

ŽILINSKÁ UNIVERZITA V ŽILINE  
Fakulta elektrotechniky a informačných technológií

28260420192005

VYHLADÁVANIE KLÚČOVÝCH RYSOV  
V OBRAZE PRE ÚČELY MEDICÍNSKEJ  
DIAGNOSTIKY

DIPLOMOVÁ PRÁCA

2019

Bc. Lucia Golierová

ŽILINSKÁ UNIVERZITA V ŽILINE  
FAKULTA ELEKTROTECHNIKY A INFORMAČNÝCH  
TECHNOLÓGIÍ  
KATEDRA TEORETICKEJ ELEKTROTECHNIKY A  
BIOMEDICÍNSKEHO INŽINIERSTVA

DIPLOMOVÁ PRÁCA

**Študijný program:** Biomedicínske inžinierstvo  
**Študijný odbor:** 5.2.47 Biomedicínske inžinierstvo  
**Školiace pracovisko** Žilinská univerzita v Žiline  
Fakulta elektrotechniky a informačných technológií  
Katedra teoretickej elektrotechniky a biomedicínskeho  
inžinierstva  
**Vedúci bakalárskej práce:** doc. Ing. Dušan Koniar, PhD.

Žilina 2019

Bc. Lucia Golierová



**KATEDRA TEORETICKEJ  
ELEKTROTECHNIKY  
A BIOMEDICÍNSKEHO INŽINIERSTVA**



ŠTUDIJNÝ ODBOR: BIOMEDICÍNSKE INŽINIERSTVO

Univerzitná 1, 010 26 Žilina

Školský rok 2018/2019

## **ZADANIE DIPLOMOVEJ PRÁCE**

**Meno, priezvisko študenta:** Bc. Lucia Golierová

**Školiace pracovisko:** KTEBI

**Názov diplomovej práce:** Vyhľadávanie kľúčových rysov v obraze pre účely medicínskej diagnostiky

**Anotácia diplomovej práce:**

Témou diplomovej práce je návrh, výber a porovnanie algoritmov pre detekciu vybraných kľúčových rysov na ľudskej tvári (oko, špička nosa, brada). Vyhľadané kľúčové body budú následne prenesené do odpovedajúceho 3D modelu hlavy pacienta za účelom automatizovanej cefalometrie a pre podporu diagnostiky obštrukčného spánkového apnoe. Práca by mala overiť využiteľnosť nových postupov strojového učenia v danej problematike.

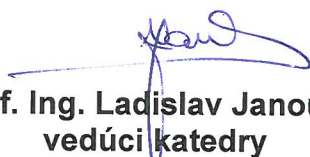
1. Aktuálny stav v oblasti využitia strojového a hĺbkového učenia v BMI aplikáciách / BMI zobrazovaní
2. Špecifikácia kľúčových rysov tváre pre podporu automatizovanej diagnostiky OSA
3. Výber algoritmu pre vyhľadávanie zvolených rysov tváre v 2D obraze
4. Verifikácia a analýza možností spárovania rysov s relevantným 3D modelom

**Školiteľ:** doc. Ing. Dušan Koniar, PhD.

**Konzultant:** Ing. Jozef Volák

**Dátum odovzdania diplomovej práce:** 7.5.2019

Žilina 21.12.2018

  
prof. Ing. Ladislav Janoušek, PhD.  
vedúci katedry

---

## Abstrakt

Zámerom tejto práce zhodnotiť význam strojového a hĺbkového učenia v biomedicínskych aplikáciách a biomedicínskom zobrazovaní a jeho využitie pri vyhľadávaní kľúčových rysov v obraze pre následnú diagnostiku obštrukčného spánkového apnoe. Praktická časť je rozdelená do štyroch častí. Prvá časť je zameraná na definovanie významných bodov tváre pre detekciu predispozície na ochorenie OSA. Úlohou nasledujúcich dvoch častí je zhodnotenie metód použitých pri realizácii vyhľadávania tváre v obraze a detekcie jednotlivých tvárových štruktúr. Posledná časť sa venuje možnostiam spárovania rysov s relevantným 3D modelom.

## Abstract

The aim of this thesis is to evaluate the importance of machine and deep learning in biomedical applications and biomedical imaging and its use in detection of key features in the image for subsequent diagnosis of obstructive sleep apnea. The practical part is divided into four parts. The first part is aimed at definition of significant points of the face for the detection of predisposition to OSA. The task of the following two parts is to evaluate the methods used for face searching in the image and for detection of facial features. The last part deals with the possibilities of matching facial features with the relevant 3D model.

## ANOTAČNÝ ZÁZNAM – BAKALÁRSKA PRÁCA

**Meno a priezvisko:** Lucia Golierová

**Akademický rok:** 2018/2019

**Názov práce:** Vyhľadávanie kľúčových rysov v obraze pre účely medicínskej diagnostiky

**Počet strán:** 45

**Počet obrázkov:** 42

**Počet tabuliek:** 2

**Počet grafov:** 0

**Počet príloh:** 0

**Počet použ. lit.:** 35

### **Anotácia v slovenskom jazyku:**

Táto práca popisuje metódy vyhľadávania kľúčových rysov tváre pre následné využitie pri diagnostike obštrukčného spánkového apnoe. Tento spôsob vyhľadávania je založený na metódach strojového učenia a počítačového videnia. Cieľom práce je vyhľadanie vhodných algoritmov slúžiacich na nájdenie tváre v obraze a následnú detekciu jednotlivých črt potrebných pri diagnostike OSA. Výsledkom je porovnanie použitých metód a zhodnotenie ich úspešnosti pri riešení danej problematiky.

### **Anotácia v anglickom jazyku:**

This paper describes methods for detection of important facial features for subsequent use in diagnosis of obstructive sleep apnea. This method of searching is based on machine learning and computer vision methods. The aim of this work is to find suitable algorithms for finding face in image and subsequent detection of particular features needed for diagnosing of OSA. The result is a comparison of used methods and evaluation of their success in solving of the given issue.

### **Kľúčové slová:**

obštrukčné spánkové apnoe, strojové učenie, počítačové videnie, Dlib, OpenCV

**Vedúci bakalárskej práce:** doc. Ing. Dušan Koniar, PhD.

**Konzultant:** Ing. Jozef Volák

**Recenzent:** \_\_\_\_\_

**Dátum odovzdania práce:** 7. Mája 2019

# Obsah

<b>1</b>	<b>Úvod .....</b>	<b>1</b>
<b>2</b>	<b>Obštrukčné spánkové apnoe .....</b>	<b>3</b>
<b>3</b>	<b>Strojové rozpoznávanie tváří .....</b>	<b>5</b>
3.1	História.....	5
3.1.1	Woody Bledsoe, Helen Chan a Charles Bisson .....	5
3.1.2	Christoph von der Malsburg .....	6
3.2	Techniky rozpoznávania tváří.....	7
3.2.1	Tradičné rozpoznávanie .....	7
3.2.2	Trojrozmerné rozpoznávanie .....	8
3.2.3	Rozpoznávanie na základe analýzy štruktúry kože.....	9
3.2.4	Rozpoznávanie pomocou tepelných kamier .....	10
3.3	Spôsoby využitia techniky rozpoznávania tváří.....	11
<b>4</b>	<b>Umelá inteligencia .....</b>	<b>13</b>
4.1	Strojové učenie.....	13
4.2	Hlboké učenie .....	15
4.3	Konvolučné neurónové siete.....	15
<b>5</b>	<b>Dlib knižnica.....</b>	<b>17</b>
<b>6</b>	<b>OpenCV knižnica .....</b>	<b>19</b>
<b>7</b>	<b>Metódy používané knižnicami OpenCV a Dlib.....</b>	<b>21</b>
7.1	Haar, LBP, HOG klasifikátory .....	21
7.2	Dlib detektory pracujúce na báze MMOD, CNN a HOG klasifikátore .....	24
<b>8</b>	<b>Praktická časť .....</b>	<b>27</b>
8.1	Vyrezávanie tváre .....	28
8.2	Detekcia kľúčových tvárových rysov .....	33
8.3	Porovnanie výsledkov OpenCV a Dlib knižníc .....	39
8.4	Spárovanie kľúčových rysov s relevantným 3D modelom .....	40
<b>9</b>	<b>Záver .....</b>	<b>44</b>

# Zoznam obrázkov a tabuliek

Obr. 2.1: Obštrukčné spánkové apnoe. ....	3
Obr. 3.1: Strojové rozpoznávanie tvárí. ....	5
Obr. 3.2: Tradičné rozpoznávanie. ....	8
Obr. 3.3: Trojrozmerné rozpoznávanie. ....	9
Obr. 3.4: Rozpoznávanie pomocou tepelných kamier. ....	10
Obr. 4.1: Umelá inteligencia. ....	13
Obr. 4.2: Umelý neurón tvoriaci perceptrón. ....	14
Obr. 4.3: Konvolučné neurónové siete. ....	16
Obr. 5.1: Detekcia tváre s použitím Dlib knižnice. ....	17
Obr. 6.1: Detekcia tváre s použitím OpenCV knižnice. ....	19
Obr. 7.1: Haar kaskády. ....	22
Obr. 7.2: LBP kaskády. ....	23
Obr. 8.1: Kľúčové body pre detekciu tváre knižnice Dlib. ....	27
Obr. 8.2: Kľúčové body detekcie tváre zobrazené na použitom modeli hlavy. ....	28
Obr. 8.3: Nastavenie parametrov pre správnu detekciu tváre. ....	29
Obr. 8.4: Korektná detekcia tváre kaskády Haar z ľavej strany. ....	30
Obr. 8.5: Korektná detekcia tváre kaskády Haar z prednej strany. ....	30
Obr. 8.6: Korektná detekcia tváre kaskády Haar z pravej strany. ....	30
Obr. 8.7: Chybná detekcia tváre kaskády Haar z ľavej strany. ....	31
Obr. 8.8: Chybná detekcia tváre kaskády Haar z pravej strany. ....	31
Obr. 8.9: Korektná detekcia tváre kaskády LBP z ľavej strany. ....	32
Obr. 8.10: Korektná detekcia tváre kaskády Dlib z pravej strany. ....	32
Obr. 8.11: Korektná detekcia tváre kaskády Dlib z prednej strany. ....	33
Obr. 8.12: Korektná detekcia tváre kaskády Dlib z ľavej strany. ....	33
Obr. 8.13: Korektná detekcia bodov tváre pomocou OpenCV č.1. ....	34
Obr. 8.14: Korektná detekcia bodov tváre pomocou OpenCV č.2. ....	34
Obr. 8.15: Korektná detekcia bodov tváre pomocou OpenCV č.3. ....	34
Obr. 8.16: Nesprávna detekcia bodov tváre pomocou OpenCV č.1. ....	35
Obr. 8.17: Nesprávna detekcia bodov tváre pomocou OpenCV č.2. ....	35
Obr. 8.18: Nesprávna detekcia bodov tváre pomocou OpenCV č.3. ....	36
Obr. 8.19: Korektná detekcia bodov tváre pomocou Dlib č.1. ....	36
Obr. 8.20: Korektná detekcia bodov tváre pomocou Dlib č.2. ....	37

Obr. 8.21: Korektná detekcia bodov tváre pomocou Dlib č.3. ....	37
Obr. 8.22: Nesprávna detekcia bodov tváre pomocou Dlib č.1. ....	38
Obr. 8.23: Nesprávna detekcia bodov tváre pomocou Dlib č.2. ....	38
Obr. 8.24: Nesprávna detekcia bodov tváre pomocou Dlib č.3. ....	38
Obr. 8.25: Prvý krok spárovania kľúčových rysov s 3D modelom. ....	40
Obr. 8.26: Druhý krok spárovania kľúčových rysov s 3D modelom. ....	40
Obr. 8.27: Tretí krok spárovania kľúčových rysov s 3D modelom. ....	41
Obr. 8.28: Štvrtý krok spárovania kľúčových rysov s 3D modelom. ....	41
Obr. 8.29: Realizácia 3D spracovania na vybranom modeli. ....	42
Obr. 8.30: Piaty krok spárovania kľúčových rysov s 3D modelom. ....	43



## Zoznam skratiek

Skratka	Anglický význam	Slovenský význam
<b>OSA</b>	Obstructive sleep apnea	Obštrukčné spánkové apnoe
<b>LBP</b>	Local binary patterns	Lokálne binárne vzory
<b>HOG</b>	Histogram of oriented gradient	Histogram orientovaného gradientu
<b>FRGC</b>	Face recognition grand challenge	Veľká výzva rozpoznávania tváre
<b>ARL</b>	Army Research Laboratory	Armádne výskumné laboratórium
<b>UI</b>	Artificial Intelligence	Umelá inteligencia
<b>CNN</b>	Convolutional neural networks	Konvolučné neurónové siete
<b>SIANN</b>	Space invariant artificial neural network	Priestorovo invariantné umelé neurónové siete
<b>MMOD</b>	Max-margin object detection	Detekcia maximálneho okraju objektu
<b>GPU</b>	Graphics processing unit	Grafický procesor
<b>ROI</b>	Region of Interest	Oblasť záujmu

# Slovník termínov

Termín	Význam termínu
<b>Obštrukčné spánkové apnoe</b>	Ochorenie charakterizované opakovanými epizódami prerušenia dýchania počas spánku.
<b>Adenoid</b>	Hltanová (hovor. „nosná“) mandľa.
<b>Nosové polypy</b>	Nezhubné novotvary vyrastajúce v nosnej dutine a prínosových dutinách, ktoré vznikajú ako následok alergie alebo chronického zápalu.
<b>Dysmorfológia</b>	Štúdium vrodených chýb.
<b>Kaskáda</b>	Spôsob spracovávania pri programovaní, pri ktorom sú aspoň dva komponenty alebo úkony usporiadané v sekvencii a výstup jedného je vstup druhého.

# Podakovanie

Predovšetkým sa chcem poďakovať vedúcemu práce doc. Ing. Dušanovi Koniarovi, PhD. a konzultantovi práce Ing. Jozefovi Volákovi za množstvo času, ktoré mi venovali pri zodpovedaní odborných a praktických otázok súvisiacich s mojou diplomovou prácou, za ich odborné vedenie, rady a trpezlivosť. V neposlednom rade patrí vďaka aj mojej rodine za ich nepretržitú morálnu podporu.

# 1 ÚVOD

Syndróm spánkového apnoe je porucha spánku, prejavujúca sa zastavením dýchania, čo znižuje nasýtenie krvi kyslíkom a prerušenie spánku. Výsledkom je narušenie prirodzených vzťahov medzi spánkom, respiračným systémom a kardiovaskulárnym systémom.

Najčastejšie sa vyskytujúce je obštrukčné spánkové apnoe (OSA), zapríčinené zlyhaním a uzatvorením horných dýchacích ciest, z dôvodu ochabnutia daných svalových skupín počas trvania spánku. Môže spôsobiť rôzne druhy srdcovocievnych ochorení a zhoršiť priebeh respiračných ochorení u pacientov.

V dnešnej dobe sa obštrukčné spánkové apnoe diagnostikuje v takzvaných spánkových laboratóriách. Táto metóda je finančne aj časovo náročná, a preto vznikla iniciatíva vyhľadávania ďalších možných spôsobov detekcie tohto ochorenia.

Predispozícia na obštrukčné spánkové apnoe sa dá určiť aj na základe tvaru pacientovej tváre a jej jednotlivých črt, a práve touto problematikou sa budeme zaoberať v diplomovej práci.

V teoretickej časti diplomovej práce rozoberieme problematiku a dôvody vzniku obštrukčného spánkového apnoe (OSA), spôsoby rozpoznávania tváří, aktuálny stav v oblasti využitia strojového a hĺbkového učenia a upresníme si kľúčové rysy tváre pre podporu automatizovanej diagnostiky OSA.

Praktická časť je zameraná na výber algoritmu pre vyhľadávanie zvolených rysov tváre v 2D obraze a samotnou verifikáciou a analýzou možností spárovania rysov s relevantným 3D modelom.

Dôležitým aspektom je automatizácia metódy, preto používame metódy strojového učenia a počítačového videnia. Detekcia hláv upresní tvorbu modelov a v rámci tváre vyhľadania kľúčových rysov pre detekciu obštrukčného spánkového apnoe. Mottom práce je následná korelácia s lekármi navrhnutými referenčnými modelmi.

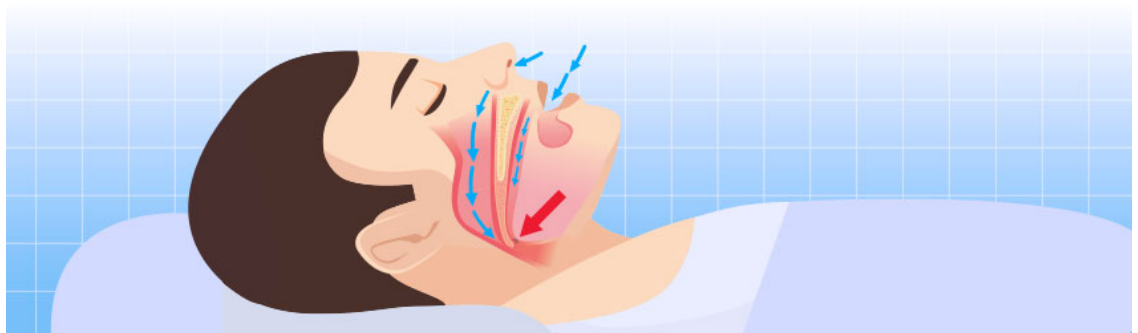
V súčasnej dobe existujú metódy na detekciu rysov v 3D obraze, v 2D obraze s nutnosťou následnej reprojekcie na model alebo analýzy v špecializovaných modalitách, napríklad termo-obraz. My sa budeme zameriavať na 2D obraz (RGB) a nájdené body sa budú reprojektovať na model zhotovený vizuálnym systémom, čo predstavuje paralelne riešenú problematiku.

Na riešenie spomínanej problematiky budeme využívať jazyk C++ a dve dostupné knižnice OpenCV a Dlib, ktorých úspešnosť pri detekcii následne porovnáme. Postup bude spočívať v dvoch krokoch: vyrezávanie tváre z obrazu a detekcia kľúčových tvárových rysov. V rámci knižnice OpenCV budú porovnávané dva kaskádové klasifikátory Haar a LBP.

## 2 OBŠTRUKČNÉ SPÁNKOVÉ APNOE

Obštrukčné spánkové apnoe (OSA) je relatívne bežný stav, keď sa steny hltanu uvoľňujú a zužujú priechodnosť dýchacích ciest. Tým vzniká prerušenie normálneho dýchania. Tento jav môže viesť k pravidelnému prerušovaniu spánku, čo môže mať veľký vplyv na kvalitu života. OSA sa môže prejavovať dvomi spôsobmi. Apnoe je stav, kedy sa svaly a mäkké tkanivá v hrdle uvoľňujú natoľko, že spôsobia úplné zablokovanie dýchacích ciest. Táto blokáda trvá minimálne 10 sekúnd. Druhým prípadom je hypopnoe, ktoré predstavuje len čiastočné zablokovanie dýchacích ciest, ktoré má za následok zníženie prúdenia vzduchu o viac ako 50%.

Ľudia trpiaci obštrukčným spánkovým apnoe môžu pociťovať opakovaný výskyt apnoe a hypopnoe počas noci. Vo vážnych prípadoch sa tento stav môže opakovať každú minútu. Termín obštrukčný rozlišuje OSA od zriedkavejších foriem spánkového apnoe, ako je centrálné spánkové apnoe, ktoré je spôsobené tým, že mozog počas spánku nevysiela signály do dýchacích svalov. [1]



Obr. 2.1 Obštrukčné spánkové apnoe [1].

### Príznaky OSA

Príznaky OSA sú často prvotne zistené partnerom, priateľom alebo členom rodiny, ktorí si problémy všimnú počas spánku. Toto ochorenie sa môže prejavovať hlasným chrápaním, hlučným a namáhavým dýchaním alebo opakovanými krátkymi obdobiami, kedy je dýchanie prerušované lapaním po dychu alebo odfrkovaním. Niektorí ľudia môžu pociťovať aj nočné potenie a môžu sa počas noci často prebúdzat' na močenie. Počas tejto blokády nedostatok kyslíka spúšťa mozog, aby osobu trpiacu syndrómom OSA vytiahol z hlbokého spánku buď do ľahšieho spánku alebo do bdlosti. Vďaka tomu sa dýchacie cesty znovu otvoria a umožnia obnoviť normálne

dýchanie. Zvyčajne však človek majúci obštrukčné spánkové apnoe nemá žiadnu spomienku na prerušené dýchane, takže ani nemusí vedieť, že má nejaký problém.

### *Príčiny OSA*

Je normálne, že svaly a mäkké tkanivá v hrdle sa počas spánku uvoľňujú a upadajú. Avšak, u väčšiny ľudí to nespôsobuje problémy s dýchaním. U ľudí s OSA sa dýchacie cesty zúžili v dôsledku mnohých faktorov. Jedným z najdôležitejších faktorov je nadváha. Nadmerný telesný tuk zvyšuje objem mäkkého tkaniva v krku, čo môže spôsobiť namáhanie svalov hrdla a nadbytok žalúdočného tuku môže tiež viesť k ťažkostiam s dýchaním. Obštrukčné spánkové apnoe sa vyskytuje častejšie u populácie mužského pohlavia. Príčina tohto javu nie je známa, ale môže súvisieť s rozdielnym rozloženým telesného tuku. Taktiež vek môže ovplyvňovať výskyt tohto ochorenia. Hoci sa OSA môže objaviť v akomkoľvek veku, častejšie je zistené u ľudí, ktorí majú viac ako 40 rokov. Ďalším parametrom môže byť veľkosť krku. Muži s veľkosťou goliera väčšou ako približne 43cm (17 palcov) majú zvýšené riziko vzniku OSA. Aj užívanie liekov so sedatívnym účinkom, ako sú napríklad tabletky na spanie alebo upokojenie môže ovplyvniť priebeh apnoe. Vrodenú predispozíciu majú ľudia s nezvyčajnou vnútornou štruktúrou hrdla, ako napríklad úzke dýchacie cesty, veľké mandle, adenoidy a jazyk, alebo naopak malá dolná čelusť. Rizikovými faktormi je aj alkohol a fajčenie. Alkohol, najmä pred spaním môže zhoršiť chrápanie a spánkové apnoe a je pravdepodobnejšie, že sa spánkové apnoe objaví u osoby, ktorá fajčí. U žien môže výskyt ovplyvniť menopauza. Zmeny hladín hormónov počas menopauzy môžu spôsobiť, že svaly hrdla sa uvoľnia viac ako obvykle. Rodinná anamnéza môže tiež ovplyvniť výskyt OSA. Veľmi dôležitým aspektom je upchávanie nosa. OSA sa vyskytuje častejšie u ľudí s upchatým nosom, ako je napríklad odklonená priehradka alebo nosné polypy, ktoré môžu byť dôsledkom zúženia dýchacích ciest. [1]

## 3 STROJOVÉ ROZPOZNÁVANIE TVÁRÍ

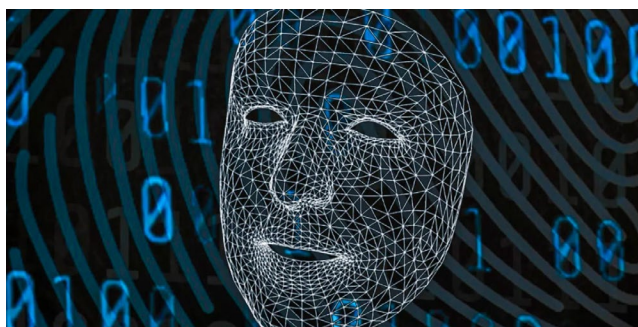
Strojové (automatizované) rozpoznávanie tváří je súbor algoritmov schopný detegovať v digitálnom obraze črty typické pre tvár, a tým pádom ju rozpoznať od iných objektov v snímanej scéne. Výsledky takejto automatizovanej detekcie možno použiť v medicíne, biometrii, identifikácii osôb a pod. [2]

### 3.1 HISTÓRIA

Medzi priekopníkov automatického rozpoznávania tváří patria Woody Bledsoe, Helen Chan Wolf a Charles Bisson. V priebehu rokov 1964 až 1965 spolu pracovali na používaní počítača na rozpoznávanie ľudských tváří. Avšak, finančné prostriedky poskytovala nemenovaná spravodajská služba, ktorá neumožnila tomuto výskumu veľkú publicitu, a preto bolo zverejnených len málo informácií. [3]

#### 3.1.1 Woody Bledsoe, Helen Chan a Charles Bisson

Na základe dostupných zdrojov sa zistilo, že pôvodný spôsob podľa Bledsoe zahŕňal manuálne označovanie rôznych rysov tváre, ako sú oči, ústa, atď. Tieto body boli potom matematicky prepočítavané počítačom. Vzdialenosti medzi danými bodmi boli automaticky vypočítané a porovnané so snímkami z databázy na určenie identity.



*Obr. 3.1 Strojové rozpoznávanie tváří [4]*

Bledsoe však opísal ťažkosti s veľkou variabilitou otáčania hlavy a sklonu, intenzitou osvetlenia a uhla, výrazom tváre, starnutím a mnohými ďalšími aspektmi. Niektoré ďalšie pokusy o rozpoznanie tváre strojom umožnili len malú alebo žiadnu variabilitu v tomto smere. Metóda korelácie (hľadanie vzoru) nespracovaných dát, ktorá je niektorými vedcami často používaná, určite zlyhá v prípade, kedy je



variabilita veľká. Najmä v prípade dvoch fotografií toho istého človeka s dvomi rôznymi natočeniami hlavy. [3]

Tento projekt bol označený ako človek – stroj, pretože človek extrahoval súradnice vybraných bodov z fotografií, ktoré potom počítač používal na rozpoznanie. Pomocou grafickej funkcie (GRAFACON alebo RAND TABLET) operátor vyberal súradnice bodov ako je stred zreníc, vnútorné okraje očí, vonkajšie okraje očí, hranica začiatku rastu vlasov na čele, atď. Z týchto súradníc bol vypočítaný zoznam 20 vzdialeností, ako napríklad šírka úst, šírka očí alebo vzdialenosť medzi zrenicami. Títo operátori boli schopní spracovať približne 40 snímok v priebehu jednej hodiny. Pri vytváraní databázy bolo meno konkrétnej osoby na fotografii asociované so zoznamom vypočítaných vzdialeností a uložený v počítači. Vo fáze rozpoznávania sa súbor vzdialeností porovnával so zodpovedajúcou vzdialenosťou pre každú fotografiu, čím sa získala podobnosť medzi fotografiou a záznamom z databázy. Záznamy s najväčšou podobnosťou boli výsledkom rozpoznávania. Pretože je veľmi nepravdepodobné, že by sa pri rotácii hlavy, pri chôdzi, naklonení a mierke (vzdialenosť od fotoaparátu), vyrovnali akékoľvek dva obrázky, každá sada vzdialeností je normalizovaná tak, aby predstavovala tvár v čelnej orientácii. Na dosiahnutie tejto normalizácie sa program najskôr pokúša určiť naklonenie, otočenie a rotáciu. Potom sa počítač vypočíta účinok týchto transformácií na vzdialenosti objektov a odstráni ho. Na výpočet týchto uhlov musí počítač poznať trojrozmernú geometriu hlavy. [3]

Po tom ako Bledsoe v roku 1996 opustil PRI (Panoramic Research, Inc.), tento projekt pokračoval na Standfordskom výskumnom inštitúte, predovšetkým pod vedením Petra Hart.

### 3.1.2 Christoph von der Malsburg

Do roku 1997 systém vyvinutý Christophom von der Malsburgom, absolventmi Univerzity Bochum v Nemecku a Univerzity v južnej Kalifornii v Spojených štátoch prekonal väčšinu systémov. Ako napríklad systémy Massachusetts Institute of Technology a Marylandskej Univerzity. Softvér bol predaný ako ZN-Face a využívaný zákazníkmi ako Deutsche Bank, operátormi letísk

a iných rušných lokalít. Bol natoľko obširny, že dokázal identifikovať tváre aj z nedokonalých záberov a dokázal prekonať prekážky ako zmena účesu, okuliare, brady a fúzy. [5]

V roku 2006 bola vyhodnotená výkonnosť najnovších algoritmov rozpoznávania tváří vo Face Recognition Grand Challenge (FRGC). V testoch sa použili obrázky s vysokým rozlíšením, 3D tvárové skenovanie a snímky s dúhovkou. Výsledky ukázali, že nové algoritmy sú 10-krát presnejšie než algoritmy rozpoznávania tváří v roku 2002 a 100-krát presnejšie než algoritmy z roku 1995. [5]

## **3.2 *TECHNIKY ROZPOZNÁVANIA TVÁRÍ***

Proces rozpoznávania tváří sa v podstate skladá z dvoch krokov. Prvý zahŕňa extrakciu a výber prvkov a druhý klasifikáciu objektov. Neskorší vývoj však zaviedol do procesu rozpoznávania rôzne technológie, ktoré budú opísané v tejto kapitole.

### **3.2.1 Tradičné rozpoznávanie**

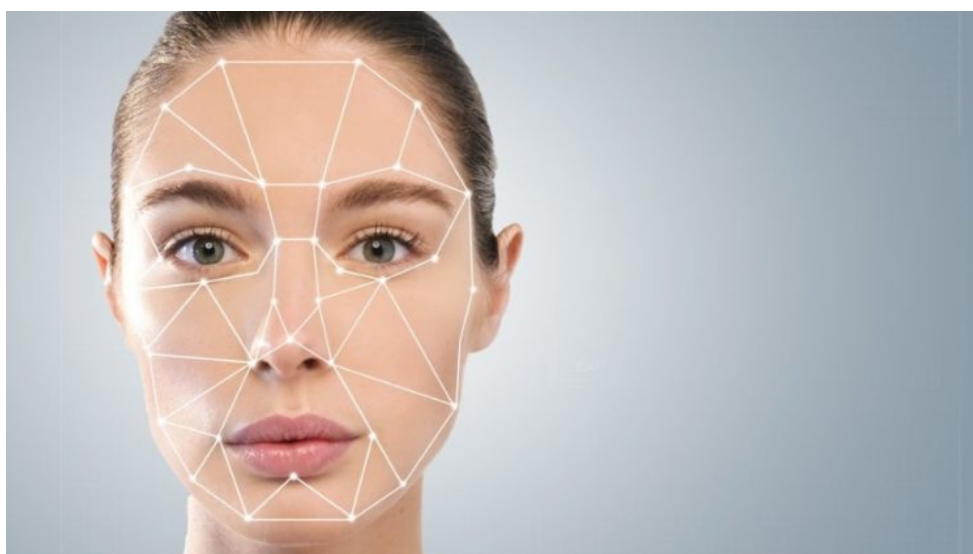
Niektoré algoritmy rozpoznávania tváre identifikujú funkcie tváre extrahovaním orientačných bodov alebo funkcií z fotografie tváre subjektu. Takéto algoritmy napríklad môžu analyzovať relatívnu polohu, veľkosť a/alebo tvar očí, nosa, lícných kostí a čeluste. Tieto parametre sa potom používajú na porovnávanie ďalších obrázkov so zodpovedajúcimi údajmi. [6]

Ostatné algoritmy normalizujú galériu tvárových obrázkov, komprimujú údaje a ukládajú len informácie, ktoré sú užitočné pre rozpoznávanie tváří. Snímok sa potom porovnáva. Jeden z najskorších úspešných systémov je založený na technikách porovnávania šablón aplikovaných na súbor výrazných tvárových funkcií, ktoré poskytujú určitý druh komprimovanej tváre. [6]

Rozpoznávacie algoritmy sa dajú rozdeliť na dve hlavné metódy. Geometrické, ktoré sa zameriavajú na rozlišovacie znaky alebo fotometrické, ktoré sa zakladajú na štatistike, ktorá transformuje obraz do hodnôt a porovnáva hodnoty so šablónami s cieľom odstrániť rozdiely. Tieto algoritmy sa dajú klasifikovať aj do dvoch širokých kategórií: holistické a funkčné modely. Prvý z nich sa pokúša rozpoznať tvár ako

celok, zatiaľ čo funkčný rozdeľuje tvár na komponenty a analyzuje ich ako aj ich priestorovú polohu vzhľadom na iné črty.

Populárne algoritmy rozpoznávania zahŕňajú analýzu základných komponentov pomocou vlastných rozhraní, lineárnu diskriminačnú analýzu, porovnávanie elastických grafov pomocou algoritmu Fisherface, skrytý model Markov, multilineárne podpriestorové učenie pomocou reprezentácie tenzorov a neuronálne motivované dynamické prepojenie. [6]



*Obr. 3.2 Tradičné rozpoznávanie [7]*

### 3.2.2 Trojrozmerné rozpoznávanie

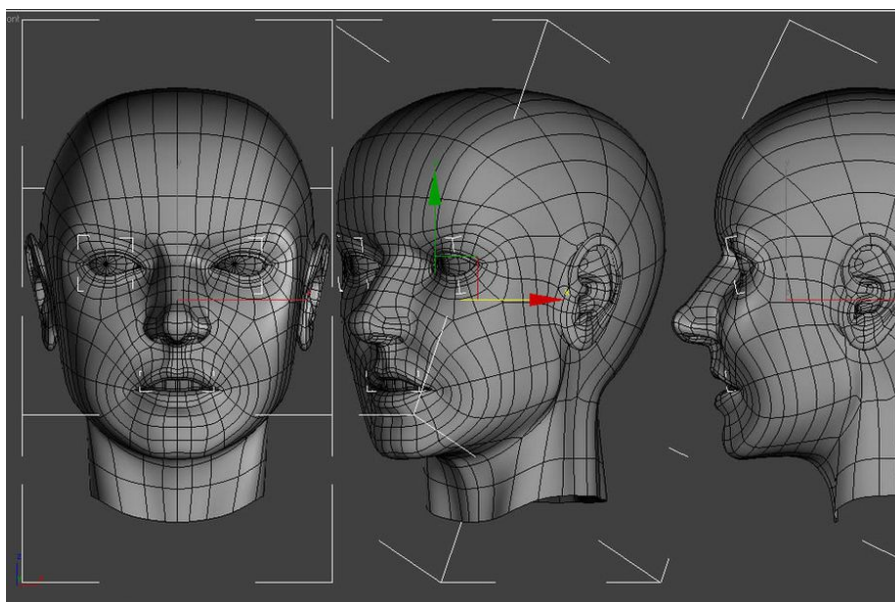
Trojrozmerná technológia rozpoznávania tváre využíva 3D snímače na zachytenie informácií o rozmeroch tváří. Tieto informácie sa potom používajú na identifikáciu charakteristických vlastností tváří, ako sú obrysy očí, nosa, brady, atď. [8]

Jednou z výhod 3D rozpoznávania tváre je to, že nie je ovplyvnená zmenami v osvetlení, ako napríklad iné techniky. Dokáže tiež identifikovať tvár z rôznych uhlov pohľadu, vrátane profilu. Trojrozmerné dátové body z tváre výrazne zlepšujú presnosť rozpoznávania tváří. 3D výskum je posilnený vývojom sofistikovaných senzorov, ktoré dokážu lepšie zachytiť 3D tvárové snímky. Snímače sú založené na premietaní štruktúrovaného svetla na tvár. Až tucet alebo viac z týchto snímačov

obrazu sa môže umiestniť na rovnaký CMOS čip – každý snímač zachytí inú časť spektra. [8]

Každá metóda má však aj svoje slabé miesta. Dokonca aj dokonalá technológia 3D prispôbovania môže byť citlivá na výrazy. Na tento cieľ aplikovala skupina vedcov z Technion nástroje metrickej geometrie pre vyriešenie problémov ako je izometria.

Nová metóda predstavuje spôsob, ako zachytiť obrázok pomocou troch sledovacích kamier, ktoré snímajú v rôznych uhloch. Jedna kamera smeruje na prednú časť objektu, druhá kolmo na jednu stranu a tretia z iného uhla. Všetky tieto kamery pracujú naraz, aby mohli sledovať tvár subjektu v reálnom čase a boli schopné detegovať črty a rozpoznávať tváre.



Obr. 3.3 Trojrozmerné rozpoznávanie [9]

### 3.2.3 Rozpoznávanie na základe analýzy štruktúry kože

Ďalším novým trendom je metóda využívajúca detaily kože, ktoré sú zachytené na štandardných digitálnych alebo skenovaných fotografiách. Táto technika, nazývaná analýza textúry kože, transformuje jedinečné línie, vzory a škvrny viditeľné na koži človeka do matematického priestoru. [10]

Analýza povrchovej textúry funguje podobne ako rozpoznávanie tvárí. Analyzuje sa snímok časti kože, nazývaný skinprint. Táto plocha sa rozdelí na ešte menšie bloky.

Pomocou algoritmov pre transformáciu textúry na matematický, merateľný priestor bude systém rozlišovať všetky čiary, póry a skutočnú textúru pokožky. Môže identifikovať aj kontrast medzi identickými pármami, čo doteraz nie je možné pomocou samotného softvéru rozpoznávania tváří. Testy ukázali, že pridanie analýzy štruktúry pokožky môže výkonnosť pri rozpoznávaní tváří zvýšiť o 20 až 25 percent. [10]

### 3.2.4 Rozpoznávanie pomocou tepelných kamier

Inou formou prijímania vstupných údajov na rozpoznávanie tváre môže byť aj metóda založená na termovíznych kamerách. Pri tomto postupe kamery zistia len tvar hlavy a ignorujú doplnky ako sú okuliare, klobúky alebo make-up. Na rozdiel od konvenčných kamier dokážu termálne kamery zachytiť snímky tváří aj v zlých a nočných podmienkach bez použitia blesku a odhalenia polohy kamery. [11]

Problém s používaním termálnych kamier na rozpoznávanie tváre však spočíva v tom, že databázy sú obmedzené. Diego Socolinsky a Andrea Selinger v roku 2004 skúmali využitie tejto metódy v reálnom živote a súčasne vytvárali novú databázu termálnych tvárových fotografií.



Obr. 3.4 Rozpoznávanie pomocou tepelných kamier [12]

V roku 2018 vyvinuli vedci z amerického armádneho výskumného laboratória (ARL) techniku, ktorá im umožňuje porovnávať fotografie tváří získané pomocou termovíznej kamery s fotografiami v databázach, ktoré boli zachytené bežnými

kamerami. Táto metóda využíva umelú inteligenciu a strojové učenie, ktoré umožňujú viditeľne porovnávať konvenčné a termálne tvárové snímky. Metóda je známa ako syntéza krížového spektra, pretože dokáže rozpoznávať tváre z dvoch rôznych spôsobov zobrazovania. Pozostáva z nelineárneho regresného modelu, ktorý mapuje konkrétny termálny obraz na zodpovedajúcu tvárovú snímku. [11]

Podľa výkonnostných testov vykonaných v ARL vedci zistili, že viacrozsahový model krížovej syntézy preukázal zlepšenie výkonnosti detekcie tvárí o približne 30% oproti predošlým metódam a približne 5% oproti najmodernejším metódam.

### **3.3 SPÔSOBY VYUŽITIA TECHNIKY ROZPOZNÁVANIA TVÁRÍ**

Aplikácie rozpoznávania tváre sa rozširujú do rôznych aspektov nášho života. Napríklad technológia rozpoznávania tváre sa teraz môže použiť namiesto hesla na odomknutie niektorých mobilných zariadení. Biometria, vrátane rozpoznávania tváre, sa môže požiť pri overovaní zákazníka pri nakupovaní online. Táto metóda je pre používateľov oveľa bezpečnejšia a vhodnejšia ako pamätať si ID používateľa a heslá. Spoločnosť facebook vyvinula rozpoznávanie tváre, aby identifikovala a označila ľudí na fotografiách uverejnených na webových stránkach. Táto aplikácia dokonca dokáže identifikovať tvár a spýtať sa určitej osoby, či sa práve ona nachádza na danej fotografii. [13]

Existuje množstvo aplikácií rozpoznávania tvárí, ktoré sa vyvíjajú na použitie v oblasti zdravotníckej starostlivosti. Pacienti sa v niektorých zdravotníckych systémoch dokážu identifikovať pomocou rozpoznania tváre. Mechanizmus detekcie tváre sa dá použiť aj na určenie úrovne bolesti pacienta, pre rozlíšenie napríklad chronickej bolesti. Tieto technológie sa môžu použiť pre identifikáciu pracovníkov aj pacientov v zdravotníckych zariadeniach, čím sa zdravotná starostlivosť stáva bezpečnejšou a efektívnejšou.

Ďalším využitím rozpoznávania tvárí je aj detekcia genetických chorôb. Tradične lekári diagnostikujú tieto typy ochorení tým, že identifikujú jedinečné vzory tvárových prvkov – dysmorfológia. Vyškolení lekári dokážu čítať tvár dieťaťa ako

mapu a zistiť akékoľvek abnormality podľa tvaru očí, veľkosti čela, plochosti nosa, natočenia uší, atď. – ktoré v kombinácii s inými príznakmi, ako je intelektuálne postihnutie alebo nezvyklá výška či hmotnosť, môžu byť špecifické pre danú genetickú poruchu. Tieto morfológické zmeny však môžu byť veľmi jemné a často ich dokážu rozpoznať len špecializovaní lekárske genetici. S tisíckami variant v genóme každej osoby je identifikácia mutácie alebo vylúčenie určitého ochorenia veľmi náročné. [13]

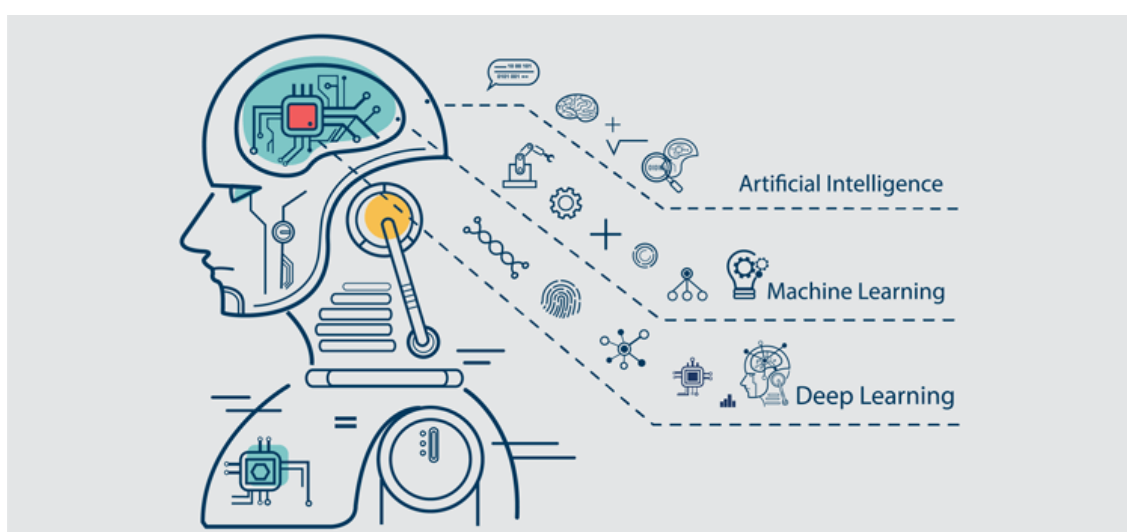
Technológia rozpoznávania tváre sa pokúša napomôcť tomuto procesu detekcie. Napríklad softvér Face2Gene, ktorý používa umelú inteligenciu dokáže porovnávať fotografie pacientov a podáva návrhy možných diagnóz.

V prípade tejto práce budeme na rozpoznávanie tváří používať 2D snímky z dôvodu dostupnosti tejto metódy. Kamery použité pre vizuálny systém na 3D skenovanie poskytujú aj RGB výstup (2D statická snímka). Porovnávať budeme knižnice Dlib a OpenCV založené na prvkoch umelej inteligencie ako je strojové učenie a kaskádové klasifikátory.

## 4 UMELÁ INTELIGENCIA

Umelá inteligencia (UI) je vedný obor alebo disciplína, ktorá sa venuje vývoju algoritmov a strojov vyznačujúcich sa prítomnosťou inteligentného správania. Strojové učenie a jeho podskupina hĺbkové učenie predstavujú kľúčové metódy pre rozširovanie pôsobnosti umelej inteligencie. [14]

História umelej inteligencie sa odvíja od 50. rokov 20. storočia. Za otca umelej inteligencie môžeme považovať amerického informatika a kognitívneho vedca Johna McCarthyho, ktorý ako prvý definoval pojem umelej inteligencie v roku 1955.



Obr. 4.1 Umelá inteligencia [15]

### 4.1 STROJOVÉ UČENIE

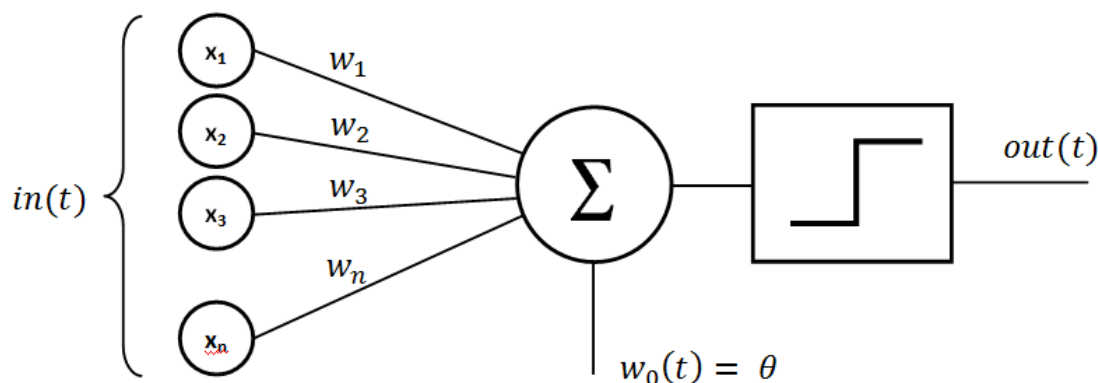
Strojové učenie je podmnožinou umelej inteligencie. To znamená, že všetky strojové učenia patria do umelej inteligencie, ale nie všetky prípady umelej inteligencie sa radia do strojového učenia.

Vývoj strojového učenia je úzko spätý s vývojom samotnej umelej inteligencie, vývojom počítačov a vývojom rôznych štatistických metód. Kanadský psychológ Donald Hebb vytvoril v roku 1949 Hebbovu metódu, ktorá charakterizuje učenie ako proces podmieňujúci vytváranie nových synapsií medzi neurónmi v mozgu. Tie sa môžu na základe opakovaní nadobudnutých poznatkov zosilňovať alebo zoslabovať.



Tento princíp inšpiroval vývoj metódy neurónových sietí, ktorá funguje rovnako ako ľudský mozog. Následne v roku 1952 navrhol Arthur Samuel prvý učiaci sa program.

Prvý perceptrón (umelá neurónová sieť) bol vytvorený Frankom Rosenblattom v roku 1957. Obsahoval však len jeden umelý neurón, a preto mal len malé využitie. To neskôr rozšíril viacvrstvový perceptrón. [16]



Obr. 4.2 Umelý neurón tvoriaci perceptrón [17]

Prvým algoritmom umožňujúcim rozoznávanie vzorov bol nearest neighbour a vznikol v roku 1967.

Jeden z aspektov, ktoré charakterizujú strojové učenie je schopnosť modifikovať sa. Toto učenie je dynamické a na uskutočnenie určitých zmien nevyžaduje zásah človeka. To ho robí menej krehkým a menej závislým od ľudských odborníkov. V zjednodušenom poňatí, učí počítače učiť sa.

Hovorí sa, že počítačový program sa učí zo skúseností  $E$  s ohľadom na určitú triedu úloh  $T$  a meranie výkonnosti  $P$ , ak sa jeho výkon pri úlohách v  $T$ , meraný pomocou  $P$ , zlepšuje so skúsenosťami  $E$ . – Tom Mitchell

V roku 1959, Arthur Samuel, jeden z priekopníkov strojového učenia, definoval strojové učenie ako „študijné pole, ktoré dáva počítačom schopnosť učiť sa bez výslovného programovania.“ Programy strojového učenia sa v istom zmysle prispôbujú v reakcii na údaje, ktorým sú vystavené. [16]

Algoritmy strojového učenia využívajú prvky matematickej štatistiky, metódy štatistickej analýzy a hĺbkovú analýzu dát. Tieto algoritmy môžu vykonávať tri

základné funkcie: klasifikáciu, regresiu a zhlukovanie. Následne sa delia do troch kategórií podľa spôsobu učenia: učenie s učiteľom, učenie bez učiteľa a učenie s posilňovaním.

Strojové učenie sa dá použiť v mnohých odboroch, napríklad v počítačovom videní. To sa špecializuje na získavanie informácií z obrázkov alebo videí, vďaka čomu dokáže rozpoznávať objekty alebo gestá, či sledovať objekty vo videu. Na základe týchto schopností sa využíva v počítačovej grafike, v robotike a dokonca aj v medicíne pri analýze medicínskych snímok.

## **4.2 HĽBOKÉ UČENIE**

V roku 2006 vznikla podskupina strojového učenia - hĺbkové učenie, zahrňujúce pokročilé algoritmy, ktoré sa snažia priblížiť úplne autonómnej umelej inteligencii. Hĺbkové učenie je tiež známe pod názvom hĺbkové štruktúrované učenie alebo hierarchické učenie. Učenie môže prebiehať pod dohľadom, polovičným dohľadom alebo bez dozoru. Hĺbkové učenie sa odlišuje od bežných strojových učebných postupov automatickým učením. [16]

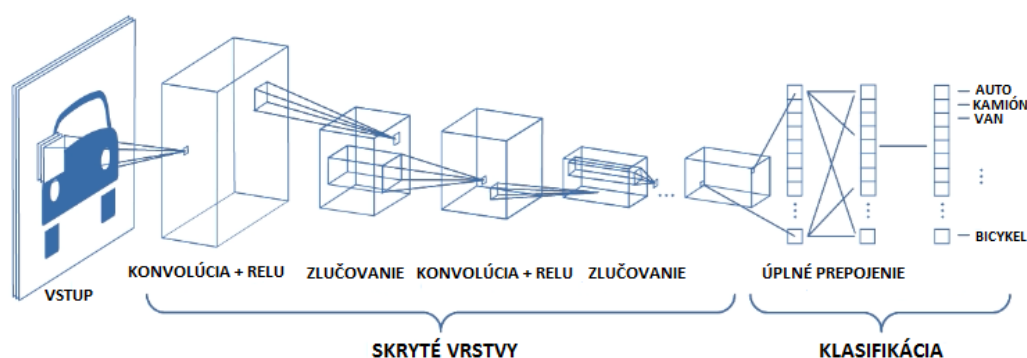
## **4.3 KONVOLUČNÉ NEURÓNOVÉ SIETE**

V terminológii hĺbkového učenia je konvolučná neurónová sieť (CNN/ConvNet) typom hlbokých neurónových sietí, ktoré sa najčastejšie používajú pri analýze vizuálnych obrazov. CNN sú regulované verzie viacvrstvových perceptrónov. Viacvrstvové perceptróny sa zvyčajne patria ku plne prepojeným sieťam, to znamená, že každý neurón v jednej vrstve je pripojený na všetky neuróny v nasledujúcej vrstve. Sú tiež známe ako posunovo invariantné alebo priestorovo invariantné umelé neurónové siete (SIANN – Space Invariant Artificial Neural Network), založené na architektúre zdieľaných váh a translačných invariantných charakteristikách. Konvolučné siete boli inšpirované biologickými procesmi. [18]

CNN využívajú relatívne málo predspracovanie v porovnaní s inými algoritmami klasifikácie obrazu. To znamená, že sieť trénuje s filtermi, ktoré boli v tradičných

algoritmoch vytvorené ručne. Táto nezávislosť od predchádzajúcich postupov pri navrhovaní prvkov je hlavnou výhodou konvolučných neurónových sietí.

CNN pozostáva zo vstupnej a výstupnej vrstvy, ako aj z viacerých skrytých vrstiev. Popis procesu ako konvolúcie v neurónových sieťach je konvenčný (matematicky sa jedná skôr o krížovú koreláciu než konvolúciu). Pri programovaní konvolučnej vrstvy v neurónovej sieti by sme mali uvažovať nasledujúce parametre. Vstup je tenzor udávaný v tvare (počet obrázkov)  $\times$  (šírka obrazu)  $\times$  (výška obrazu)  $\times$  (hĺbka obrazu). Ďalej je to počet konvolučných jadier (kernelov) a ich rozmery (šírka a výška). Hĺbka jadier musí byť rovnaká ako hĺbka obrazu. Konvolučné vrstvy aplikujú operáciu konvolúcie na vstup, pričom výsledok odovzdajú ďalšej vrstve. [18]



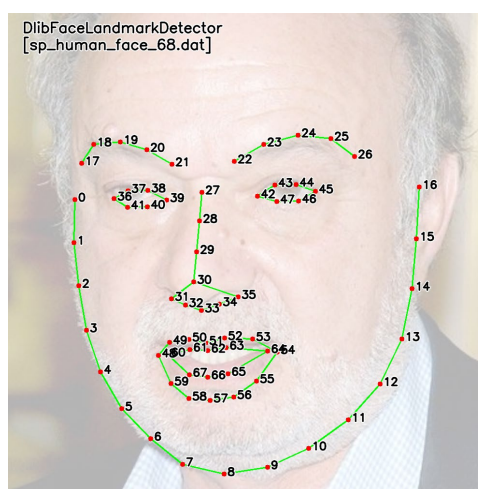
Obr. 4.3 Konvolučné neurónové siete [19]

Tieto siete sa uplatňujú najmä v oblasti rozpoznávania obrazu a videa, klasifikácie obrazu a lekárskej analýzy obrazu.

## 5 DLIB KNIŽNICA

Dlib je moderná všeobecne použiteľná softvérová knižnica v programovacom jazyku C++. Používa sa v priemysle, ako aj v akademickej oblasti v širokej škále vrátane robotiky, vstavaných zariadení, mobilných telefónov a veľkých výkonných výpočtových prostredí. [20]

Od vývoja začatého v roku 2002 sa spoločnosť Dlib rozrástla so širokou škálou nástrojov. Do roku 2016 obsahuje softvérové komponenty pre riešenie problémov so sieťami, vláknami, grafickými používateľskými rozhraniami, dátovými štruktúrami, lineárnou algebrou, strojovým vzdelávaním, spracovaním obrazu, získavaním dát, analýzou XML a textu, numerickou optimalizáciou, Bayesovými sieťami a mnohými ďalšími úlohami. V posledných rokoch bola veľká časť vývoja zameraná na vytvorenie širokého súboru nástrojov štatistického strojového učenia a v roku 2009 bola publikovaná spoločnosť Dlib v časopise Journal of Machine Learning Research [2]. Odvtedy sa používa v širokej škále oblastí.



Obr. 5.1 Detekcia tváre s použitím Dlib knižnice [21]

Na rozdiel od mnohých projektov s otvoreným zdrojovým kódom poskytuje úplnú a presnú dokumentáciu pre každú triedu a funkciu. Obsahuje aj režimy ladenia, ktoré kontrolujú zdokumentované predpoklady pre funkcie. Ak je to povolené, zachytí sa veľká väčšina chýb spôsobených nesprávnym volaním funkcií alebo nesprávnym použitím objektov. Poskytuje taktiež veľa príkladových programov. [20]

Dlib knižnica poskytuje dobré pokrytie skúšky zariadení. Pomer jednotkových testovacích riadkov kódu na riadky knižnice kódu je približne 1 na 4. Knižnica je pravidelne testovaná na systémoch MS Windows, Linux and Mac OS X. Mal by však fungovať na akomkoľvek POSIX systéme a používa sa aj na systémoch Solaris, HP-UX a BSD. Na používanie knižnice nie sú potrebné žiadne ďalšie balíky. Potrebné sú aplikácie, ktoré sú poskytované mimo balíka OS. Taktiež nie je potrebné vykonať ďalšie kroky inštalácie alebo konfigurácie pred používaním knižnice. Celý kód špecifický pre operačný systém je izolovaný vo vnútri OS odberových vrstiev, ktoré sú udržiavané aby boli čo najmenšie. Zvyšok knižnice je buď vrstvený na vrchole vrstiev abstrakcie OS alebo priamo v ISO štandarde C++.

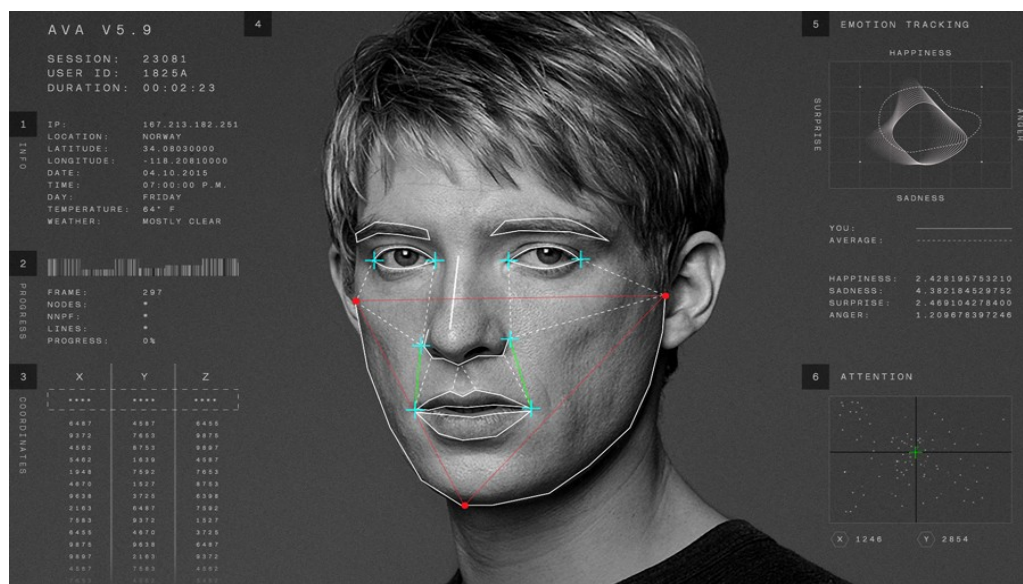
Dlib je všeobecne použiteľná softvérová knižnica napísaná v programovacom jazyku C++. Jeho dizajn je silne ovplyvnený myšlienkami od konštrukčného návrhu a softvérového inžinierstva na báze komponentov. Ide teda predovšetkým o súbor nezávislých softvérových komponentov. Je to softvér s otvoreným zdrojom, ktorý je uvoľnený pod licenciou Boost Software License. [20]

## 6 OPENCV KNIŽNICA

OpenCV (open source) je knižnica programovacích funkcií zameraná hlavne na počítačové videnie v reálnom čase. Pôvodne bola vyvinutá spoločnosťou Intel, neskôr ju podporoval Willow Garage a potom Itseez (ktorý neskôr získal Intel). Knižnica je multiplatformová a bezplatná na použitie pod licenciou BSD s otvoreným zdrojovým kódom. [22]

OpenCV je napísané v C++ a aj jeho primárne rozhranie je v C++, ale stále si zachováva menej komplexné, hoci rozsiahle staršie rozhranie C. Existujú však aj väzby v jazyku Python, Java a MATLAB/OCTAVE. Rozhranie API pre tieto tieto rozhrania je možné nájsť v online dokumentácii. Všetky nové trendy a algoritmy v OpenCV sú teraz vyvinuté v rozhraní C++.

Knižnica OpenCV bola vybudovaná tak, aby poskytovala spoločnú infraštruktúru pre aplikácie počítačového videnia a urýchlila používanie strojového vnímania v komerčných produktoch. Ako produkt s licenciou BSD, OpenCV uľahčuje podnikom využívať a upravovať kód. [22]



Obr. 6.1 Detekcia tváre s použitím OpenCV knižnice [23]

Obsahuje viac ako 2500 optimalizovaných algoritmov, ktoré zahŕňajú komplexnú sadu klasických aj modernejších algoritmov strojového učenia. Tieto algoritmy môžu byť použité na detekciu a rozpoznávanie tváre, identifikáciu objektov, klasifikáciu

ľudských činností vo videách, sledovanie pohybov kamery, sledovanie pohybujúcich sa objektov, extrahovanie 3D modelov objektov, vytváranie 3D bodových zoskupení zo stereofónnych kamier, spájanie záberov a vytvorenie vysokého rozlíšenia. [

## 7 METÓDY POUŽÍVANÉ KNIŽNICAMI OPENCV A DLIB

Pri procese vyhľadávania tváre a rovnako aj pri detekcii jednotlivých kľúčových črt pre diagnostiku obštrukčného spánkového apnoe boli použité rôzne metódy využívajúce knižnice OpenCV a Dlib. Medzi metódy pracujúce s pomocou knižnice OpenCV patria Haar, LBP a Hog kaskádové klasifikátory, zatiaľ čo Dlib pracuje s detektormi využívajúcimi dmm mód, konvolučné neurónové siete a Hog kaskádový klasifikátor.

### 7.1 HAAR, LBP A HOG KLASIFIKÁTORY

Základnou myšlienkou detekcie tváří založenej na kaskádových klasifikátoroch je čo najrýchlejšie vylúčenie nepotrebného obsahu obrazu v skorých štádiách pre minimalizovanie počtu výpočtových operácií v neskorších štádiách.

V ranom štádiu detekcie tváří by mal algoritmus spĺňať tri požiadavky. Prvou požiadavkou je rýchla detekcia. Tváre sú v obraze umiestnené v rôznych polohách a veľkostiach a detegujú sa mierkou a posuvným oknom. V každej polohe sú použité rôzne mierky, takže rýchlosť skorej detekcie určuje rýchlosť celého detekčného systému. Druhým aspektom je vysoká miera detekcie (DR). Včasná detekcia je predbežný skrining. V tomto bode je región obrazu, ktorý je vyhodnotený ako oblasť bez prítomnosti tváre, odstránený. Treťou požiadavkou na algoritmy kaskádových klasifikátorov je vysoká schopnosť rozlíšenia oblasti obrazu bez tváre. Druhá a tretia požiadavka môžu byť protichodné, a preto je dôležité, aby medzi nimi nastala rovnováha. Preto sa zaviedli dva koncepty: miera detekcie tváre (FDR) a miera odmietnutia pozadia (BRR). FDR predstavuje aké percento zistených vzoriek tváre tvorí celkovú rozlohu tváre a BRR reprezentuje percentné zastúpenie vzoriek nedetekujúcich tvár z celkového počtu týchto negatívnych vzoriek. [30]

#### *HAAR KASKÁDY*

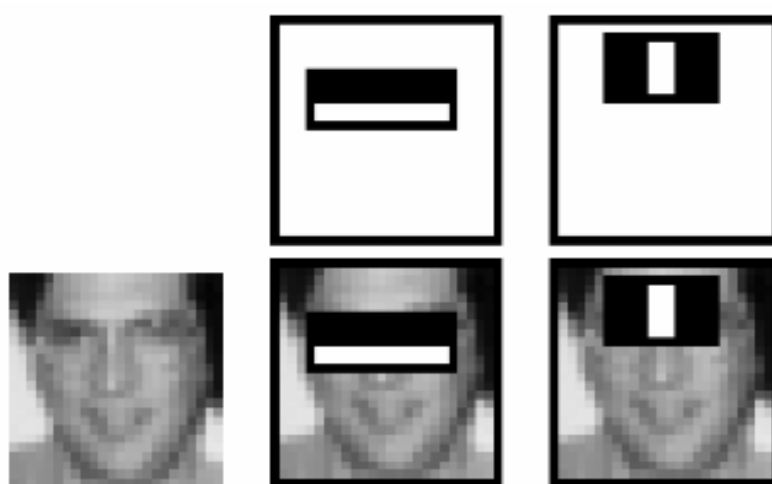
Detekcia objektov pomocou Haarových kaskádových klasifikátorov je efektívna metóda, ktorú v roku 2001 navrhli Paul Viola a Michael Jones vo svojej práci nazvanej Rýchla detekcia objektov pomocou zosilnenej kaskády jednoduchých



funkcií. Ide o metódu založenú na strojovom učení, pri ktorom sa kaskádová funkcia učí z databáz pozitívnych a negatívnych obrázkov a potom sa následne využíva na detekciu objektov v iných obrázkoch. [24]

Algoritmus má štyri stupne: výber funkcie Haar, vytváranie integrálnych obrázkov, adaboost tréning a kaskádové klasifikátory. Haar kaskády sú známe tým, že sú schopné detekovať tváre a časti tela v obraze, ale môžu byť vytrénované na identifikáciu takmer akéhokoľvek objektu. [25]

Ako príklad môže byť uvedené rozpoznávanie tváre. Spočiatku algoritmus potrebuje veľa pozitívnych obrázkov tvárí a negatívnych obrázkov bez tvárí, aby mohol trénovať klasifikátor. Potom je potrebné extrahovať funkcie. Haarova funkcia uvažuje susedné obdĺžnikové oblasti na špecifickom mieste v detekčnom okne, sumarizuje intenzity pixlov v každej oblasti a vypočíta rozdiel medzi týmito sumami. Integrálne obrázky slúžia k urýchleniu celého tohto procesu. [26]

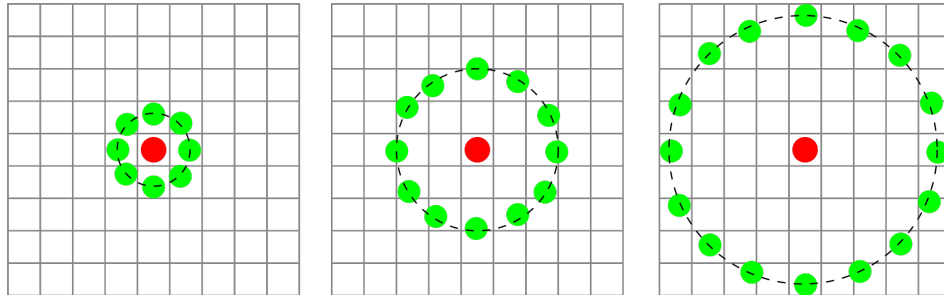


*Obr. 7.1 Haar kaskády [27]*

Ako príklad môže byť uvedené rozpoznávanie tváre. Spočiatku algoritmus potrebuje veľa pozitívnych obrázkov tvárí a negatívnych obrázkov bez tvárí, aby mohol trénovať klasifikátor. Potom je potrebné extrahovať funkcie. Haarova funkcia uvažuje susedné obdĺžnikové oblasti na špecifickom mieste v detekčnom okne, sumarizuje intenzity pixlov v každej oblasti a vypočíta rozdiel medzi týmito sumami. Integrálne obrázky slúžia k urýchleniu celého tohto procesu.

**LBP (LOCAL BINARY PATTERNS) KASKÁDY**

Lokálne binárne vzory (LBP) sú typom vizuálneho deskriptora používaného na klasifikáciu v počítačovom videní. LBP bol prvýkrát opísaný v roku 1994. Od tej doby sa zistilo, že je účinnou metódou pre klasifikáciu textúr a v prípade jeho kombinácie s deskriptorom histogramu orientovaných gradientov (HOG) značne zlepšuje výkon detekcie. [28]



Obr. 7.2 LBP kaskády [29]

Vo svojej najjednoduchšej forme sa vektor funkcie LBP vytvorí nasledujúcim spôsobom. Vyšetrované okno sa rozdelí do buniek (napríklad 16x16 pixelov pre každú bunku). Každý pixel v bunke sa porovná s každým z jeho ôsmich susedov (v ľavom hornom rohu, v ľavom strede, vľavo dole, vpravo hore, atď). Ak je hodnota stredového pixelu väčšia ako hodnota suseda, zapíše sa 0, v opačnom prípade 1. Takto sa vytvorí 8-miestne binárne číslo (ktoré je zvyčajne pre ľahšie spracovanie prevedené na desiatkové číslo). Potom nasleduje zostavenie histogramu frekvencie výskytu každého čísla pre každú bunku. Tento histogram môže byť chápaný ako 256-rozmerný príznakový vektor. Následne je histogram normalizovaný a vzniká zlúčený (normalizovaný) histogram všetkých buniek. To udáva vektor funkcie pre celé okno. Po týchto úpravách môže byť vektor spracovaný pomocou určitého strojového učiaceho algoritmu na klasifikáciu obrázkov. Takéto klasifikátory sa môžu použiť na rozpoznávanie tváre alebo analýzu textúry. [28]

Po získaní LBP obrazu  $f(x,y)$  sa môže LBP histogram definovať ako:

$$H_i = \sum_{x,y} I\{f(x,y) = i\}, i = 0, \dots, n-1 \quad (6.1)$$

Kde  $n$  je počet rozdielov označených LBP operátorom a  $I\{A\}$  je 1 ak je  $A$  pravdivé a 0, ak nie je.

### ***HOG KASKÁDY***

Histogram orientovaného gradientu (HOG) je diskriminačný lokálny deskriptor s invariantnosťou 2D rotácie, ktorý bol aplikovaný na širokú oblasť počítačového videnia, napríklad na detekciu chodcov, rozpoznávanie tváre, detekcia tváre, atď.

Klasifikátory HOG boli upravené na generovanie fondov funkcií s rôznou veľkosťou a zrnitosťou, kvôli očakávaniu, že rôzne funkcie sa môžu prispôbiť zmenám v rôznych štádiách. [31]

Základnou koncepciou HOG klasifikátora je nahradenie pôvodnej hodnoty pixelu gradientom susedných pixelov na zlepšenie odolnosti voči meniacemu sa osvetleniu.

Znížením počtu počítaných histogramov sa znižuje aj schopnosť identifikovať tvár a pozadie, zvýšenie robí výpočet zložitejší, ale výsledok je zároveň citlivejší na rotáciu.

Funkcia HOG reprezentuje počet pixlov v každom smere gradientu, ale nevyhodnocuje rozloženie pozície každého gradientu pixlov. Z toho vyplýva, že tvár je rozdelená na pravidelné, neprekrývajúce sa siete. Predpokladajme, že  $r$  predstavuje počet riadkov a  $s$  počet stĺpcov. Tvár bude tým pádom rozdelená do  $r$  a  $s$  regiónov. V každom z týchto regiónov a potom počíta histogram gradientu. [31]

## ***7.2 DLIB DETEKTORY PRACUJÚCE NA BÁZE MMOD, CNN A HOG KLASIFIKÁTORY***

Pri akomkoľvek druhu spracovania obrazu, počítačového videnia alebo strojového učenia je veľmi pravdepodobné, že užívateľ narazí a použije funkcie dostupné pomocou knižnice Dlib. Knižnica je pôvodne napísaná v jazyku C++, ale má dobrú, ľahko použiteľnú väzbu na programovací jazyk Python. Je veľmi často a úspešne používaná na detekciu tváre a detekciu orientácie tváre. [32]

Pred niekoľkými rokmi bola do knižnice Dlib pridaná implementácia algoritmu detekcie max-margin objektu (MMOD). Tento nástroj sa od tej doby stal veľmi populárny, pretože uľahčuje prácu užívateľa. Vďaka nemu stačí definovať útvary v obraze a detektor sa ich naučí vyhľadávať. Táto metóda nevykonáva žiadne čiastkové vzorkovanie, ale namiesto toho optimalizuje všetky pod-okná. MMOD sa môže použiť na zlepšenie akejkoľvek metódy detekcie objektov. [33]

Dlib detektor využívajúci MMOD je detektor tváre založený na báze konvolučných neurónových sietí. Načíta predtrénovaný model a použije ho na vyhľadanie tváre v obrazoch. Model využívajúci konvolučné siete je omnoho presnejší ako model založený na HOG kaskádovom klasifikátore. Nevýhodou však je, že tento mód vyžaduje oveľa viac výpočtovej sily a má byť vykonaný na GPU (ang. Graphics processing unit – grafický procesor) na dosiahnutie primeranej rýchlosti. [33]

Prvý krok, ktorý je vykonávaný pri tréňovaní tohto módu je definovanie konvolučnej neurónovej siete. Bude sa hodnotiť konvolučne cez celú obrazovú pyramídu. Principiálne funguje ako klasifikátor posuvného okna. To znamená, že je nutné siete definovať rozmery obrazu na ktorom má objekt vyhľadávať a následnej na obraze určiť, či je daný objekt v obraze prítomný.

Pre vytrénovanie danej siete je potrebné vytvorenie databázy pozitívnych a negatívnych snímok, čiže fotografií kde sa nachádzajú tváre v obraze a fotografií, kde sa tváre nenachádzajú vôbec. Vytrénovaná sieť potom používa nadobudnuté poznatky na detekciu ľudskej tváre v nových snímkach.

Detektor fungujúci na báze konvolučnej neurónovej siete je schopný detegovať tváre takmer vo všetkých uhloch. Bohužiaľ však nie je vhodný pre video pracujúce v reálnom čase. [32]

Druhým testovaným spôsobom s použitím knižnice Dlib je funkcia face detection. Tento detektor tváre je vytvorený pomocou klasického HOG (histogram orientovaného gradientu) klasifikátora, ktorý je kombinovaný s lineárnym klasifikátorom, obrazovou pyramídou a systémom detekcie pomocou posuvného okna.

Na začiatku je potrebné aj nastaviť číslo, koľkokrát by sa mal pôvodný obraz zo snímky zväčšiť. V predvolenom nastavení je toto číslo 1 a pracuje vhodne pre väčšinu prípadov. Zvýšenie tohto čísla pomáha detektoru vyhľadať a rozpoznať aj menšie tváre.

Po detekcii vyhľadávanej tváre v obraze je oblasť záujmu zvýraznená obdĺžnikom. Rozmery tohto obdĺžnika je možné upravovať. V prípade tejto diplomovej práce nastala drobná prekážka, keď obdĺžnik detegujúci tvár v niektorých prípadoch nezahŕňal hornú časť čela a spodnú časť brady, a preto bolo potrebné prednastaviť väčšie hranice tohto útvaru. [32]

Pri následnom použití tohto detektora je zrejmé, že nedeteguje tváre vo všetkých uhloch. Jeho cieľom je byť dobrým frontálnym detektorom tváre a tým naozaj je. A prekvapivo dokáže detegovať tvár aj v nie úplne dokonalých podmienkach.

Pre porovnanie času potrebného pre spracovanie a detekciu tváre v obraze, pri 620x420 fotografii trvá detekcia s použitím konvolučných neurónových sietí približne 3,3 sekundy, zatiaľ čo v prípade použitia Hog klasifikátora 0,2 sekundy. Čas spracovania závisí na hardvérovom nastavení a rozmeroch snímok. [32]

## 8 PRAKTICKÁ ČASŤ

Pre správnu a účelnú detekciu tváre pri diagnostike obštrukčného spánkového apnoe je dôležité definovanie kľúčových tvárových rysov pre toto ochorenie. K dispozícii máme body, ktoré je možné detekovať pomocou Dlib knižnice (Obr...). Pre diagnostiku OSA je najdôležitejších 19 tvárových rysov: koreň nosa (Dlib model č.28), špička nosa (Dlib model č.34), špička brady (Dlib model č.9), vnútorný kútik pravého oka (Dlib model č.40), vnútorný kútik ľavého oka (Dlib model č.43), vonkajší kútik pravého oka (Dlib model č.37), vonkajší kútik ľavého oka (Dlib model č.46), horná kontúra pravého oka (Dlib model č.38 a č.39), horná kontúra ľavého oka (Dlib model č.44 a č.45), dolná kontúra pravého oka (Dlib model č.41 a č.42), dolná kontúra ľavého oka (Dlib model č.47 a č.48), ohraničenie nosa z pravej strany (Dlib model č.32), ohraničenie nosa z ľavej strany (Dlib model č.36), pravý kútik úst (Dlib model č.49), ľavý kútik úst (Dlib model č.55), najvyšší bod vrchnej pery (Dlib model č.52), najnižší bod spodnej pery (Dlib model č.58) a líčne kosti po oboch stranách tváre (Dlib model č.1, č.2, č.3, č.15, č.16, č.17).



Obr. 8.1 Kľúčové body pre detekciu tváre knižnice Dlib [34]



*Obr. 8.2 Kľúčové body detekcie tváre zobrazené na použítom modeli hlavy*

## **8.1 VYREZÁVANIE TVÁRE**

V prvom kroku je cieľom správna detekcia štruktúry hlavy a tváre a jej následné vyrezanie. Pre porovnanie sú použité dve metódy: pomocou knižnice Dlib a pomocou OpenCV knižnice. OpenCV knižnica pracuje na základe využitia Haar a LBP kaskádových klasifikátorov, zatiaľ čo Dlib knižnica využíva hlboké neurónové siete, dnn.mmode detektor a tvárový detektor založený Hog klasifikátore.

Po vytrénovaní danej neurónovej siete môžeme upravovať jej nastavenie pomocou piatich parametrov. Prvým parametrom je Scale Factor, ktorý udáva mierku koľkokrát je veľkosť obrázku zmenšená oproti originálu. Druhým parametrom je minNeighbours, určujúci, koľko susedov by si mal každý potenciálny obdĺžnik zachovať. Flags je tretím parametrom s rovnakým významom pre starú kaskádu ako vo funkcii cvHaarDetectObjects. Tá sa používa na vyhľadávanie objektov alebo tvarov nielen pre detekciu tváre, ale jej význam závisí od klasifikátora Haar. Poslednými dvomi parametrami sú minSize a maxSize, ktoré určujú minimálnu

a maximálnu veľkosť objektu. Objekty menšie ako hodnota `minSize` sú ignorované, rovnako ako objekty presahujúce hodnotu parametra `maxSize`.

```
float scale = 1.2;  
int neighbour = 4;  
int rozmerx = 24;  
int rozmery = 24;  
int mrozmerx = 350;  
int mrozmery = 350;
```

*Obr. 8.3 Nastavenie parametrov pre správnu detekciu tváre*

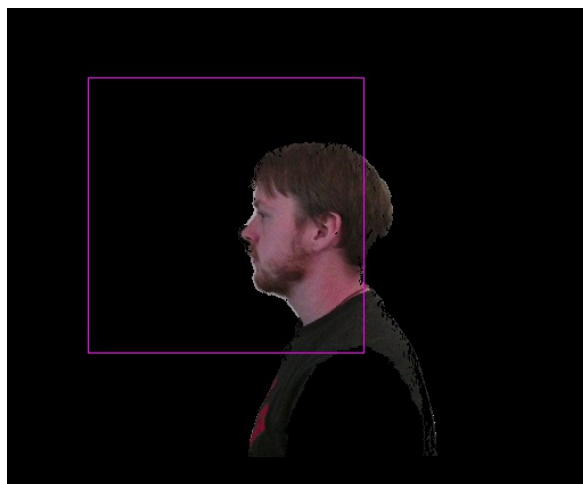
Obrázok (Obr. 8.3) predstavuje časť kódu slúžiaceho na detekciu tváre, v ktorom prebieha nastavenie vyššie spomínaných parametrov. Zobrazené hodnoty sa pri testoch preukázali ako najvhodnejšie pre správne fungovanie daného algoritmu.

Samotné vyrezávanie tváre spočíva v dvoch krokoch. Z prvotnej identifikácie tváre a jej ohraničením pomocou obdĺžnika a z následného vyrezania vygenerovaného útvaru.

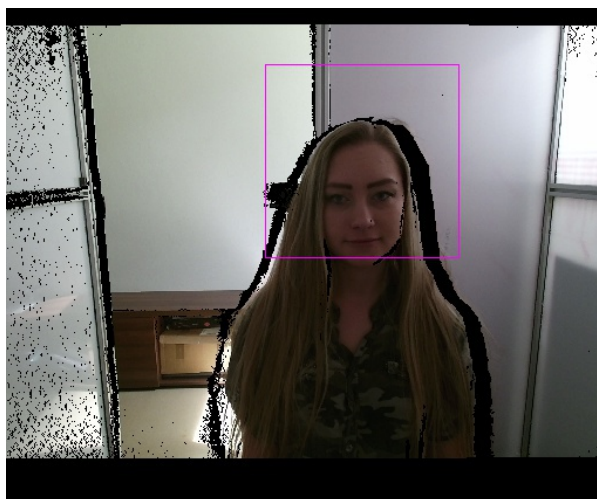
Z hľadiska porovnávania úspešnosti OpenCV a Dlib knižnice v oblasti detekcie a vyrezania tváre môžeme vidieť rozdiely už na úplnom začiatku. Zatiaľ čo knižnica Dlib poskytuje užívateľovi percentuálne väčšie množstvo správne detegovaných tvárí, knižnica OpenCV pracuje s výrazne viditeľnejšou rýchlosťou. Spracovanie snímky pomocou OpenCV trvá maximálne 5 milisekúnd, kým detekcia pomocou obidvoch metód využívajúcich Dlib knižnicu môže trvať až 500 milisekúnd.

Pri detekcii pomocou knižnice OpenCV sme používali dva kaskádové klasifikátory. LBP (Local binary patterns) klasifikátory percentuálne v detekcii tvárí oveľa menej úspešné ako Haar klasifikátory, čo mohlo byť spôsobené prísnejšie stanovenými kritériami pri trénovaní tejto siete. Haar kaskádové klasifikátory síce vyhládali tvár v obraze častejšie, avšak rovnako častejšie aj detegovali chybné časti v obraze. Na nasledujúcich obrázkoch môžeme vidieť správnu detekciu pomocou Haar kaskád.

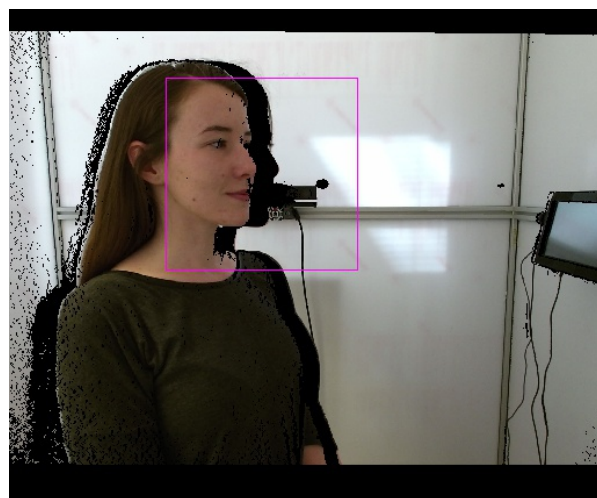




*Obr. 8.4 Korektná detekcia tváre kaskády Haar z ľavej strany*

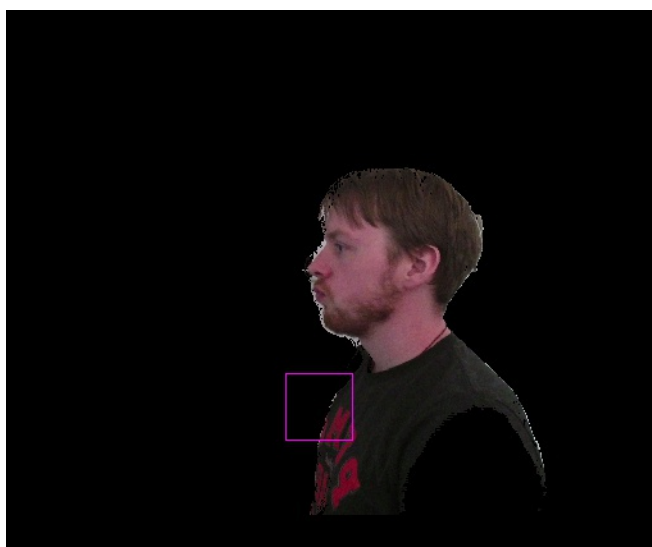


*Obr. 8.5 Korektná detekcia tváre kaskády Haar z prednej strany*



*Obr. 8.6 Korektná detekcia tváre kaskády Haar z pravej strany*

Z obrázkov môžeme vyhodnotiť, že Haar klasifikátor dokázal pomerne úspešne vyhľadať tvár aj v rôznych uhloch a natočeniach. Výborné výsledky nadobudol dokonca aj v prípade, že daná osoba bola zachytená kamerou z profilu. Avšak, ako bolo spomenuté už v texte vyššie, Haar kaskády často identifikovali ako tvár aj nesprávne objekty. V takýchto prípadoch sa väčšinou jednalo o určitú časť oblečenia, zle vyhodnotenú časť pozadia, záber protichodnej kamery alebo klasifikátor označil len určitú časť tváre, nie celú požadovanú oblasť.

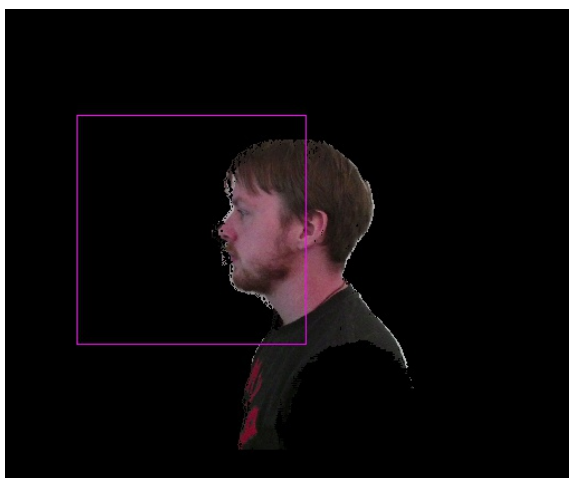


*Obr. 8.7 Chybná detekcia tváre kaskády Haar z ľavej strany*



*Obr. 8.8 Chybná detekcia tváre kaskády Haar z pravej strany*

V prípade LBP klasifikátora boli výsledky menej uspokojivé a početnosť správnej identifikácie tváre bola znateľne nižšia.

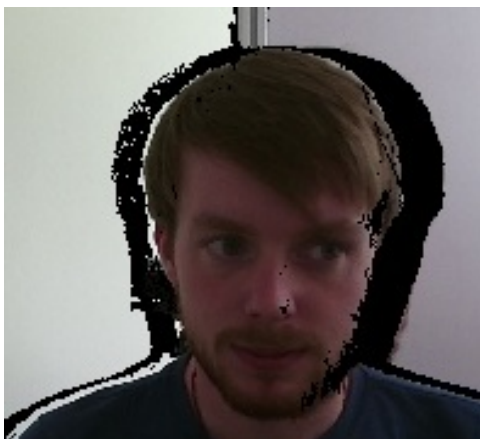


*Obr. 8.9 Korektná detekcia tváre kaskády LBP z ľavej strany*

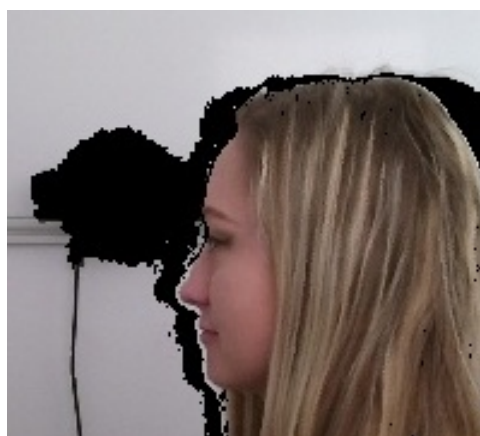
V prípade Dlib knižnice bola funkcia vyhľadávania tváre automaticky spojená s jej vyrezaním. V konečnom dôsledku môže byť metóda využívajúca Dlib knižnicu, aj napriek dlhšej dobe spracovávania vyhodnotená ako najúspešnejšia, pretože má najlepší pomer medzi počtom správne vyhodnotených snímok a počtom fotografií v databáze.



*Obr. 8.10 Korektná detekcia tváre pomocou Dlib z pravej strany*



*Obr. 8.11 Korektná detekcia tváre pomocou Dlib z prednej strany*



*Obr. 8.12 Korektná detekcia tváre pomocou Dlib z ľavej strany*

## **8.2 DETEKCIA KLÚČOVÝCH TVÁROVÝCH RYSOV**

Po vyhľadání tváre nasleduje samotná detekcia klúčových tvárových rysov. Je najvýznamnejším krokom pre už spomínanú možnosť diagnostiky obštrukčného spánkového apnoe. Vďaka vyhľadaným bodom a vzdialenostiam medzi nimi je možné určovať predispozíciu na toto ochorenie. Rovnako ako pri prvom kroku vyhľadávania tváre boli použité dve knižnice pre porovnanie: OpenCV a Dlib.

Nasledujúce snímky zobrazujú správnu detekciu bodov tváre pomocou knižnice OpenCV. Vo vyhľadávaní samotných klúčových rysov bola táto knižnica prekvapivo úspešnejšia ako v hľadaní tváre ako celku. Výsledné percento úspešne detegovaných snímkov je výrazne vyššie.



*Obr. 8.13 Korektná detekcia bodov tváre pomocou OpenCV č.1*

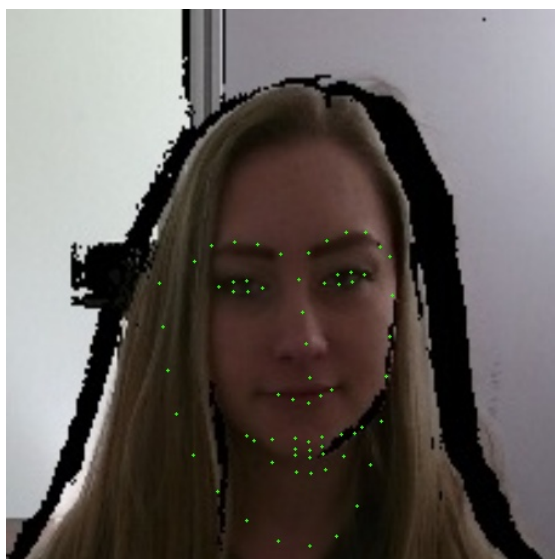


*Obr. 8.14 Korektná detekcia bodov tváre pomocou OpenCV č.2*

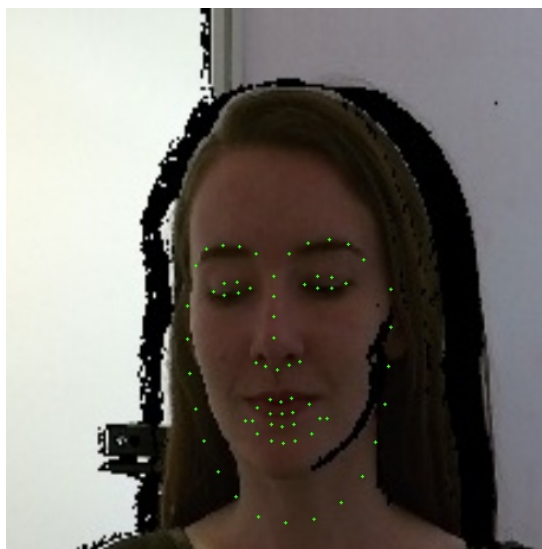


*Obr. 8.15 Korektná detekcia bodov tváre pomocou OpenCV č.3*

Nastali však opäť aj prípady nesprávnej detekcie kľúčových bodov tváre, kedy vyhl'adané črty neboli identické s danou predlohou. Niektoré z týchto chybných výsledkov boli len nepatrne odlišné od očakávaných detekcií, ale nachádzali sa tu aj prípady detegujúce diametrálne odlišné oblasti obrazu, ako bolo potrebné.

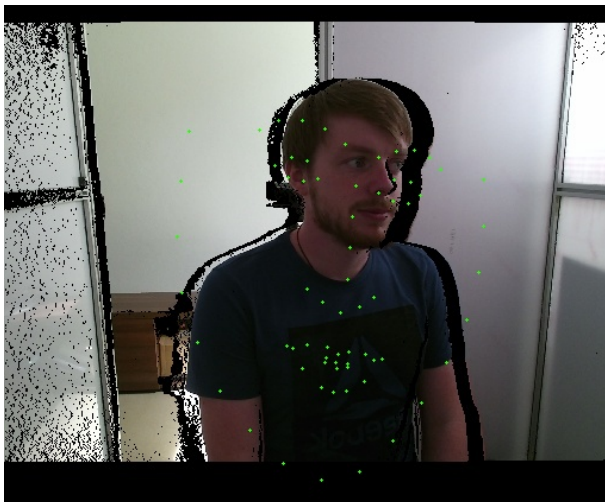


*Obr. 8.16 Nesprávna detekcia bodov tváre pomocou OpenCV č.1*



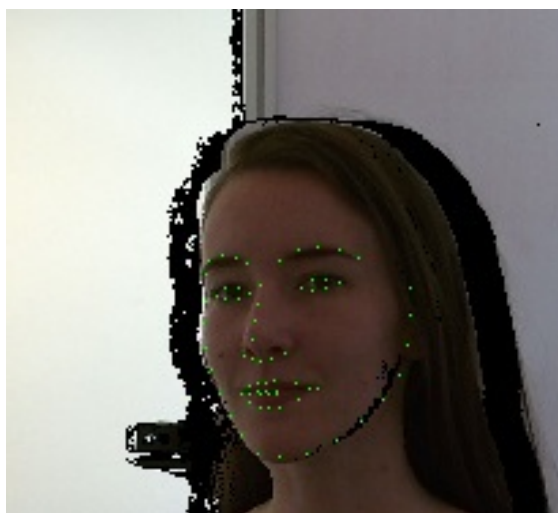
*Obr. 8.17 Nesprávna detekcia bodov tváre pomocou OpenCV č.2*





*Obr. 8.18 Nesprávna detekcia bodov tváre pomocou OpenCV č.3*

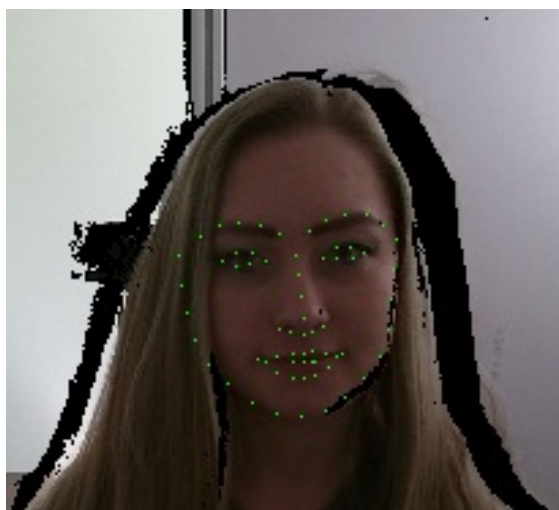
Nasledujúce snímky zobrazujú správnu detekciu bodov tváre pomocou knižnice Dlib. Táto knižnica sa potvrdila ako výrazne úspešnejšia na tieto účely ako predošlá OpenCV knižnica. Väčšina fotografií bola vyhodnotená správne a bez väčších odchýlok od očakávaných výsledkov. Vysoká úspešnosť sa objavila dokonca aj u fotografií zosnímaných z rôznych uhlov.



*Obr. 8.19 Korektná detekcia bodov tváre pomocou Dlib č.1*



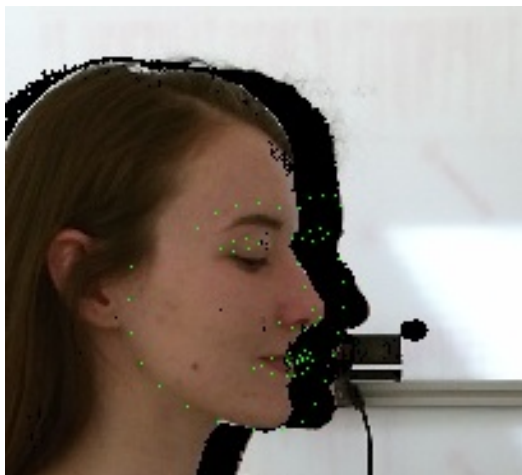
*Obr. 8.20 Korektná detekcia bodov tváre pomocou Dlib č.2*



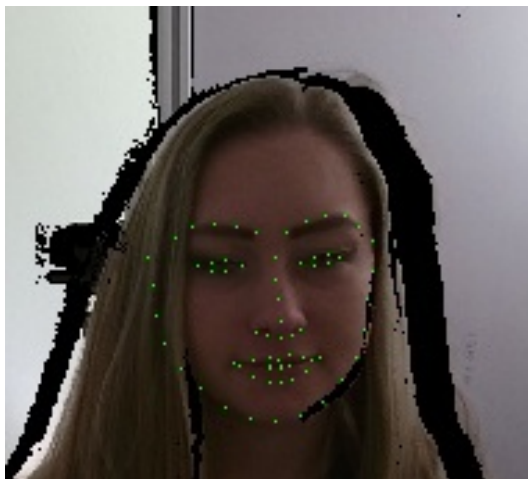
*Obr. 8.21 Korektná detekcia bodov tváre pomocou Dlib č.3*

Samozrejme, aj pri použití knižnice Dlib nastali prípady nie úplne presnej detekcie jednotlivých čŕt, ale tieto rozdiely neboli natoľko výrazné ako pri predošlom spôsobe pomocou knižnice OpenCV. Prípadov nesprávnej detekcie oblasti záujmu v použitých snímkach bolo výrazne menej a väčšinou sa jednalo o minimálne odchýlky od predpokladaných výsledkov, zatiaľ čo pri použití klasifikátorov knižnice OpenCV boli výsledkami vyhľadávania často úplne rozdielne predmety alebo časti obrazu.

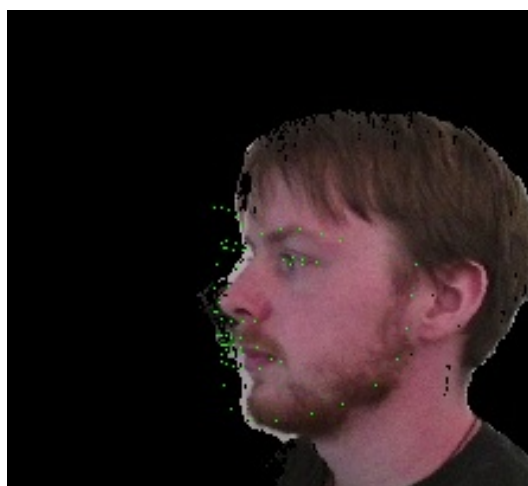




*Obr. 8.22 Nesprávna detekcia bodov tváre pomocou Dlib č.1*



*Obr. 8.23 Nesprávna detekcia bodov tváre pomocou Dlib č.2*



*Obr. 8.24 Nesprávna detekcia bodov tváre pomocou Dlib č.3*

### 8.3 POROVNANIE VÝSLEDKOV OPENCV A DLIB KNIŽNÍC

Veľmi podstatnou súčasťou praktickej časti je samotné porovnanie úspešnosti detekcie kľúčových tvárových črt pomocou využitia OpenCV a Dlib knižnice. Výsledky hodnotenia sú udávané v percentách.

Prvá tabuľka (Tab. 8.1) vyjadruje množstvo správne, respektíve nesprávne určených tvárí v obraze. Už na prvý pohľad je viditeľné, že výsledky Dlib knižnice sú výrazne úspešnejšie. Zatiaľ čo OpenCV knižnica detegovala správne len 16% skúmaných snímok, Dlib knižnica bola úspešná v 86% prípadov. Pre toto vyhodnotenie boli spracované výsledky z celkového počtu 235 snímok.

Tab. 8.1: Percentuálne vyjadrenie úspešnosti vyhľadania tvárových rysov

	Správna detekcia	Nesprávna detekcia
<b>OpenCV knižnica</b>	16%	84%
<b>Dlib knižnica</b>	86%	14%

Druhá tabuľka (Tab. 8.2) zobrazuje percentuálne množstvo správne detegovaných bodov tváre na piatich vybraných vzorkách. Aj v tomto porovnaní vykazovala knižnica Dlib výrazne lepšie výsledky, a to vo všetkých piatich prípadoch.

Tab. 8.2: Percentuálne vyjadrenie množstva detegovaných tvárových rysov

	OpenCV knižnica		Dlib knižnica	
	Správna detekcia	Nesprávna detekcia	Správna detekcia	Nesprávna detekcia
Obrázok č.1	0%	100%	100%	0%
Obrázok č.2	0%	100%	83%	17%
Obrázok č.3	24%	76%	98%	2%
Obrázok č.4	0%	100%	100%	0%
Obrázok č.5	74%	26%	71%	29%

## **8.4 SPÁROVANIE KLÚČOVÝCH RYSOV S RELEVANTNÝM 3D MODELOM**

Posledným bodom diplomovej práce je verifikácia a analýza možností spárovania rysov s relevantným 3D modelom. Táto časť pozostáva z piatich krokov.

V prvom kroku sme použili nami vytvorenú modelovú RGB snímku, v ktorej v ktorej sme aplikovali algoritmus pre detekciu a následné vyrezanie tváre. RGB obraz poskytuje informáciu o farbe snímanej scény založenú na princípe aditívneho RGB modelu.



*Obr. 8.25 Prvý krok spárovania kľúčových rysov s 3D modelom*

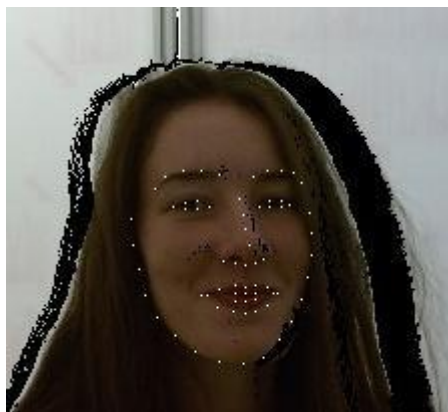
Druhý krok spočíva v rovnakom postupe aplikovanom na obraz nesúci informáciu o hĺbke tzv. D-snímka, pričom je detegovaná a vyrezaná rovnaká oblasť záujmu (ROI – region of interest). Obraz hĺbky je reprezentovaný sivým obrazom, kde hodnota jednotlivých pixelov predstavuje absolútnu vzdialenosť kamier od snímaného objektu.



*Obr. 8.26 Druhý krok spárovania kľúčových rysov s 3D modelom*

Po vykonaní týchto dvoch krokov nám vznikli dve snímky s identickými rozmermi. Každý pixel v prvom obrázku odpovedá príslušnému pixelu v obraze vytvorenom v druhom kroku.

V treťom kroku sme aplikovali algoritmus detekcie kľúčových črt tváre na vyrezanú RGB snímku z kroku jeden.



*Obr. 8.27 Tretí krok spárovania kľúčových rysov s 3D modelom*

Nasledujúci štvrtý krok spočíval v rovnakom postupe ako predošlý, ale jeho realizácia prebiehala na existujúcej D-snímke. Nájdené pozície pixelov v RGB obraze pritom odpovedajú pozíciám pixelov v D-obraze.



*Obr. 8.28 Štvrtý krok spárovania kľúčových rysov s 3D modelom*

V poslednom kroku sme preniesli 2D obraz do 3D obrazu pomocou reprojekcie založenej na transformačných rovniciach. Vstupnými údajmi sú snímky získané z RGB-D kamery. Spracovaním hĺbkovej mapy a RGB snímku do príslušného RGB-D

obrazu sme získali informácie o farbe a hĺbke potrebné pre implementáciu do príslušného 3D (X,Y,Z) priestoru.

Pre realizáciu daných výpočtov je potrebné použitie nasledujúcich rovníc:

$$X = \frac{(x - c_x) \cdot hlbka}{f_x \cdot (X_{\tan\ gential} + X_{radial})} \quad (8.1)$$

$$Y = \frac{(y - c_y) \cdot hlbka}{f_y \cdot (Y_{\tan\ gential} + Y_{radial})} \quad (8.2)$$

$$Z = hlbka \quad (8.3)$$

kde  $c_x$  a  $c_y$  predstavujú hlavné body, zvyčajne orientované v strede obrazu a sú to takzvané vnútorné parametre,  $f_x$  a  $f_y$  predstavujú ohniskové vzdialenosti vyjadrené v pixeloch a  $x, y$  sú polohy pixelov v normalizovaných súradniciach obrazu, pričom parametre  $X$  a  $Y$  sú korekcie skreslenia šošovky. [35]



*Obr. 8.29 Realizácia 3D spracovania na vybranom modeli*

Na predchádzajúcom obrázku (Obr. 8.29) môžeme vidieť realizáciu daného procesu vytvorenia 3D zobrazenia na nami používanom modeli hlavy, zatiaľ čo na

nasledujúcom obrázku (Obr. 8.30) je zobrazený piaty a posledný krok spárovania detegovaných kľúčových rysov s relevantným 3D modelom, pričom horšia kvalita zobrazeného modelu je spôsobená slabým rozlíšením použitých kamier.



*Obr. 8.29 Piaty krok spárovania kľúčových rysov s 3D modelom*

## 9 ZÁVER

Zámerom tejto diplomovej práce je zhodnotenie významu využitia strojového a hlbokého učenia v biomedicínskych aplikáciách a v biomedicínskom zobrazovaní. Konkrétnejšie sa venuje špecifikácii a detekcii kľúčových rysov tváre v 2D obraze pre uľahčenie a urýchlenie automatizovanej diagnostiky obštrukčného spánkového apnoe (OSA). Náplňou práce je výber a porovnanie algoritmov pre vyhľadávanie zvolených rysov tváre. Posledným bodom je verifikácia a analýza možností spárovania rysov s relevantným 3D modelom.

Prvá časť práce sa venuje definícii syndrómu obštrukčného spánkového apnoe, jeho prejavmi, príznakmi a príčinami jeho vzniku. Jednou z príčin je genetická predispozícia založená na rozmeroch tváre pacienta. Práve táto informácia je základným kameňom pre výskum tejto diplomovej práce.

Ďalšia kapitola nadväzuje na tú predošlú, avšak venuje sa technickej stránke danej problematiky. Opisuje strojové rozpoznávanie tvárí, celú históriu a vývoj tohto odvetvia a následne jednotlivé techniky slúžiace na túto detekciu. Pre zhodnotenie, táto práca využíva takzvané tradičné rozpoznávanie, čo znamená, že na detekciu tvárí využíva 2D snímky.

Dôležitým bodom je definícia umelej inteligencie a jej podmnožín strojového a hlbokého učenia, ktoré sú využívané pri trénovaní algoritmov využívaných pre správnu detekciu tváre a jej kľúčových bodov. Do tejto časti patria aj konvolučné neurónové siete, ktoré sú súčasťou hlbokého učenia a sú regulovanou verziou viacvrstvových perceptrónov.

V tejto diplomovej práci sú používané dve softvérové knižnice: Dlib a OpenCV. Obe sú napísané v programovacom jazyku C++. Obidve knižnice ponúkajú viacero metód slúžiacich na detekciu tváre ako celku ale aj vyhľadanie jednotlivých črt. OpenCV knižnica využíva kaskádové klasifikátory Haar, LBP a Hog. Po vykonaní spomínanej detekcie je možné zhodnotiť, že najúspešnejším klasifikátorom bol Haar a najmenej úspešným Hog klasifikátor. Výsledky však neboli uspokojivé ani v prípade použitia LBP klasifikátora, ktorý bol schopný detegovať tvár len na jednej snímke zo 400 použitých exemplárov.

---

Dlib knižnica používa dva detektory. Prvý je založený na metóde MMOD, využívajúci konvolučné neurónové siete na vytrénovanie a detekciu, zatiaľ čo druhý používa už vyššie spomínaný Hog a lineárny klasifikátor. Druhý detektor je starší a je navrhnutý hlavne na frontálnu detekciu tváre. Výhodou však je vyššia rýchlosť vyhľadávania oblastí záujmu. Prvá metóda je novšia a vykazuje výrazne lepšie výsledky aj vďaka možnosti detegovať tvár a črty z viacerých uhlov natočenia.

V praktickej časti tejto práce sa nachádzajú zadané body umiestnené na tvári potrebné pre konkrétnu diagnostiku ochorenia obštrukčného spánkového apnoe a pre porovnanie aj výpis bodov, ktoré je možné detegovať pomocou knižnice Dlib.

Celý proces detekcie realizovanej v tejto diplomovej práci je rozdelený na dve základné časti: detekcia a vyrezávanie tváre a následné vyhľadávanie kľúčových rysov v obraze vzniknutom v prvom kroku.

Posledným bodom práce je spárovanie kľúčových rysov s relevantným 3D modelom. Táto oblasť sa skladá z piatich krokov, ktoré sú v práci opísané a ku každému je priložená aj názorná ukážka.

V závere môžeme konštatovať, že knižnica Dlib a v nej obsiahnuté detektory sa potvrdili ako viditeľne úspešnejšie pri riešení danej problematiky, čo je názorne spracované aj pomocou dvoch tabuliek (Tab. 8.1 a Tab. 8.2) v praktickej časti diplomovej práce.

---



## Zoznam použitej literatúry

1. Sound sleeper – Obstructive Sleep Apnoea (OSA). [Online] 8.Máj 2015. <http://www.soundsleeper.com.au/obstructive-sleep-apnoea>.
  2. **URBAN, N.** Skryté nebezpečenstvá rozpoznávania tváre. [Online] 15.Júl 2017 <https://touchit.sk/skryte-nebezpecenstva-rozpoznavania-tvare/122612>
  3. **DE LEEUW, K.; BERGSTR, J.** The History of Information Security: A Comprehensive Handbook, 2007. s.264-265 ISBN 9780444516084
  4. **GEBHART, A.** Facial recognition: Apple, Amazon, Google and the race for your face. [Online] 18.Marec 2019 <https://www.cnet.com/how-to/facial-recognition-apple-amazon-google-and-the-race-for-your-face-facebook/>
  5. University of Southern California. Mugspot can find a face in the crowd – face recognition software prepares to go to work in the streets [Online] 12.November 2007 <https://www.sciencedaily.com/releases/1997/11/971112070100.htm>
  6. **BRUNELLI, R.** Template Matching Techniques in Computer Vision: Theory and Practice, Wiley, 2009 ISBN 978-0-470-51706-2
  7. **LO, J.** Facial Recognition [Online] 20.December 2018. <https://australiancybersecuritymagazine.com.au/facial-recognition/>
  8. **WILLIAMS, M.** Better Face-Recognition Software [Online] 30.Máj 2007. <https://www.technologyreview.com/s/407976/better-face-recognition-software/>
  9. **ZIBREG, CH.** Qualcomm to bring 3D facial recognition and depth sensing to 2018 Android phones [Online] 22.August 2017. <https://www.idownloadblog.com/2017/08/22/qualcomm-to-bring-3d-facial-recognition-depth-sensing-to-2018-android-phones/>
  10. **BONSOR, K.; JOHNSON, R.** How facial recognition system work [Online] 9.Apríl 2018. <https://electronics.howstuffworks.com/gadgets/high-tech-gadgets/facial-recognition3.htm>
  11. AZoRobotics Army builds face recognition technology that works in low-lights conditions [Online] 18.Apríl 2018. <https://www.azorobotics.com/News.aspx?newsID=9840>
  12. **HOZZA, S.** Arc alapú azonosítás a biometriában [Online] 24.Marec 2014. <https://www.securinfo.hu/termekek/biometria/1160-arc-alapu-azonositas-a-biometriaban.html>
  13. **HEATER, B.** Don't rely on face unlock to keep your phone secure [Online] 2.September 2017. [https://techcrunch.com/2017/09/06/dont-rely-on-face-unlock-to-keep-your-phone-secure/?guccounter=1&guce\\_referrer\\_us=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnLw&guce\\_referrer\\_cs=ES9qVjGvHHEjyCpO3zE8Fw](https://techcrunch.com/2017/09/06/dont-rely-on-face-unlock-to-keep-your-phone-secure/?guccounter=1&guce_referrer_us=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnLw&guce_referrer_cs=ES9qVjGvHHEjyCpO3zE8Fw)
-

14. AI Umelá inteligencia [Online] 2.Február 2019. <https://www.alza.sk/ai-umela-inteligencia#inteligentne-telefony>
  15. **SIGNORELLI, P.** Artificial Intelligence: Transforming the nature of work, learning, and learning to work [Online] 21.August 2018. <https://www.thetambellinigroup.com/artificial-intelligence-transforming-the-nature-of-work-learning-and-learning-to-work/>
  16. **CIBULA, M.** Strojové učenie [Online] 2.Máj 2017. [https://smnd.sk/mcibula/zakl\\_info/definicia.html](https://smnd.sk/mcibula/zakl_info/definicia.html)
  17. Wikipedia: Perceptrón [Online] 7.Marec 2015. <https://sk.wikipedia.org/wiki/Perceptr%C3%B3n>
  18. **SAHA, S.** A Comprehensive guide to Convolutional Neural Networks [Online] 15.December 2018. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
  19. **CORNELISSE, D.** An Intuitive guide to Convolutional Neural Networks [Online] 24.Apríl 2018. <https://medium.freecodecamp.org/an-intuitive-guide-to-convolutional-neural-networks-260c2de0a050>
  20. Wikipedia: Dlib [Online] 11.Marec 2019. <https://en.wikipedia.org/wiki/Dlib>
  21. Dlib FaceLandmark Detector [Online] 2016. <https://enoxsoftware.com/dlibfacelandmarkdetector/documentation/q-a/>
  22. Tutorialspoint: Opencv – Overview [Online] 2019. [https://www.tutorialspoint.com/opencv/opencv\\_overview.htm](https://www.tutorialspoint.com/opencv/opencv_overview.htm)
  23. **MANWANI, N.** Face-Recognition Using OpenCV: A step-by-step guide to build a facial recognition system [Online] 23.Október 2018. <https://hackernoon.com/face-recognition-using-opencv-a-step-by-step-guide-to-build-a-facial-recognition-system-8da97cd89847>
  24. Wikipedia: AdaBoost [Online] 13.Január 2018. [En.wikiipedia.org/wiki/AdaBoos](http://en.wikiipedia.org/wiki/AdaBoos)
  25. Docs, OpenCV. Face Detection Using Haar Cascades. [Online] 4.August 2017. [docs.opencv.org/3.3.0/d7/d8b/tutorial\\_py\\_face\\_detection.html](https://docs.opencv.org/3.3.0/d7/d8b/tutorial_py_face_detection.html)
  26. Mathworks: Train a Cascade Object Detector [Online] 2017. [www.mathworks.com/help/vision/ug/train-a-cascade-object-detector.html](http://www.mathworks.com/help/vision/ug/train-a-cascade-object-detector.html)
  27. **BERGER, W.** Deep Learning Haar Cascade explained [Online] <http://www.willberger.org/cascade-haar-explained/>
  28. **PIETIKAINEN, M.** Local Binary Patterns [Online] 2010. [http://www.scholarpedia.org/article/Local\\_Binary\\_Patterns](http://www.scholarpedia.org/article/Local_Binary_Patterns)
  29. Wikipedia: Local binary patterns [Online] 6.Marec 2019. [https://en.wikipedia.org/wiki/Local\\_binary\\_patterns](https://en.wikipedia.org/wiki/Local_binary_patterns)
  30. **CRUZ, SHIGUEMORI, GUIMARAES** A comparison of Haar-like, LBP and HOG approaches to concrete and asphalt runway detection in high resolution imagery [Online] Júl 2015. [http://epacis.net/jcis/PDF\\_JCIS/JCIS11-art.0101.pdf](http://epacis.net/jcis/PDF_JCIS/JCIS11-art.0101.pdf)
-

31. **YANG, H.; WANG, X.** Cascade classifier for face detection [Online] 31.Júl 2015. <https://journals.sagepub.com/doi/pdf/10.1177/1748301816649073>
32. **KING, D.** Easily Create High Quality Object Detectors with Deep Learning [Online] 11.Október 2016. [http://blog.dlib.net/2016/10/easily-create-high-quality-object.html?fbclid=IwAR23tc5SgeEGoP1-FS4VJdXv7gBFIdnjANrmKo\\_DnFCzI5NbwnMvw4tLmRM](http://blog.dlib.net/2016/10/easily-create-high-quality-object.html?fbclid=IwAR23tc5SgeEGoP1-FS4VJdXv7gBFIdnjANrmKo_DnFCzI5NbwnMvw4tLmRM)
33. **KING, D.** Max-Margin Object Detection [Online] 31.Január 2015. [https://arxiv.org/abs/1502.00046?fbclid=IwAR3LpB-1qtHxjubfL23Sa282EHdts-M1eo-om\\_QT73AB5clop0pJTlf4J6k](https://arxiv.org/abs/1502.00046?fbclid=IwAR3LpB-1qtHxjubfL23Sa282EHdts-M1eo-om_QT73AB5clop0pJTlf4J6k)
34. **ROSEBROCK, A.** Facial landmarks with dlib, OpenCV, and Python [Online] 3.Apríl 2017. <https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/>
35. **VOLAK, J.; KONIAR, D.; JABLONCIK, F.; HARGAŠ, L.** 3D Scanning support for obstructive sleep apnea diagnostics [Online] 2019. <file:///C:/Users/Lucia/Desktop/Downloads/Transcom.pdf>

## ČESTNÉ VYHLÁSENIE

Vyhlasujem, že som zadanú diplomovú prácu vypracoval samostatne, pod odborným vedením vedúceho bakalárskej práce doc. Ing. Dušana Koniara, PhD.a používal som len literatúru uvedenú v práci.

Súhlasím so zapožičiavaním diplomovej práce.

V Žiline dňa 7. 5. 2019

---

podpis

