

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Evidenčné číslo: FEI-100863-111119

AUTENTIFIKÁCIA EMÓCIÍ OPERÁTORA NA ZÁKLADE
VÝRAZU TVÁRE
DIPLOMOVÁ PRÁCA

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Evidenčné číslo: FEI-100863-111119

AUTENTIFIKÁCIA EMÓCIÍ OPERÁTORA NA ZÁKLADE
VÝRAZU TVÁRE
DIPLOMOVÁ PRÁCA

Študijný program:	Robotika a kybernetika
Názov študijného odboru:	kybernetika
Školiace pracovisko:	Ústav robotiky a kybernetiky
Vedúci záverečnej práce:	prof. Ing. Jarmila Pavlovičová, PhD.
Konzultant:	Ing. Michal Tölgyessy, PhD.

Bratislava 2025

Bc. Maroš Kocúr



ZADANIE DIPLOMOVEJ PRÁCE

Študent: **Bc. Maroš Kocúr**
ID študenta: 111119
Študijný program: robotika a kybernetika
Študijný odbor: kybernetika
Vedúci práce: Ing. Michal Tölgyessy, PhD.
Vedúci pracoviska: prof. Ing. František Duchoň, PhD.
Miesto vypracovania: Ústav robotiky a kybernetiky

Názov práce: **Autentifikácia emócií operátora na základe výrazu tváre**

Jazyk, v ktorom sa práca vypracuje: slovenský jazyk

Špecifikácia zadania:

Pri interakcii človeka s robotickým systémom je dôležité, aby robot vedel aj to, v akom emočnom rozpoložení sa operátor nachádza. Cieľom práce je vytvoriť modul, ktorý robotu takúto informáciu poskytne. Predpokladá sa využitie RGB kamery, ale je možnosť využiť aj RGB-D kameru.

Úlohy práce:

1. Analyzujte existujúce metódy analýzy emócií na základe výrazu tváre.
2. Naštudujte princípy tvorby biometrických modelov tváre a metódy detekcie a rozpoznávania tváre.
3. Navrhňte a implementujte systém pre identifikáciu emócií operátora na základe jeho tváre.
4. Testujte a validujte systém na simulovaných aj reálnych dátach.
6. Vytvorte ROS2 balík pre daný systém.
7. Vyhodnoťte experimenty a dosiahnuté výsledky.

Termín odovzdania diplomovej práce: 16. 05. 2025
Dátum schválenia zadania diplomovej práce: 16. 10. 2024
Zadanie diplomovej práce schválil: prof. Ing. Jarmila Pavlovičová, PhD. – garantka študijného programu

SÚHRN

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Študijný program:	Robotika a kybernetika
Autor:	Bc. Maroš Kocúr
Diplomová práca:	Autentifikácia emócií operátora na základe výrazu tváre
Vedúci záverečnej práce:	prof. Ing. Jarmila Pavlovičová, PhD.
Konzultant:	Ing. Michal Tölgyessy, PhD.
Miesto a rok predloženia práce:	Bratislava 2025

Interakcia človeka s robotom v dynamickom prostredí si čoraz viac vyžaduje pochopenie emocionálneho stavu operátora s cieľom optimalizovať komunikáciu a rozhodovacie procesy. Cieľom tejto práce je navrhnúť a implementovať modul, ktorý poskytuje robotickému systému emocionálnu spätnú väzbu a umožňuje mu zisťovať výrazy tváre operátora a odvodzovať jeho emocionálne stavy. S využitím kamery RGB s možnosťou integrácie kamery RGB-D bude systém využívať biometrické modely tváre a techniky rozpoznávania tváre na identifikáciu emócií v reálnom čase. Medzi kľúčové úlohy patrí analýza súčasných metód detekcie emócií výrazu tváre, štúdium princípov tvorby biometrických modelov tváre a implementácia robustného systému na detekciu emócií. Systém bude overený prostredníctvom testovania na simulovaných aj reálnych súboroch údajov. Okrem toho bude vyvinutý balík ROS2, ktorý zabezpečí bezproblémovú integráciu v rámci robotických systémov. Výsledky budú kriticky posúdené prostredníctvom experimentov s cieľom zabezpečiť presnosť a efektívnosť výkonu v reálnych aplikáciách.

Kľúčové slová: RGB kamera, neurónová sieť

ABSTRACT

SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA

FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY

Study Programme:	Robotics and cybernetics
Author:	Bc. Maroš Kocúr
Master's thesis:	Operator Emotion Authentication Based on Facial Expression
Supervisor:	prof. Ing. Jarmila Pavlovičová, PhD.
Consultant:	Ing. Michal Tölgyessy, PhD.
Place and year of submission:	Bratislava 2025

Human-robot interaction in dynamic environments increasingly requires an understanding of the operator's emotional state to optimize communication and decision-making processes. This work aims to design and implement a module that provides emotional feedback to a robotic system, enabling it to detect the operator's facial expressions and infer emotional states. Leveraging an RGB camera, with the option to integrate an RGB-D camera, the system will employ biometric facial models and facial recognition techniques to identify emotions in real-time. Key tasks include analyzing current facial expression emotion detection methods, studying the principles of facial biometric model creation, and implementing a robust system for emotion detection. The system will be validated through testing on both simulated and real datasets. Additionally, a ROS2 package will be developed to ensure seamless integration within robotic systems. The outcomes will be critically assessed through experiments to ensure accuracy and performance efficiency in real-world applications.

Keywords: RGB camera, neural network

Pod'akovanie

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

Obsah

Úvod	1
1 Úvod	2
1.1 Motivácia	3
1.2 Ciele práce	3
2 Teoretické základy	5
2.1 Emócie a ich prejav	5
2.1.1 Univerzálne emócie	5
2.1.2 Kultúrne rozdiely v prejave emócií	5
2.1.3 Výrazy tváre ako indikátory emócií	5
2.2 Analýza obrazu	5
2.2.1 Detekcia tváre	5
2.2.2 Extrakcia príznakov	7
2.2.3 Klasifikácia	7
2.3 Biometria	7
2.3.1 Princípy biometrických systémov	7
2.3.2 Identifikácia vs. verifikácia	8
3 Existujúce metódy analýzy emócií	9
3.1 Ručne značenie	9
3.2 Automatická analýza emócií	9
3.2.1 Konvolučné neurónové siete	9
3.2.2 Typy vhodných neurónových sietí	9
3.2.3 Príklady použitia počítačového videnia	10
4 Návrh riešenia	11
4.1 Porovnanie vybraných metód	11
4.2 Architektúra systému	12
4.3 Výber dát	14
4.4 Extrakcia príznakov	14
4.5 Klasifikácia	16
4.6 Vyber hyperparametrov	16
5 Implementácia riešenia	17

5.1	Výber nástrojov	17
5.2	Implementácia jednotlivých komponentov	17
5.3	Vizualizácia výsledkov	17
6	Implementácia v ROS2	18
6.1	Konverzia modelu	18
6.2	Integrácia do robotického systému	18
7	Exprerimenty a vyhodnotenie	19
7.1	Dátová sada	19
7.2	Metriky	19
7.3	Výsledky	20
7.4	Analýza výsledkov	22
8	Záver	23
8.1	Zhodnotenie práce	23
8.2	Obmedzenia práce	23
8.3	Budúce smerovanie	23
9	Doplňujece poznamky	24
10	Plán práce na ďalší semester	25
	Záver	26
	Zoznam použitej literatúry	27
	Prílohy	29

Zoznam obrázkov a tabuliek

Obrázok 1.1	Schéma systému na rozpoznávanie emócií operátora pomocou RGB kamery.	4
Obrázok 2.1	Detekcia tváre pomocou Haar Cascade.	6
Obrázok 2.2	Detekcia tváre pomocou SSD modelu.	7
Obrázok 4.1	porovnanie rôznych metód na rozpoznávanie emócií. [9]	12
Obrázok 4.2	Architektúra systému na rozpoznávanie emócií operátora pomocou RGB kamery.	13
Obrázok 7.1	Confusion matrix.	19
Obrázok 7.2	Trénovacie grafy.	20
Obrázok 7.3	Examples of different facial expressions.	21

Zoznam algoritmov

Zoznam výpisov

Úvod

Ukazka upraveného šablony FEIStyle.cls (<https://github.com/Kyslik/FEIStyle>) s použitím Times New Roman fontu. Nastavenie fontov som prebral z oficiálneho IEEE šablony, vložil som ho do FEIStyle.cls na riadky 228-230.

Tu bude krásny úvod s diakritikou atď.

A možno aj viac riadkový úvod.

1 Úvod

S rozvojom umelej inteligencie a strojového učenia sa otvárajú nové možnosti pre interakciu medzi človekom a strojom. Jednou z najdôležitejších oblastí výskumu je rozpoznávanie emócií na základe výrazu tváre, ktoré umožňuje strojom porozumieť emocionálnemu stavu používateľ'a. V kontexte robotických systémov je dôležité, aby roboty boli schopné rozoznať emócie človeka, čo môže zlepšiť komunikáciu, kooperáciu a bezpečnosť pri spoločnej práci. [article04]

Psychológ Paul Ekman pomenoval šesť základných emócií - šťastie, smútok, hnev, strach, prekvapenie a znechutenie - ktoré sú často základom pre zavedené kategórie emócií pri rozpoznávaní tváre. Výber týchto emócií bol založený na ich univerzálnom rozpoznávaní a pozorovaní naprieč kultúrami, čo ich predurčuje na použitie v systémoch rozpoznávania tváre.

Podľa Ekmanovho výskumu sa tieto pocity odrážajú v konkrétnych výrazoch tváre, ktoré dokážu automaticky rozpoznať ľudia zo všetkých kultúrnych prostredí. Jeho práca vytvorila fundamentálny základ pre strojové učenie a psychológiu, najmä pri vytváraní systémov, ktoré dokážu dešifrovať výrazy tváre na určenie emocionálneho stavu jednotlivca.

V záujme konzistentnosti a presnosti v aplikáciách, ako je zdravotníctvo, robotika a služby zákazníkom, je možné do systémov rozpoznávania tváre zahrnúť konzistentnú metódu analýzy emócií. Pochopenie úrovne spokojnosti alebo podráždenia používateľ'a môže napríklad pomôcť upraviť reakcie systému a zlepšiť výsledky interakcií. [1]

Emócie zohrávajú dôležitú úlohu v procese rozhodovania, riadenia a interakcie. Schopnosť robotického systému porozumieť emocionálnemu stavu používateľ'a umožňuje jeho prispôbenie konkrétnym podmienkam a potrebám operátora. Napríklad v priemysle môžu robotické systémy identifikovať stres alebo únavu operátora, čím prispievajú k zvýšeniu bezpečnosti a efektivity. Okrem toho, v oblasti zdravotnej starostlivosti môže rozpoznávanie emócií pomôcť monitorovať psychický stav pacientov a prispieť k ich lepšej starostlivosti. [2] []

Rozpoznávanie emócií je možné dosiahnuť rôznymi metódami, ktoré zahŕňajú spracovanie obrazu, analýzu textu, reč a gestá. Výraz tváre je však najvýznamnejším a najpresnejším indikátorom emócií, pretože vyjadruje okamžitý emocionálny stav človeka. Emócie, ako sú šťastie, smútok, hnev alebo prekvapenie, sú viditeľné prostredníctvom zmien vo svaloch tváre, ktoré sú merateľné a analyzovateľné pomocou technológií strojového učenia, najmä pomocou hlbokých neurónových sietí (CNN).[article04]

Súčasná metóda na rozpoznávanie emócií zahŕňa viacero prístupov. Tradičné prístupy, ako napríklad metódy založené na geometrických črtách a textúrach, boli doplnené modernými metódami založenými na hlbokom učení, ktoré dosahujú vysokú presnosť. Neurónové siete sú

schopné automaticky extrahovať črty tváre bez potreby manuálneho zásahu, čo výrazne zvyšuje efektivitu systému. Tieto pokročilé modely dosahujú vysokú mieru úspešnosti v rôznych aplikáciách, ako sú zdravotná starostlivosť, priemyselná automatizácia alebo monitorovanie únavy vodičov. [2] [3]

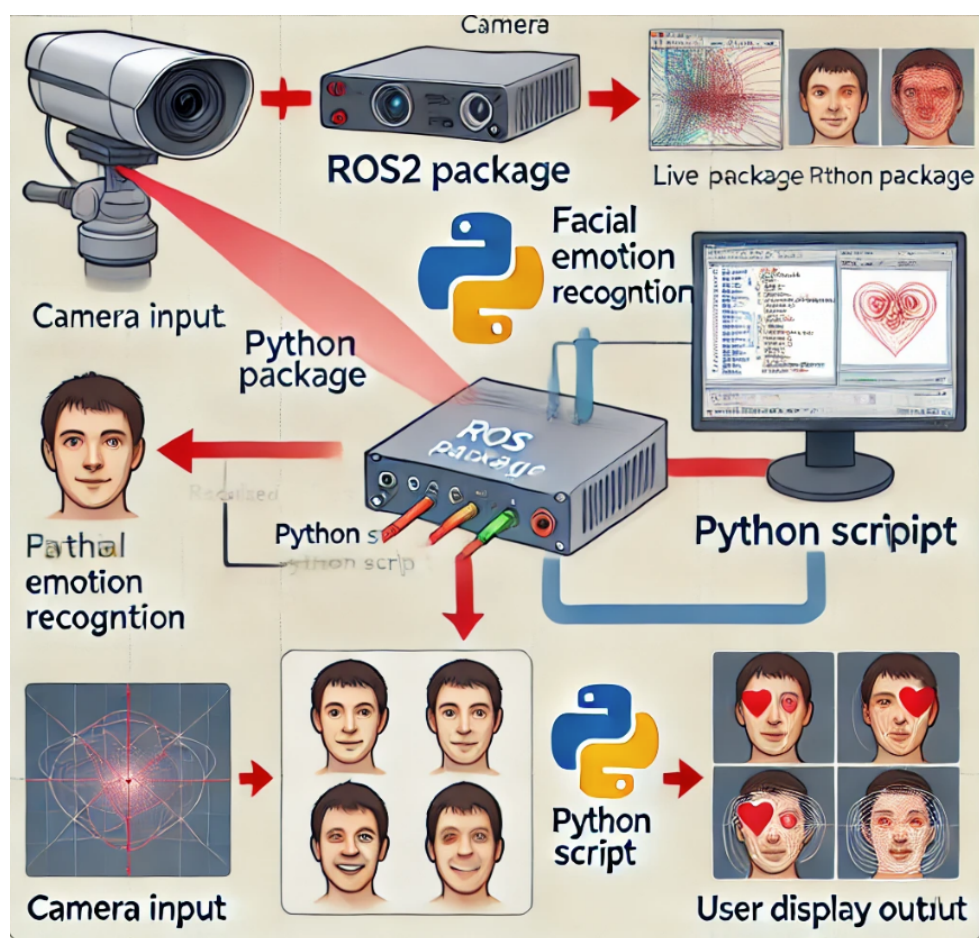
1.1 Motivácia

Motiváciou pre rozpoznávanie emócií tváre je jeho potenciál zlepšiť interakciu medzi človekom a počítačom, zlepšiť monitorovanie duševného zdravia a vytvoriť adaptívne systémy pre rôzne oblasti, ako je vzdelávanie, marketing a robotika. [4]

1.2 Ciele práce

Táto práca sa zameriava na návrh systému na rozpoznávanie emócií operátora pomocou RGB kamery, ktorý umožní robotickým systémom analyzovať a prispôbiť sa emocionálnemu stavu používateľa v reálnom čase. Systém bude testovaný na simulovaných aj reálnych dátach a integrovaný do robotických platforiem cez ROS2 pre jeho nasadenie v priemyselných a zdravotných aplikáciách.

Cieľom práce je vytvoriť systém, ktorý bude schopný rozpoznať emócie v reálnom čase.



Obr. 1.1: Schéma systému na rozpoznávání emocí operátora pomocí RGB kamery.

2 Teoretické základy

2.1 Emócie a ich prejav

Emócie sú komplexné psychologické stavy, ktoré zahŕňajú subjektívne zážitky, fyziologické reakcie a behaviorálne prejavy. V priebehu výskumu boli emócie definované rôznymi spôsobmi, ale všeobecne sa považujú za reakcie na podnety, ktoré ovplyvňujú ľudské správanie a myslenie. Emócie môžu byť pozitívne alebo negatívne a ovplyvňujú naše rozhodovanie, pamäť a vnímanie sveta okolo nás. [4]

2.1.1 Univerzálne emócie

Jednou z najvýznamnejších teórií o emóciách je teória univerzálnych emócií, ktorú vyvinul psychológ Paul Ekman. Podľa tejto teórie existuje šesť základných emócií, ktoré sú univerzálne rozpoznateľné na základe výrazu tváre: radosť, smútok, hnev, prekvapenie, strach a odpor. Tieto emócie sú nezávislé od kultúrnych vplyvov a prejavujú sa podobným spôsobom naprieč rôznymi kultúrami a etnickými skupinami. [2]

2.1.2 Kultúrne rozdiely v prejave emócií

Napriek existencii univerzálnych emócií existujú významné kultúrne rozdiely v tom, ako sú emócie prejavované a vnímané. Niektoré kultúry, ako napríklad západné, sú viac orientované na individualizmus, kde je prejav emócií otvorenejší a priamy, zatiaľ čo v kolektivistických kultúrach, ako sú východné ázijské krajiny, sú emócie častejšie potláčané alebo prejavované menej intenzívne. [4]

2.1.3 Výrazy tváre ako indikátory emócií

Výraz tváre je jedným z hlavných spôsobov, ako sú emócie vonkajšie prejavované. Svalové pohyby tváre, ktoré zahŕňajú zmeny v oblasti očí, obočia, úst a líc, sú kľúčovými indikátormi emočných stavov. Tento typ neverbálnej komunikácie je extrémne efektívny, pretože umožňuje okamžitý a intuitívny prenos emocionálnych informácií [5]. Výskum ukázal, že až 55 % emočných informácií je prenášaných prostredníctvom výrazov tváre, čo zdôrazňuje ich význam v sociálnej interakcii. [2]

2.2 Analýza obrazu

Analýza obrazu je kľúčová pre proces rozpoznávania emócií na základe tváre. Tento proces zahŕňa detekciu tváre, extrakciu príznačkov a následnú klasifikáciu emócií

2.2.1 Detekcia tváre

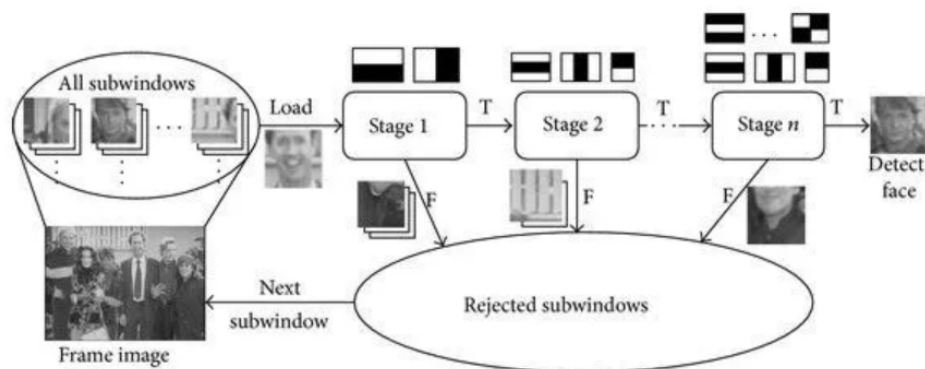
Detekcia tváre je prvým krokom v procese rozpoznávania emócií. Tento krok zahŕňa lokalizáciu tváre v obraze a je rozhodujúci pre ďalšie spracovanie. Moderné metódy detekcie

tváre, ako je algoritmus Viola-Jones, používajú rýchle a efektívne prístupy k lokalizácii tvárových oblastí, čo je nevyhnutné pre následné kroky. Vývoj hlbokých neurónových sietí, ako sú konvolučné neurónové siete (CNN), výrazne zlepšil presnosť detekcie tváre, čo umožnilo rozpoznávať tváre aj v rôznych svetelných podmienkach a uhloch.[2]

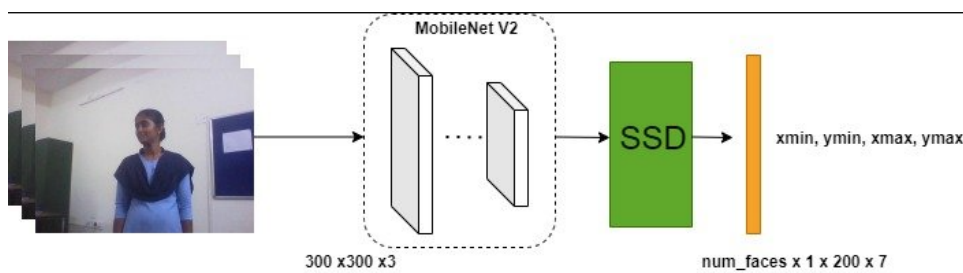
Metódy:

Metóda Haar Cascade využíva sériu funkcií odvodených z Haar-like vzorov, ktoré sa používajú na identifikáciu oblastí záujmu. Tieto vzory umožňujú rýchlu detekciu, pričom proces je optimalizovaný pomocou "Integral Image". Napriek svojej robustnosti pri čelných záberoch sme identifikovali výzvy pri zmene osvetlenia a uhla záberu, ktoré sme riešili dodatočnou normalizáciou jasu a kontrastu. Algoritmus využíva preddefinované Haar-like príznaky (ako vzory okienka pre kontrastné oblasti, napr. oči vs. líca). Používa koncept Integral Image, ktorý znižuje výpočtovú náročnosť výberu príznakov. Tréning klasifikátora sa vykonáva pomocou algoritmu AdaBoost, ktorý identifikuje najvýznamnejšie príznaky. Detekcia prebieha cez viacero okienok "rôznej veľkosti na snímke.

Detekcia tváre pomocou face_net (SSD model) využíva predtrénovaný model SSD (Single Shot MultiBox Detector) s architektúrou Caffé na detekciu tvárí. Táto metóda je modernejšia a efektívnejšia oproti Haar Cascade, pretože využíva hlboké neurónové siete na extrakciu robustných príznakov. Princíp fungovania je optimalizovaný na detekciu viacerých objektov naraz, pričom predikcie sú vykonávané na rôznych úrovniach vrstiev CNN. Model využíva vstupnú veľkosť obrazu 300x300 pixelov a predspracovanie obrazu na odstránenie šumu (normalizácia s hodnotami RGB: 104.0, 177.0, 123.0). Výstup modelu obsahuje confidence score (pravdepodobnosť detekcie) a súradnice ohraničujúceho boxu.



Obr. 2.1: Detekcia tváre pomocou Haar Cascade.



Obr. 2.2: Detekcia tváre pomocou SSD modelu.

2.2.2 Extrakcia príznačkov

Toto dat asi jak podkapitotlu ku detekcii tvare pomocou CV a NN. Po detekcii tváre nasleduje extrakcia príznačkov, kde sú identifikované kľúčové črty tváre, ako sú oči, nos, ústa a obočie. Tieto črty sú dôležité pre analýzu výrazov tváre, pretože zmeny v ich polohe alebo napätí súvisia s rôznymi emočnými stavmi. [5] Typické algoritmy používané pri extrakcii príznačkov zahŕňajú Gaborove filtre a histogramy orientovaných gradientov (HOG), ktoré zameriavajú pozornosť na zmeny v textúre a tvaroch.[2].

2.2.3 Klasifikácia

Klasifikácia emócií je záverečným krokom, kde sú extrahované príznačky spracované a priradené k určitým emočným kategóriám. Moderné metódy klasifikácie používajú algoritmy strojového učenia, ako sú Support Vector Machines (SVM), ale najúčinnnejšie sú konvolučné neurónové siete (CNN), ktoré dokážu automaticky klasifikovať výrazy do kategórií, ako sú šťastie, smútok alebo hnev. [5] [2]

2.3 Biometria

Biometria sa zaoberá rozpoznávaním osôb na základe jedinečných fyziologických alebo behaviorálnych charakteristík. V oblasti rozpoznávania tváre ide o identifikáciu alebo verifikáciu osôb na základe tvárových črt. [5] [2]

2.3.1 Princípy biometrických systémov

Biometrické systémy sú založené na zhromažďovaní a analýze údajov, ktoré sú pre jednotlivca jedinečné, ako sú odtlačky prstov, dúhovka alebo tvár. Tieto systémy musia byť schopné spoľahlivo identifikovať alebo overiť osobu na základe týchto údajov. V kontexte rozpoznávania tváre systém spracováva obraz tváre, extrahuje relevantné črty a porovnáva ich s uloženými údajmi. [4]

2.3.2 Identifikácia vs. verifikácia

Identifikácia a verifikácia sú dva hlavné prístupy v biometrických systémoch. Identifikácia zahŕňa určenie identity osoby na základe údajov o tvári v porovnaní s databázou, zatiaľ čo verifikácia porovnáva údaje jednej osoby s predtým zaznamenanými údajmi, aby potvrdila, či ide o tú istú osobu[5]. Rozpoznávanie tváre je často používané v aplikáciách na bezpečnosť, kde verifikácia hrá kľúčovú úlohu pri autentifikácii používateľov, zatiaľ čo identifikácia sa používa na vyhľadávanie osôb v rozsiahlych databázach.[2]

3 Existujúce metódy analýzy emócií

V oblasti rozpoznávania emócií na základe výrazu tváre existuje mnoho prístupov, ktoré môžeme rozdeliť na manuálne a automatizované metódy. Kým tradičné manuálne prístupy spočívajú v ručnom označovaní výrazov tváre, moderné metódy využívajú automatické algoritmy, často založené na neurónových sieťach (NN).

3.1 Ručne značenie

Ručne značenie (manuálna anotácia) spočíva v označovaní kľúčových bodov na tvári a následnom priradení výrazov tváre k určitým emočným kategóriám. Tento proces je časovo náročný a vyžaduje expertov na interpretáciu dát. Avšak, ručné značenie je stále dôležité pre tvorbu datasetov, ktoré sú nevyhnutné na tréning automatických systémov. Dôležité datasetové projekty, ako sú Cohn-Kanade alebo AffectNet, sa opierajú o ručné značenie výrazov tváre. Manuálna anotácia má významnú úlohu v počiatočných fázach výskumu, ale pre aplikácie, ktoré vyžadujú veľké množstvo dát, je neefektívna. [4]

3.2 Automatická analýza emócií

Automatická analýza emócií využíva pokročilé algoritmy počítačového videnia a strojového učenia, aby bola schopná rozpoznať emócie na základe výrazu tváre bez potreby manuálneho zásahu. Moderné systémy rozpoznávania emócií sa vo veľkej miere spoliehajú na neurónové siete (NN), najmä na konvolučné neurónové siete (CNN), ktoré dokážu automaticky extrahovať a klasifikovať príznaky výrazu tváre.

3.2.1 Konvolučné neurónové siete

Neurónové siete sú inšpirované biologickými mozgovými štruktúrami a sú schopné učiť sa z obrovských množstiev dát. Pre úlohy rozpoznávania obrazu, vrátane rozpoznávania emócií, sú najbežnejšie využívané konvolučné neurónové siete (CNN). CNN majú schopnosť automaticky extrahovať črty tváre bez potreby manuálnej definície a v kombinácii s ďalšími typmi sietí, ako sú rekurentné neurónové siete (RNN) alebo dlhodobé pamäte (LSTM), umožňujú ešte lepšiu interpretáciu časovo premenlivých dát, ako sú sekvencie výrazov tváre. [4] [6]

3.2.2 Typy vhodných neurónových sietí

Pre aplikácie rozpoznávania emócií sa osvedčili rôzne typy neurónových sietí:

Konvolučné neurónové siete (CNN): CNN sa často používajú na extrakciu priestorových príznakov z obrazov tváre, ako sú oči, ústa a obočie. CNN sú obzvlášť účinné pri identifikácii týchto príznakov z rôznych uhlov a svetelných podmienok. [7]

Rekurentné neurónové siete (RNN) a LSTM: Tieto siete sú vhodné pre analýzu sek-

vencií, ako sú videozáznamy alebo opakujúce sa výrazy tváre. Použitím týchto sietí je možné zohľadniť časové zmeny v tvári, čo je dôležité pre interpretáciu dynamických emócií. [8]

Deep Convolutional Neural Networks (DCNN): DCNN je špeciálny typ CNN, ktorý dosahuje vysokú presnosť v úlohách rozpoznávania emócií, najmä pri kombinácii s technológiami počítačového videnia.

3.2.3 Príklady použitia počítačového videnia

rozpoznávanie emócií je úzko prepojené s oblasťou počítačového videnia. Počítačové videnie používa algoritmy na interpretáciu vizuálnych informácií. V súčasnosti sa CNN často kombinujú s technológiami, ako sú techniky extrakcie príznakov (napr. HOG alebo SIFT), aby sa zlepšila presnosť rozpoznávania výrazu tváre. Tieto systémy sú schopné identifikovať a klasifikovať tvárové príznaky aj v náročných podmienkach, ako sú premenlivé svetelné podmienky alebo čiastočné zakrytie tváre.[**Huang2023**]

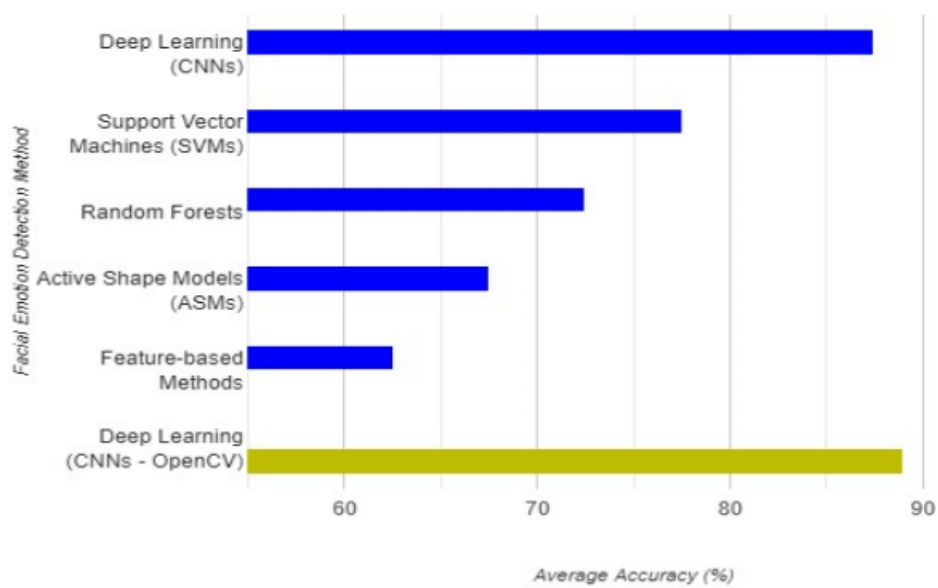
4 Návrh riešenia

4.1 Porovnanie vybraných metód

V dokumente sa ako hlavné metódy použili kombinácia OpenCV a konvolučných neurónových sietí (CNN), pričom tieto technológie boli zvolené kvôli schopnosti riešiť výzvy spojené so zmenami osvetlenia, rôznymi pózami a potrebou rýchleho spracovania v reálnom čase. OpenCV slúžil na efektívne predspracovanie dát a detekciu tvárí pomocou Haar cascade classifier, zatiaľ čo CNN umožnili extrakciu komplexných črt a klasifikáciu emócií. Ako základ bola použitá predtrénovaná architektúra VGGFace, ktorá výrazne zlepšila presnosť systému a skrátila čas potrebný na tréning. Model bol trénovaný na datasete FER2013, ktorý obsahoval obrázky kategorizované podľa siedmich základných emócií (šťastie, smútok, hnev, prekvapenie, strach, znechutenie a neutrálne). Na optimalizáciu modelu bol použitý Adam algoritmus a cross-entropy loss funkcia. Výsledný systém dosiahol vysokú presnosť 95,2 % a bol schopný spracovať dáta v reálnom čase pri rýchlosti 25 snímok za sekundu na priemernej PC zostave. V porovnaní s tradičnými metódami, ako sú ručne vytvárané črty spojené so strojovým učením, sa tento prístup ukázal ako oveľa presnejší a robustnejší, čím ponúka široké možnosti využitia v oblastiach ako sociálna robotika, zdravotníctvo a interakcia človek-stroj.[9]

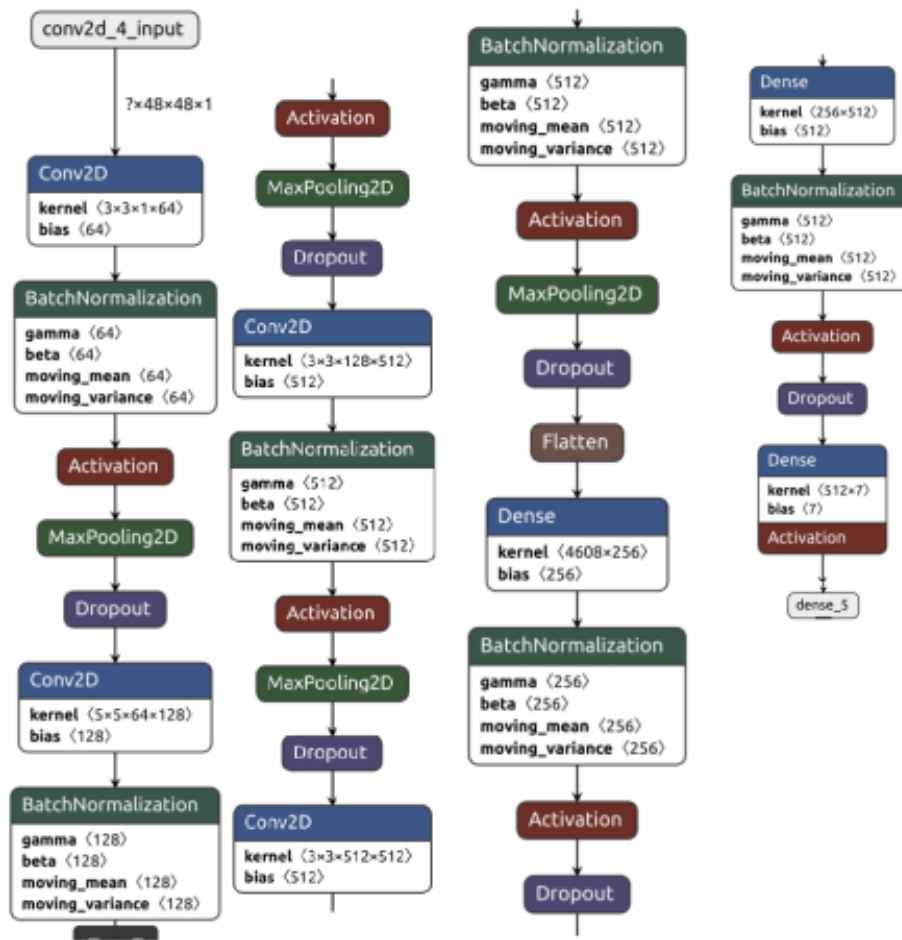
V experimentoch bol model testovaný na klasifikáciu emócií, pričom dosiahnuté výsledky ukázali rôznu úroveň presnosti pre jednotlivé emócie. Najvyššiu presnosť dosiahol model pri detekcii neutrálneho výrazu, a to 88,5 %, čo je pripisované jasným črtám a minimálnej nejednoznačnosti tohto výrazu. Pre kategóriu šťastia bola presnosť 85,2 %, pričom niektoré chyby mohli byť spôsobené vplyvom osvetlenia. Klasifikácia smútku dosiahla presnosť 82,7 %, čo ukazuje, že model si poradí aj so subtílnymi črtami, ako sú klesnuté kútiky úst či slzy. Pri kategórii hnevu bola presnosť 79,4 %, kde variácie intenzity emócie predstavovali výzvu.[9]

Pre emóciu prekvapenia dosiahol model presnosť 77,1 %, avšak jej prechodná povaha a možná podobnosť s inými emóciami spôsobovali ťažkosti. Najnižšiu presnosť model vykazoval pri kategóriách strachu (74,8 %) a znechutenia, ktoré sú charakteristické jemnými a komplexnými výrazmi tváre. Tieto výsledky naznačujú, že model je veľmi efektívny pri detekcii výrazných emócií, no u jemnejších a zriedkavejších prejavov si vyžaduje ďalšiu optimalizáciu. [9]



Obr. 4.1: porovnanie rôznych metód na rozpoznávanie emócií. [9]

4.2 Architektúra systému



Obr. 4.2: Architektúra systému na rozpoznávanie emócií operátora pomocou RGB kamery.

4.3 Výber dát

Dataset FER2013 (Facial Expression Recognition 2013) je jedným z najčastejšie používaných datasetov na tréovanie a testovanie modelov na rozpoznávanie emócií. Obsahuje 35 887 obrázkov v odtieňoch sivej, pričom každý má rozlíšenie 48x48 pixelov. Obrázky sú kategorizované do siedmich emócií: šťastie, smútok, hnev, prekvapenie, strach, znechutenie a neutrálne. Dataset je rozdelený na tréovaciu, validačnú a testovaciu časť, čo umožňuje komplexné testovanie a optimalizáciu modelov. Najviac obrázkov obsahuje kategória „šťastie“, zatiaľ čo kategória „znechutenie“ má najmenej obrázkov, čo vedie k nerovnomernému rozloženiu dát medzi triedami.

Ďalšie podobné datasety

- **AffectNet:** Obsahuje viac ako 1 milión obrázkov, ktoré sú anotované na 11 kategórií emócií vrátane neutrálnych a jemných variácií ako „záujem“ či „zvedavosť“.
- **CK+ (Cohn-Kanade Extended):** Obsahuje 593 sekvencií tvárových výrazov, ktoré pokrývajú 7 základných emócií. Obrázky sú vysokej kvality a často sa používajú na analýzu prechodov medzi neutrálnymi a výrazovými stavmi.
- **KDEF (Karolinska Directed Emotional Faces):** Obsahuje 4900 obrázkov od 70 jedincov, pričom každý jedinec zobrazuje 7 emócií. Je špecificky navrhnutý na psychologické experimenty a rozpoznávanie emócií.
- **EmoReact:** Dataset určený na rozpoznávanie emócií z videí, obsahuje multimodálne dáta vrátane audio a vizuálnych záznamov. Je užitočný pri štúdiu komplexnejších interakcií medzi emóciami a dynamickými prejavmi.

Datasety ako FER2013 a AffectNet sú preferované pre svoju veľkosť a rozmanitosť, zatiaľ čo menšie datasety ako JAFFE a CK+ sa často používajú na overenie základných hypotéz a pilotné štúdie. Výber datasetu závisí od konkrétneho cieľa výskumu a požiadaviek na presnosť a robustnosť modelu.

4.4 Extrakcia príznakov

Extrakcia príznakov je kľúčovým krokom v procese analýzy tvárových emócií, pretože umožňuje modelu identifikovať a rozpoznať črty tváre, ktoré súvisia s konkrétnymi emóciami. Tento proces sa zvyčajne skladá z niekoľkých fáz, ktoré sú založené na kombinácii počítačového videnia a metód strojového učenia:

1. **Predspracovanie obrazu:** Pred extrakciou sa obraz upravuje na zjednodušenie analýzy:

- **Normalizácia:** Normalizácia farieb a kontrastu na zjednotenie vlastností obrazu.
 - **Orezanie:** Odstránenie okrajov a pozadia, aby sa zvýraznili relevantné časti tváre.
 - **Zmena veľkosti:** Zmena rozlíšenia obrazu na zjednotenie veľkosti vstupov.
2. **Detekcia kľúčových bodov tváre:** V tejto fáze sa identifikujú charakteristické body tváre, ako sú oči, nos, ústa a obočie:
 - **Algoritmy detekcie:** Bežne sa používajú metódy ako Haar Cascade, Dlib, MTCNN, OpenPose, SSD model. Tieto metódy poskytujú presné súradnice bodov tváre, ktoré sú základom na analýzu zmien v jednotlivých oblastiach.
 3. **Extrakcia príznakov:** Geometrické črty sa zameriavajú na vzdialenosti, uhly a proporcie medzi rôznymi bodmi na tvári. Tieto črty sú často použité na identifikáciu rozdielov medzi neutrálnymi a emocionálnymi výrazmi.
 4. **Použitie textúrových vlastností:** Textúrové vlastnosti poskytujú detailnejšie informácie o štruktúre tváre. Často sa využívajú metódy ako Gaborove filtre, HOG alebo LBP na extrakciu textúrnych príznakov.
 5. **Využitie konvolučných neurónových sietí (CNN):** Moderné prístupy nahrádzajú manuálnu extrakciu črt konvolučnými neurónovými sieťami:
 - **Konvolučné vrstvy:** Automaticky detegujú a extrahujú črty relevantné pre konkrétne úlohy.
 - **Pooling vrstvy:** Znižujú rozmer vstupov a zvyšujú invariantnosť voči posunom a zmenám veľkosti.
 - **ReLU aktivačné funkcie:** Zvyšujú nelinearitu a umožňujú modelu učiť sa zložitejšie vzory.
 - **Úplne prepojené vrstvy:** Kombinujú extrahované príznaky a klasifikujú ich do konkrétnych kategórií.
 6. **Tvorba príznakového vektora:** Všetky extrahované črty (geometrické, textúrové, výstupy CNN) sa kombinujú do vysokodimenziálneho vektora, ktorý reprezentuje obraz tváre. Tento vektor je vstupom pre klasifikátor (napr. Softmax, SVM), ktorý priradí tvár k jednej z kategórií emócií.

Extrahované črty sú základom na rozpoznávanie jemných rozdielov medzi emóciami, ako napríklad zmena polohy obočia môže naznačovať hnev, zatiaľ čo stlačenie pier môže naznačovať

smútok. Kombinácia črt a textúr umožňuje modelu pochopiť komplexné výrazy a správne ich kategorizovať. Táto postupnosť krokov zaisťuje, že model má prístup k najrelevantnejším informáciám z obrazu, čo zvyšuje presnosť klasifikácie emócií.

4.5 Klasifikácia

4.6 Vyber hyperparametrov

5 Implementácia riešenia

5.1 Výber nástrojov

Programovací jazyk, knižnice (OpenCV, TensorFlow, PyTorch).

5.2 Implementácia jednotlivých komponentov

Podrobný popis implementácie.

5.3 Vizualizácia výsledkov

Vizualizácia výsledkov analýzy. Grafy, tabuľky.

6 Implementácia v ROS2

6.1 Konverzia modelu

Konverzia trénovaného modelu do formátu vhodného pre ROS2.

6.2 Integrácia do robotického systému

Popis integrácie do ROS2, komunikácia s ostatnými modulmi.

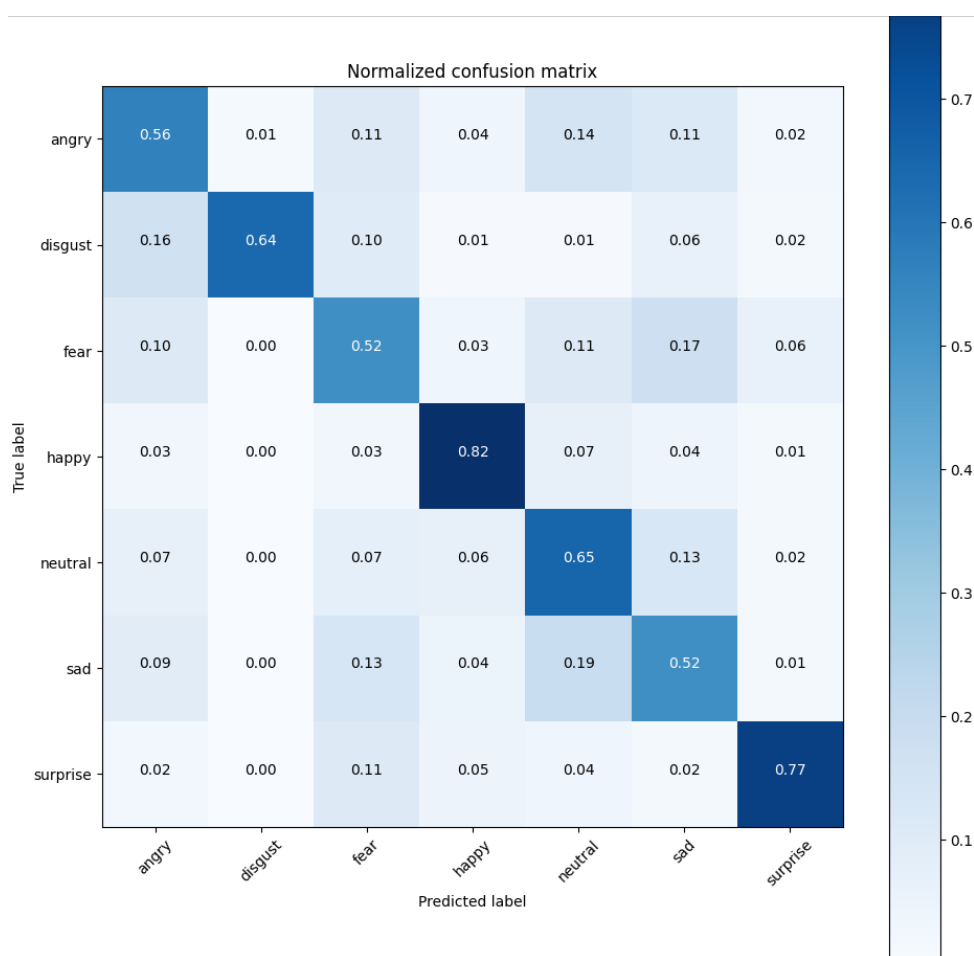
7 Experimenty a vyhodnotenie

7.1 Dátová sada

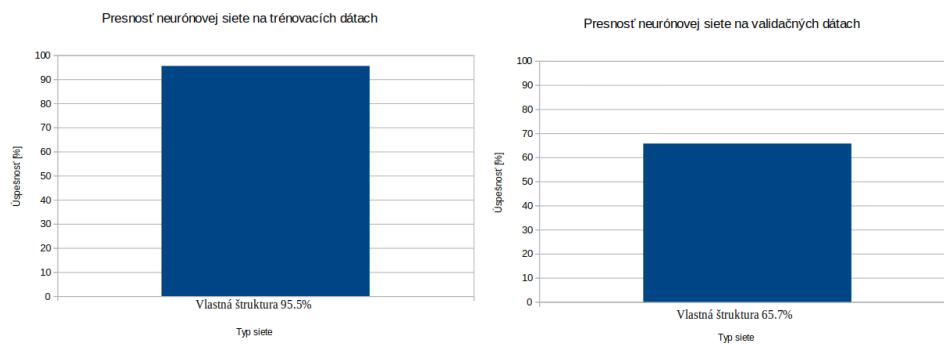
Popis použitého dataset-u (veľkosť, rozdelenie tried, kvalita).

7.2 Metriky

Výber vhodných metrik (presnosť, úplnosť, F1-skóre, ROC krivka).



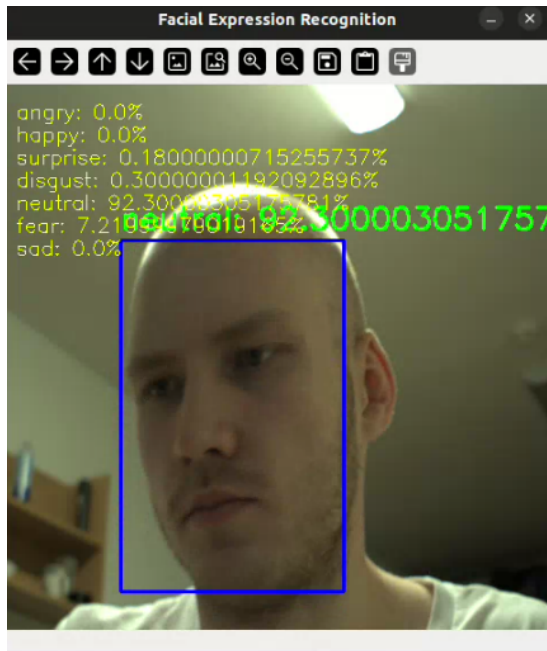
Obr. 7.1: Confusion matrix.



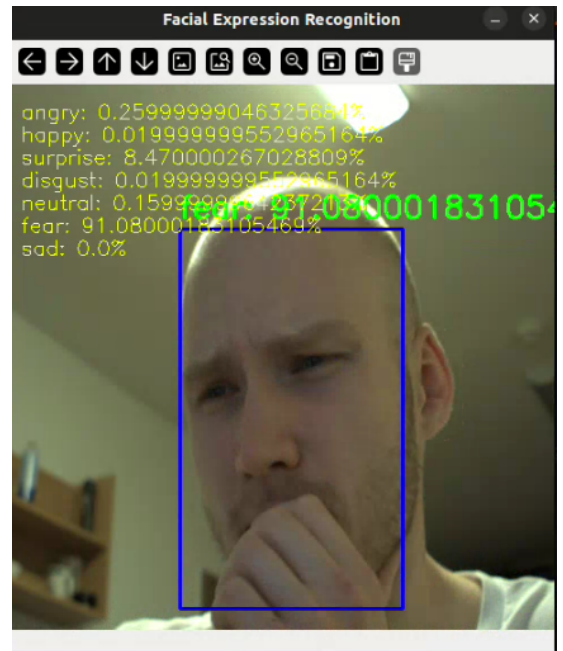
Obr. 7.2: Trénovacie grafy.

7.3 Výsledky

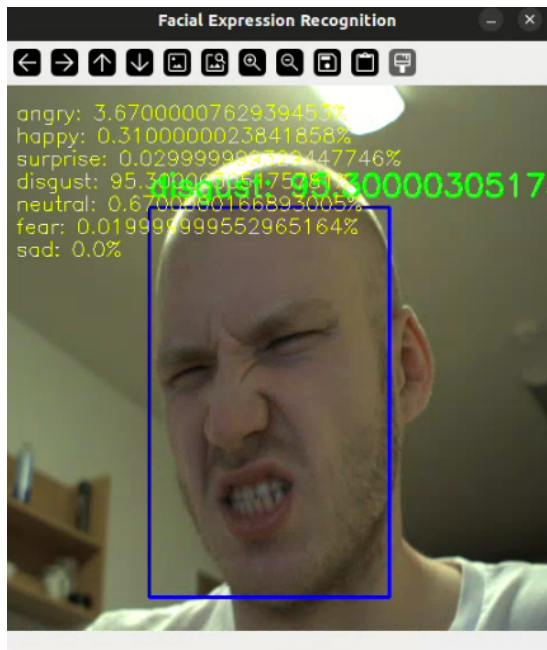
Vyhodnotenie výsledkov experimentov. Prehľadné zhrnutie výsledkov, porovnanie s inými prácami.



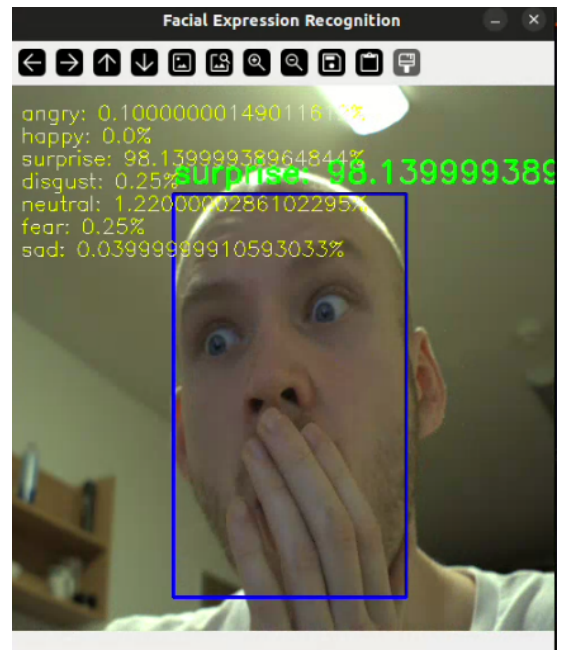
(a) Neutral



(b) Fear



(c) Disgust



(d) Surprise

Obr. 7.3: Examples of different facial expressions.

7.4 Analýza výsledkov

Analýza výsledkov, príčiny chýb, možné zlepšenia.

8 Záver

8.1 Zhodnotenie práce

Zhodnotenie dosiahnutých výsledkov.

8.2 Obmedzenia práce

Obmedzenia práce, možné zlepšenia.

8.3 Budúce smerovanie

Možné smerovanie ďalej práce.

9 Doplnujece poznamky

Literatúra: Pravidelne citujte relevantnú literatúru. Obrázky a diagramy: Používajte obrázky a diagramy na ilustráciu komplexných konceptov. Kód: Ak je to možné, pridajte ukážky kódu. Tabuľky: Používajte tabuľky na porovnanie výsledkov. Táto štruktúra poskytuje komplexný rámec pre vašu prácu. Môžete ju prispôbiť podľa svojich konkrétnych potrieb a zistení.

10 Plán práce na ďalší semester

- Implementácia inej štruktúry neuronovej siete pre dosiahnutie lepších výsledkov.
- Testovanie systému na simulovaných a reálnych dátach.
- Integrácia systému do robotických platforiem.
- Experimenty a vyhodnotenie výsledkov.
- Napísanie záverečnej práce.
- Obhajoba záverečnej práce.

Záver

Conclusion is going to be where?

Here.

Zoznam použitej literatúry

1. EKMAN, Paul. Are there basic emotions? *Psychological Review*. 1992, roč. 99, č. 3, s. 550–553. Dostupné z DOI: 10.1037/0033-295x.99.3.550.
2. BISOGNI, Carmen, CASTIGLIONE, Aniello, HOSSAIN, Sanoar, NARDUCCI, Fabio a UMER, Saiyed. Impact of Deep Learning Approaches on Facial Expression Recognition in Healthcare Industries. *IEEE Transactions on Industrial Informatics*. 2022, roč. 18, č. 8, s. 5619–5627. Dostupné z DOI: 10.1109/TII.2022.3141400.
3. MARTINEZ, Brais a VALSTAR, Michel F. *Advances in Face Detection and Facial Image Analysis*. Advances, Challenges, and Opportunities in Automatic Facial Expression Recognition. Ed. KAWULOK, Michal, CELEBI, M. Emre a SMOLKA, Bogdan. Cham: Springer International Publishing, 2016. ISBN 978-3-319-25958-1. Dostupné z DOI: 10.1007/978-3-319-25958-1_4.
4. CANAL, Felipe Zago, MÜLLER, Tobias Rossi, MATIAS, Jhennifer Cristine, SCOTTON, Gustavo Gino, DE SA JUNIOR, Antonio Reis, POZZEBON, Eliane a SOBIERANSKI, Antonio Carlos. A survey on facial emotion recognition techniques: A state-of-the-art literature review. *Information Sciences*. 2022, roč. 582, s. 593–617. ISSN 0020-0255. Dostupné z DOI: <https://doi.org/10.1016/j.ins.2021.10.005>.
5. WANG, Huan-Huan a GU, Jing-Wei. The Applications of Facial Expression Recognition in Human-computer Interaction. In: *2018 IEEE International Conference on Advanced Manufacturing (ICAM)*. 2018, s. 288–291. Dostupné z DOI: 10.1109/AMCON.2018.8614755.
6. ROY, Arnab Kumar, KATHANIA, Hemant Kumar, SHARMA, Adhitiya, DEY, Abhishek a ANSARI, Md. Sarfaraj Alam. *ResEmoteNet: Bridging Accuracy and Loss Reduction in Facial Emotion Recognition*. 2024. Dostupné z arXiv: 2409.10545 [cs.CV].
7. ZHANG, Saining, ZHANG, Yuhang, ZHANG, Ye, WANG, Yufei a SONG, Zhigang. A Dual-Direction Attention Mixed Feature Network for Facial Expression Recognition. *Electronics*. 2023, roč. 12, č. 17. ISSN 2079-9292. Dostupné z DOI: 10.3390/electronics12173595.
8. KO, Byoung Chul. A Brief Review of Facial Emotion Recognition Based on Visual Information. *Sensors*. 2018, roč. 18, č. 2. ISSN 1424-8220. Dostupné z DOI: 10.3390/s18020401.

9. PRADEEP, V, MADHUSHREE, SUMUKHA, B S, RICHARDS, Gary R a PRASHANT, Shetty Prajakta. Facial Emotion Detection using CNN and OpenCV. In: *2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS)*. 2024, s. 1–6. Dostupné z DOI: 10.1109/ICETCS61022.2024.10543993.

Prílohy