdelenie

Ako aj ostatné biometrické systémy, aj rozpoznávanie tváre funguje v jednom alebo oboch z režimov:

- Verifikácia (autentifikácia)
- Identifikácia tváre

Pri verifikácii tváre ide o porovnanie jedna k jednej, čo znamená, že jeden snímok tváre sa porovnáva len jedným záznamom identity v databáze, za ktorú sa prehlasuje. Typickým využitím tohoto režimu je samoobslužná kontrola identity prostredníctvom elektronického pasu[2].

Identifikácia tváre zahŕňa porovnanie jedna k mnohým, čo znamená, že jeden snímok tváre sa porovnáva s viacerými záznamami identít v databáze a vyberie jednu[2]. V niektorých prípadoch využitia je postačujúce nájsť len najpodobnejšiu identitu. V iných prípadoch, ako napríklad sledovanie podozrivých osôb, je okrem nájdania najpodobnejšej tváre potrebné zaviesť aj prah spoľahlivosti, a tie tváre ktoré dosiahli mieru podobnosti väčšiu ako je prah, sú zaznamenané.

Úspešnosť systému na rozpoznávanie tváre závisí vo veľkej miere na množstve variabilných faktorov, ako je osvetlenie, poloha tváre, mimika, vek, make-up, účes, brada či pohyb tváre. Na základe týchto faktorov Jain rozdeľuje[2] na 2 kategórie vzhľadom na miery spolupráce užívateľov:

- Scenár so spolupracujúcim užívateľom
- Scenár s nespolupracujúcim užívateľom

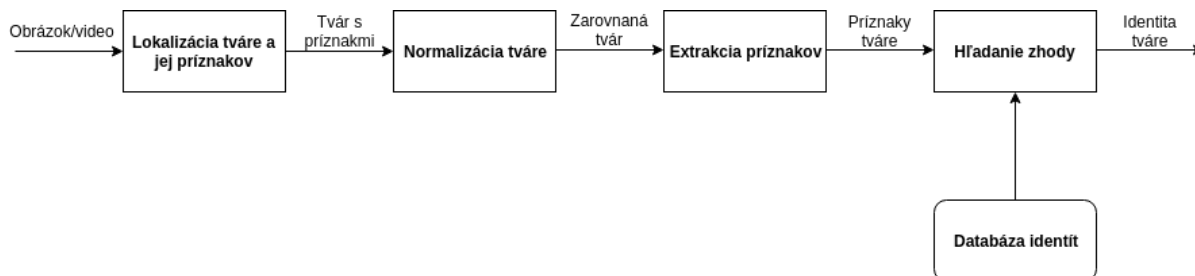
Prípád spolupracujúceho užívateľa je využívaný napríklad pri prihlasovaní do počítača, riadenie fyzického prístupu, elektronické pasy (e-passport), teda prípady kedy má užívateľ záujem spolupracovať na správnom zoznámaní tváre (napr. pod správnym zoznámaním tváre môžeme rozumieť napríklad snímok tváre z predu s neutrálnym výrazom a otvorenými očami) kvôli prístupu alebo povoleniu.

V prípade nespolupracujúceho užívateľa, ktorý je typické pre už spomínané sledovanie podozrivých, si osoba nie je vedomá toho, že podlieha identifikácii. Čo sa týka vzdialenosti medzi tvárou osoby a kamerou, sa v spolupracujúcom scenári využíva krátka vzdialenosť, typicky o jedného metra, a ide teda o oveľa jednoduchšiu úlohu v porovnaní s identifikáciou nespolupracujúcej osoby často na väčšiu vzdialenosť.

## 1.2 Proces identifikácie

Jain popisuje rozpoznávanie tváre takto: “Rozpoznávanie tváre sa zaraďuje medzi problémami rozpoznávania vzorov, kde tvár, ktorá je reprezentovaná ako trojzormerný objekt, podlieha odlišnostiam vo svetle, postoji, výraze a iných faktoroch, potrebuje byť identifikovaná na základe zozbieraných obrázkov“[2]. Zatiaľ čo rozpoznávanie tváre z dvojrozmerného snímku je dnes bežne používané vo väčšine prípadov, v niektorých, najmä tých ktoré vyžadujú vyššiu bezpečnosť, sa využíva trojrozmerný snímok tváre, prípadne snímky tváre mimo bežne viditeľného spektra - napr. termografický snímok tváre. Systém na rozpoznávanie tváre sa podľa Jaina[2] vo všeobecnosti skladá zo štyroch základných častí, ako je ukázané na obrázku 1:

- Detekcia tváre a lokalizácia bodov tváre
- Normalizácia tváre
- Extrakcia príznakov
- Hľadanie zhody tváre



Obr. 1: Proces rozpoznávania tváre

Pri detekcii tváre ide primárne o oddelenie oblasti tváre od pozadia snímku. V prípade videa je potrebné sledovať (track) detekovanú tvár naprieč niekoľkými snímkami videa pomocou komponentu na sledovanie pohybu tváre. Zatiaľ čo detekcia tváre poskytuje len hrubý odhad polohy a veľkosti tváre, lokalizácia bodov tváre najde už konkrétne časti tváre, ako sú napríklad oči, nos, obrys tváre a podobne. Lokalizácia bodov tváre je zväčša vykonávaná osobitným komponentom na lokalizáciu bodov alebo komponentom na vyrovnanie (alignment) tváre.[2]

Normalizácia tváre ide o normalizáciu tváre v geometrickom a fotometrickom zmysle. Tento krok je nevyhnutný, pretože sa od najnovších a najlepších (state-of-the-art) rozpoznávacích metód očakáva, že dokážu rozpoznať tvár v rôznych polohách a rôznom svetle. Geometrická

normalizácia vykonáva transformáciu tváre do štandardného formátu snímku pomocou orezania (crop) tváre. Snímok je následne zakrivený (warp) a upravený (morph) kvôli ešte lepšej a presnejšej normalizácii tváre. Úlohou fotometrickej normalizácie je spracovanie snímku na základe osvetlenie či farebnej škály.[2]

Extrakcia príznakov je vykonávaná na normalizovanom snímku tváre, s cieľom vybrať charakteristické informácie, ktoré sú užitočné pri rozlišovaní tvárí rozdielnych osôb, pričom odolný voči odchýlkam v geometrickej a fotometrickej normalizácii. Extrahované príznaky sú následne použité pri hľadaní zhody s identitou.[2]

Pri hľadaní zhody tváre sa porovnávajú extrahované príznaky zo vstupnej tváre s jednou alebo viacerými tvármi ktoré už sú zapísané v databáze. Výsledkom porovnania s jednou tvárou je výsledkom odpoveď áno alebo nie (verifikácia). V prípade porovnávania s viacerými tvármi je výsledkom indetita vstupnej tváre, za predpokladu, že identita ktorá je nájdená presiahne prah spoľahlivosti, inak vráti informáciu, že tvár je neznáma. V súčasnosti je v tejto oblasti najväčšou výzvou nájsť spoľahlivé meranie vhodné na určenie podobnosti príznakov tváre.[2]

Presnosť systému na rozoznávanie tváre vo veľkom závisí na správnosti extrahovaných príznakov tváre, ktoré naopak, závisia od správnej lokalizácii a normalizácii tváre.

### 1.3 Techniky rozpoznávania tvári

Zhao rozdeľuje[9] algoritmy rozpoznávania do dvoch základných kategórií, vzhľadom na spôsob extrakcie príznakov:

- Metódy založené na príznakoch (feature-based)
- Medódy založené na vzhľade (appearance-based)

Metódy založené na príznakoch využívajú rôzne vlastnosti a geometrické atribúty na popis tváre, ako sú napríklad vzdialenosti či uhly medzi bodmi tváre. Na druhej strane, metódy založené na vhlade využívajú globálne vlastnosti tvárového vzoru. Typickou črtou algoritmov založených na vhlade je výpočet bázových vektorov, na ktoré je následne tvár premietnutá. Koeficienty takejto projekcie sú potom použité k efektívnej reprezentácii údajov tváre[10]. Oblúbené algoritmy ako sú Principal Component Analysis PCA, Linear Discriminant Analysis LDA, Indenpendent Component Analysis ICA, Local Feature Analysis LFA, Manifolds, Correlation Filters alebo Tensorfaces sú založené práve na vhlade tváre. Holistický prístup k rozpoznávaniu tváre má však často problémy pri rôznych polohách tváre[10].



### 1.3.1 Eigenfaces

Podstatou metódy Eigenfaces navrhutej Turkom a Pentlandom[5], tiež známej ako PCA, je hľadanie najmenej kvadratickej chyby lineárneho podpriestoru, ktorý mapuje dáta z originálneho N-rozmerného priestoru na M-rozmerný priestor príznakov, kde  $M \ll N$ . Týmto dosiahneme redukcii dát do M rozmerného priestoru využitím M vlastných vektorov matice kovariance, ktoré korešpondujú s najväčšími hodnotami vlastných čísel[10]. Konečné bázové vektory ktoré sú najvhodnejšie na popis dát, sú nájdené pri procese optimalizácie, ktorej podstatou je maximalizácia variance premietnutých dát. Jain popisuje[10] proces výberu bázových PCA vektorov  $W$  optimalizačnou funkciou(1), kde  $S_T$  označuje úplne maticu rozptylu, ktorá obsahuje kovariancie dát tváre.

$$W_{PCA} = \operatorname{argmax} |W^T S_T W| = [w_1, w_2, \dots, w_m] \quad (1)$$



Obr. 2: Prvých 6 bázových vektorov Eigenfaces, prebraté z[10, p. 45]

Na obrázku2 vidíme príklad vlastných vektorov vybraných metódou Eigenfaces, na obrázkoch po procese normalizácie. Metóda PCA je vhodná na reprezentáciu dát, čo neznamená, že je vhodná aj na rozdeľovanie do tried.

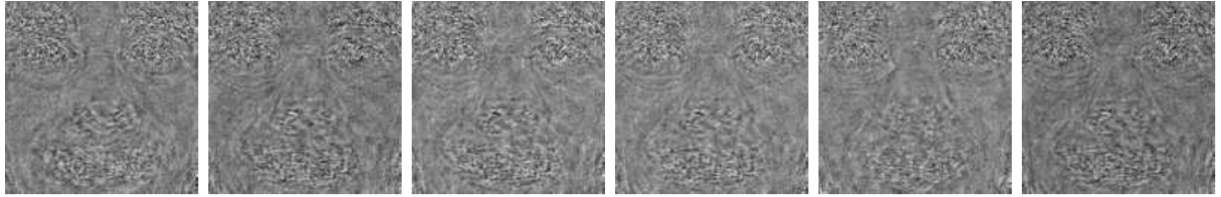
### 1.3.2 Linear Discriminant Analysis, Fisherfaces

Metóda LDA[11] je oveľa vhodnejšia k hľadaniu projekcií, ktoré dobre oddelujú rozdielne triedy. Táto metóda je založená na hľadaní optimálnych projekčných vektorov, ktoré optimalizujú pomer medzitriednych (within class) a vnútrotriednych (between class) vzdialeností - maximalizuje oddelenie tried v premietnutom priestore. Optimálne bázové vektory LDA, môžu byť podľa[10] definované ako rovnica(2), kde  $S_B$  značí medzitriednu maticu rozptylu a kde  $S_W$  vnútrotriednu maticu rozptylu.

$$W_{LDA} = \operatorname{argmax} \frac{|W^T S_B W|}{|W^T S_W W|} \quad (2)$$

Zvyčajne pri riešení problematiky rozpoznávania tváre (a väčšine ostatných problémov rozpoznávania vzorov obrázkov) je množstvo tréningových obrázkov menší ako počet pixelov (dimenzionalita dát), čo znamená, že vnútrotriedová matica rozptylu  $S_W$  je singulárna,

čo je problémom pre LDA. Na vyriešenie tohoto problému singularnej matice sa na najprv vykonáva PCA na redukovanie dimenzionality dát a až následne sa aplikuje LDA v menej rozmernom podpriestore PCA. NA základe týchto zmien bolo dosiahnuté zlepšenie výsledkov v porovnaní s tradičnou PCA metódou. Projekčné vektory, ktorých príklad môžeme vidieť na obrázku3 Fisherfaces sú tie, ktorú spĺňajú maximalizujú vysledok optimalizačnej funkcie(3).



Obr. 3: Prvých 6 bázových vektorov Fisherfaces, prebraté z[10, p. 46]

$$W_{LDA} = \underset{W}{\operatorname{argmax}} \frac{|W^T W_{PCA}^T S_B W_{PCA} W|}{|W^T W_{PCA}^T S_W W_{PCA} W|} \quad (3)$$

### 1.3.3 Independent Component Analysis

Podstatou metódy ICA, je hľadanie takých neortogonálnych báz, aby boli na nich založené transformácie príznakov boli štatisticky nezávislé, zatiaľ čo metóda PCA hľadá také ortogonálne bázy tváre, aby príznaky po transformácii neboli korelované[10]. Bázové vektory tváří metódy PCA sú založené len na štatistike druhého rádu. ICA zovšeobecňuje koncept PCA a vytvára tak model so vzťahmi štatisticky vyššie rádu. Pôvodná motivácia tejto metódy vychádza z potreby zaradiť zvukové vlny do nezávislých zdrojov, bez priamej znalosti procesu spájania (mixovania) týchto vln[10]. Bartlett vo svojej publikácii[12], v ktorej aplikoval metódu ICA na rozpoznávanie tváre uvádza, že ICA dosahuje oveľa lepšie výsledky v porovnaní s metódou PCA, pri rozpoznávaní tváre v rôznych častiach dňa a rôznych výrazoch tváre.

### 1.3.4 Local Feature Analysis

LFA vytvára skupinu lokálne prepojených detektorov príznakov, založených na rozložení charakteristikého podpriestoru. Vzniknú tak minimálne korelované a topograficky indexované podskupiny príznakov, ktoré definujú podpriestor záujmu. Lokálna reprezentácia vytvára odolnosť voči zmenám v lokalizovaných oblastiach objektov. Príznaky použité metódou LFA sú menej náchylné na zmeny vo svetle a zároveň je vďaka nim jednoduchšie odhadnúť rotáciu objektu. Algoritmus LFA bol použitý ako kľúčový komponent algoritmu FaceIt, ktorý je jedným z komerčne využívaných systémov na rozpoznávanie tváre.[10]

### **1.3.5 Neurónové siete a Support Vector Machines**

Neurónové siete NS a Support Vector Machines SVM sú zvyčajne využívané v nízko-dimenzionálnych priestoroch príznakov, najmä kvôli výpočtovej náročnosti spracovania, v prípade viacdimeziionálnych dát tváre[10]. Neurónové siete podliehajú širokému záujmu o výskum, najmä čo sa týka lepšej reprezentácie príznakov tváre a rozpoznávania tváre. Avšak, so zvyšujúcim sa množstvom osôb, ktoré sa neurónová sieť učí rozpoznávať, narastá výpočtová zložitosť exponenciálne. Spojením viacerých neurónových sietí sa podarilo zlepšiť celkový výkon pri rozpoznávaní tváre.

Vo všeobecnosti nevieme povedať, čo presne sa neurónová sieť naučila, alebo akým spôsobom bude fungovať. NS zvyčajne vyžaduje veľké množstvo tréningových dát kvôli dobrej generalizácii dát, s čím zároveň súvisí značné množstvo výpočtov uskutočnených v procese tréningovania. SVM sa ukazuje ako úspešná metóda pri rozpoznávaní objektov, použitím kernelového triku (kernel trick), ktorý mapuje dáta do viacdimeziionálneho priestoru príznakov. SVM v tomto priestore potom nájde nadrovinu, ktorá maximalizuje šírku hranice oddeľujúcej triedy klasifikácie, a znižuje tak riziko zlej klasifikácie nielen pre tréningové dáta, ale tiež k dosiahnutiu lepšej generalizácie neznámych dát.[10]

### **1.4 Detekcia tváre**

### **1.5 Hľadania bodov tváre**

### **1.6 Meranie úspešnosti**

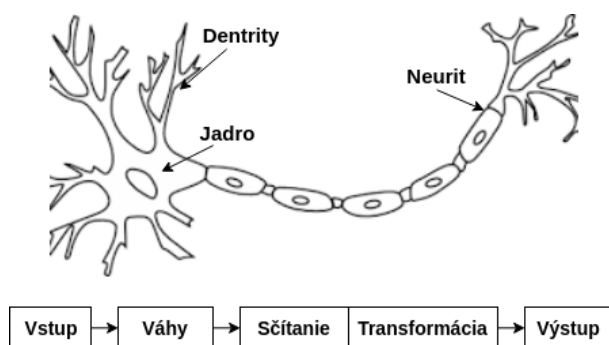
### **1.7 Súčasný stav**

## 2 Neurónové siete

### 2.1 Neurón

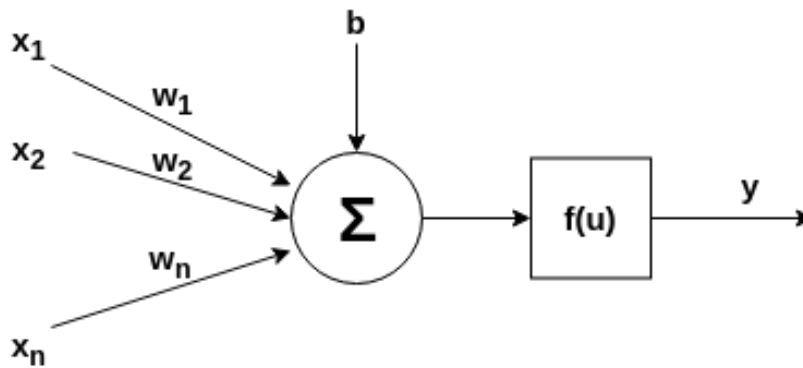
Základnou jednotkou z ktorej sa skladá mozog je neurón. Kúsok mozgu, o veľkosti malého zrnka ryže, obsahuje približne 10000 neurónov, pričom každý neurón vytvára v priemere 6000 prepojení s inými neurónmi, a ide teda v podstate o veľkú biologickú sieť prepojení, ktorá nám dovoľuje rozumieť svetu okolo nás[13].

V samotnej podstate neurónu, prichádza k príjmu informácií z ostatných neurónov, ich spracovaniu a odoslaniu výsledku spracovania ďalším bunkám. Neurón prijíma vstupné informácie pomocou dendritov, čo sú v výbežky na konci neurónu ktoré sa správajú ako antény a zachytávajú signály. Každé z týchto prepojení je dynamicky zosilnené alebo zoslabené, závislosti od frekvencie používania daného spojenia, pričom sila prepojenia definuje, aký veľký prínos má daný dendrit k výstupu neurónu[13]. Po zvážení sily(váhy) prepojení sú vstupné informácie sčítané v jadre neurónu a transformované na nový signál(výsledok), ktorý je cez neurit odoslaný ďalším neurónom. Zjednodušený proces fungovania biologického neurónu vidíme na obrázku4.



Obr. 4: Zjednodušený popis fungovania biologického neurónu

Fungovanie biologického neurónu sa stalo inšpiráciou pre návrh umelého neurónu. Rovnako ako biologický neurón, aj umelý neurón prijíma vstupné údaje -  $x_1, x_2, \dots, x_n$  ktoré sú násobené špecifickou váhou  $w_1, w_2, \dots, w_n$  prislúchajúcou k neurónu. Váhované vstupy sú podobne ako pri biologickom neuróne sčítané, je k nim pripočítaný základ(bias), a následne transformované aktivačnou funkciou, ktorej výsledok je odoslaný ďalším neurónom. Fungovanie neurónu sa dá zapísať ako rovnica(4), kde vstup, a jeho schému vidíme na obrázku5.

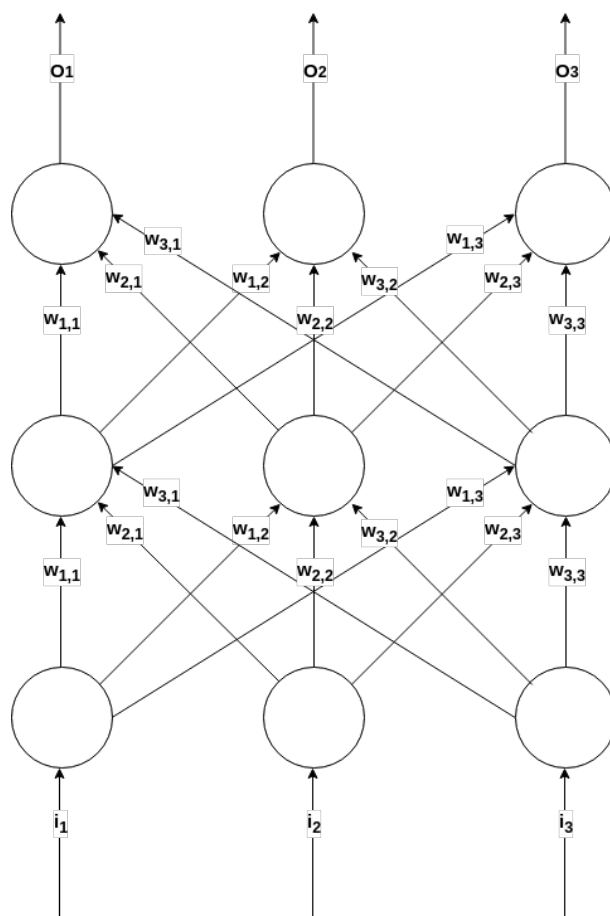


Obr. 5: Schéma fungovania umelého neurónu

$$y = f\left(\sum_{i=1}^n x_i w_i + b\right) \quad (4)$$

Jedným neurónom vieme aproximovať jednoduchú lineárnu funkciu. Na aproximáciu zložitejších problémov, napríklad na rozpoznanie ručne písaných znakov, ich však potrebujeme viac. Ľudský mozog má usporiadané neuróny vo vrstvách. Buduma tvrdí, že mozgová kôra, ktorá je zodpovedná za ľudské myslenie, je tvorená šiestimi takýmito vrstvami[13], a informácia sa prenáša z jednej vrstvy do druhej, až kým vstupná informácia nedáva zmysel. Príkladom je ľudské videnie, kde najspodnejšia vrstva prijíma informácie zo siete oka, táto informácia sa presúva vrstvami mozgovej kôry, až nakoniec v poslednej vrstve rozpozná, o aký predmet ide.

Aplikovaním rovnakého konceptu spájania neurónov do vrstiev, vytvárame umelé neurónové siete. Obrázok jednoduchej neurónovej siete vidíme na obrázku 6, kde spodná vrstva prijíma vstupné dáta a vrchná vrstva neurónov počíta finálny výsledok. Stredná vrstva neurónov je nazývaná skrytá vrstva, kde  $w_{i,j}^{(k)}$  označuje váhu prepojenia medzi  $i$ -tým v  $k$ -tej vrstve a  $j$ -tým neurónom v  $k+1$  vrstve, pričom schopnosť neurónovej siete riešiť problémy priamo závisí na nájdení optimálnych hodnôt váh, pričom vďaka skrytej vrstve je možné vypočítať komplexnejšie vzťahy vo vstupných dátach. V ukážkovej sieti sú existujú prepojenia neurónov v smere do nasledujúcej vrstvy, bez prepojení medzi neurónmi v rovnakej vrstve alebo do predchádzajúcej, tento typ siete sa označuje ako dopredná - feed forward. Lineárne neuróny sú nenáročné na výpočet, avšak akákoľvek dopredná sieť len s lineárnymi neurónmi môže byť vyjadrená ako sieť bez skrytej vrstvy, a teda je limitovaná na naučenie sa zložitejších vzťahov. Riešením je vnesenie nelineárnych vzťahov do neurónovej siete, prostredníctvom aktivačných funkcií.[13]

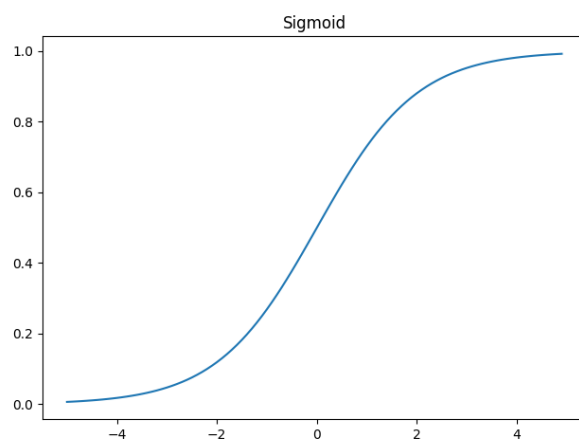


Obr. 6: Príklad doprednej neurónovej vrstvy skladajúcej sa z troch vrstiev a tromi neurónmi vo vrstve.

## 2.2 Aktivačné funkcie

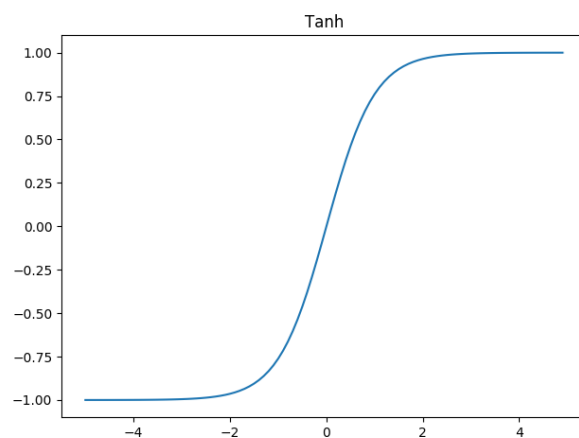
Poznáme 3 základné nelineárne aktivačné funkcie - sigmoid, tanh a relu. Sigmoidná funkcia, je definovaná ako(5), jej priebeh je podobný písmenu s, čo spôsobuje, že ak je vstup veľmi malý, výstup bude blízky 0, a pokiaľ je veľmi veľký, výsledok sa blíži k 1.

$$f(z) = \frac{1}{1 + e^{-z}} \quad (5)$$



Obr. 7: Priebek sigmoidnej funkcie

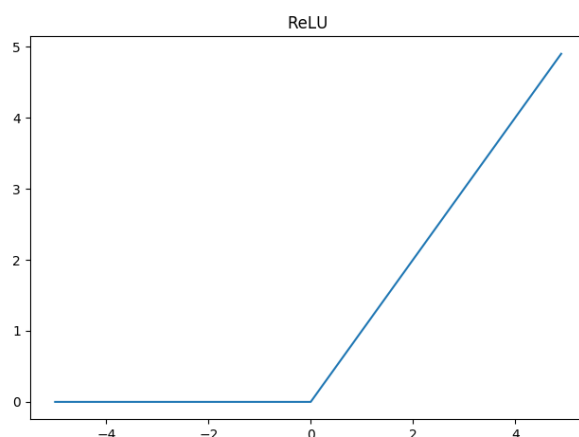
Funkcia tanh - hyperbolický tangens, je podobná sigmoidnej čo sa týka priebehu, s tým že jej rozsah je od -1 do 1, má centrum priebehu v 0, kvôli čomu je často preferovanou aktivačnou funkciou.



Obr. 8: Priebek funkcie tanh

Iným tipom nelineárnej funkcie je ReLU, vyjadrený rovnicou (6), a priebehom ako na obrázku9 a je používanou funkciou najmä v oblasti strojového videnia.

$$f(z) = \max(0, z) \quad (6)$$



Obr. 9: Priebeh ReLU funkcie

## 2.3 Trénovanie neurónovej siete

Pod pojmom trénovania neurónovej siete rozumieme proces optimalizácie vektora váhových prepojení medzi neurónmi. Počas trénovania sa zvyčajne na vstup neurónovej siete prináša veľké množstvo tréningových vzoriek, pričom opakovanne upravujeme hodnoty váh s cieľom minimalizovať chybné výsledky neurónovej siete. Chybu výsledku neurónovej siete vypočítame chybovou funkciou (objective function, error function loss function) (7), a kde  $E$  je výsledná chyba pre  $i$ -tu tréningovú vzorku,  $t^i$  označuje pravdivú hodnotu výsledku pre vzorku  $i$  a  $y^i$  označuje výstup neurónovej siete.

$$E = \frac{1}{2} \sum_i (t^{(i)} - y^{(i)})^2 \quad (7)$$

Výsledok chybovej funkcie je rovný nule, pokiaľ náš model neurónovej siete vracia perfektné výsledky pre každú tréningovú vzorku. Cieľom trénovania je teda optimalizovať váhy neurónov tak, aby sa výsledok chybovej funkcie  $E$  čo najviac približoval k 0.

## 2.4 Gradient descent

Po vypočítaní chybovej funkcie je potrebné upraviť samotné hodnoty váh, pomocou metódy gradient descent. Predpokladajme, že máme funkciu  $y = f(x)$ , kde obe  $x$  a  $y$  sú reálne čísla a derivácia tejto funkcie je označená ako  $f'(x)$ . Derivácia  $f'(x)$  nám dáva sklon alebo smerovanie funkcie  $f(x)$  v bode  $x$ , inými slovami nám udáva, ako upraviť malú zmenu vo vstupe tak, aby sme dosiahli požadovanú funkciu na výstupe (??). [goodfellow2016deep]

$$f(x + e) \approx f(x) + e f'(x) \quad (8)$$

Derivácie sú vhodným prostriedkom na minimalizáciu funkcií, pretože nám udáva, ako



uraviť  $x$  tak, aby sme dosiahli zlepšenie v  $y$ .

# Záver

Conclusion is going to be where?

Here.

# Zoznam použitej literatúry

1. HEADQUARTERS, ACI WORLD. *The Application of Biometrics at Airports*. 2005. (Accessed on 04/04/2018).
2. JAIN, Anil K a LI, Stan Z. *Handbook of face recognition*. Springer, 2011.
3. KANADE, Takeo. Picture processing system by computer complex and recognition of human faces. 1974.
4. KIRBY, Michael a SIROVICH, Lawrence. Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern analysis and Machine intelligence*. 1990, roč. 12, č. 1, s. 103–108.
5. TURK, Matthew a PENTLAND, Alex. Eigenfaces for recognition. *Journal of cognitive neuroscience*. 1991, roč. 3, č. 1, s. 71–86.
6. BELHUMEUR, Peter N., HESPANHA, João P a KRIEGMAN, David J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*. 1997, roč. 19, č. 7, s. 711–720.
7. WISKOTT, Laurenz, KRÜGER, Norbert, KUIGER, N a VON DER MALSBURG, Christoph. Face recognition by elastic bunch graph matching. *IEEE Transactions on pattern analysis and machine intelligence*. 1997, roč. 19, č. 7, s. 775–779.
8. VIOLA, Paul a JONES, Michael. Rapid object detection using a boosted cascade of simple features. 2001, roč. 1, s. I–I.
9. ZHAO, Wenyi, CHELLAPPA, Rama, PHILLIPS, P Jonathon a ROSENFELD, Azriel. Face recognition: A literature survey. *ACM computing surveys (CSUR)*. 2003, roč. 35, č. 4, s. 399–458.
10. JAIN, Anil K, FLYNN, Patrick a ROSS, Arun A. *Handbook of biometrics*. Springer Science & Business Media, 2007.
11. DUDA, Richard O, HART, Peter E a STORK, David G. *Pattern classification*. John Wiley & Sons, 2012.
12. BARTLETT, Marian Stewart, MOVELLAN, Javier R a SEJNOWSKI, Terrence J. Face recognition by independent component analysis. *IEEE Transactions on neural networks*. 2002, roč. 13, č. 6, s. 1450–1464.
13. BUDUMA, Nikhil a LOCASCIO, Nicholas. *Fundamentals of deep learning: designing next-generation machine intelligence algorithms*. Ö'Reilly Media, Inc.", 2017.

# Prílohy