DHBW
Duale Hochschule
Baden-Württemberg

Implementation on a smile tracking AI model

Studienarbeit

Course of Studies Informatik

Duale Hochschule Baden-Württemberg Karlsruhe

Jeremie Bents

Marvin Lindner

Submitted on: 31.01.2025

Student ID, Course: 1941564, TINF22B2
4274538, TINF22B2

Supervisor at DHBW: Prof. Dr. Roland Schätzle

Contents

1

Introduction

1.1 Motivation

The primary motivation behind this study is to develop a robust software system capable of automatically detecting smiling and laughing faces. The project is inspired by the popular series Last One Laughing (LOL), where the ability to monitor and analyze face expressions in real-time can significantly enhance the viewing experience. This study aims to leverage advanced face recognition and emotion detection technologies to create a service that can accurately identify and respond to face expressions. The literature review will cover existing methods and technologies in face recognition and emotion detection, highlighting the advancements and challenges in the field. The scope of this study includes the development, implementation, and evaluation of the proposed system.

1.2 Problem

The current state of face recognition and emotion detection technologies presents several challenges and limitations. While there have been significant advancements, existing systems often struggle with accuracy and real-time performance, especially in dynamic environments. The primary problem addressed in this study is the need for a reliable and efficient system that can detect smiling and laughing faces in real-time with high accuracy. The gap identified is the lack of a specialized solution tailored for the specific requirements of the LOL series, including real-time monitoring, immediate response, and high reliability. This

study aims to bridge this gap by developing a system that meets these specific needs.

1.3 Research Framework and Objectives

The importance of this research lies in its potential to enhance the viewing experience of the LOL series by providing a reliable and efficient face expression detection system. The research problem is to develop a system that can accurately detect smiling and laughing faces in real-time. The research aims and objectives include:

- Developing a face recognition module to detect faces in camera streams.
- Implementing an emotion detection module to identify smiling and laughing expressions.
- Ensuring real-time performance and immediate response to detected expressions.
- Evaluating the system's accuracy and reliability through comprehensive testing.

The hypotheses of this study are:

1. The proposed system will accurately detect smiling and laughing faces in real-time.
2. The system will provide immediate responses to detected expressions with minimal latency.

The methodology includes the development of the system using Python, leveraging existing libraries and frameworks for face recognition and emotion detection. The study will also involve extensive testing and evaluation to ensure the system meets the defined requirements.

1.4 Thesis Structure

The order of information in this thesis will follow a structured approach:

1. **Introduction**: Provides the motivation, problem statement, and research objectives.
2. **Literature Review**: Reviews existing methods and technologies in face recognition and emotion detection.
3. **Methodology**: Details the development process, tools, and techniques used.
4. **Implementation**: Describes the implementation of the face recognition and emotion detection modules.

5. **Testing and Evaluation**: Presents the testing procedures, results, and evaluation of the system's performance.
6. **Conclusion**: Summarizes the findings, discusses the implications, and suggests future work.

> **TODO**
>
> Diese Struktur ist sehr generisch und im Verlaufe der Dokumentation unserer Ergebnisse und der Entwicklung des Systems werden wir die Struktur anpassen müssen.

2

Fundamentals

- Fundamentals of face Recognition and Emotion Detection
- Psychological Theories and Mechanisms of Smiling and Laughter
- Overview of Machine Learning and Deep Learning Techniques in Visual Analysis

TODO

Marvin: planned structure:
- Theoretically we need to explain what human emotions are (emphasize on smiling)
- Machine Learning | Done
- Neural Networks | Done
  - Convolutional Neural Networks | Done
    - Feature Based
    - Image Based

Image Processing
- Technologies

further topics that need to be covered:
- Haar Feature

2.1 History of Face Recognition

Face recognition technology started in the 1960s, when early computer programs could analyze face features by measuring the distances between key points, such as the eyes and nose. A major milestone came in the 1990s with the intro-

duction of the “eigenfaces” approach, which used mathematical algorithms to identify unique face characteristics automatically. This made face recognition significantly more accurate and practical, opening doors for its use in security and surveillance. By the early 2000s, it was implemented in airports and other high-security areas, particularly for public safety.

Today, face recognition has become a common feature worldwide. In countries like China, it’s used extensively in public spaces for surveillance, shopping, and even as a payment method. In the United States, law enforcement agencies use it to identify suspects, while smartphones like iPhones and many Android devices offer face unlock features for secure access. Many airports globally also use face recognition to streamline passport control.

Public reactions to face recognition are mixed. On the one hand, people appreciate its convenience for personal use and the added layer of security it provides in certain settings. However, there are significant privacy concerns. In the European Union, strict data privacy laws have slowed down its adoption for surveillance, and in the United States, several cities and states have passed legislation limiting its use. Many worry about data security, the potential for mass surveillance, and possible biases that could lead to misidentification. This ongoing debate has led to calls for more regulation to ensure that face recognition technology is used responsibly and ethically.

[1], [2], [3]

Certainly! Here is the revised text with a focus on smooth reading and scientific understanding:

—

2.2 Emotion Recognition

Emotions play a significant role in this thesis and need to be described in detail to understand their importance and the technology used to recognize them.

2.2.1 Introduction to Emotions

They are complex psychological states characterized by physiological changes, cognitive processes, and behavioral responses. They play a pivotal role in our daily lives, influencing communication, decision-making, and social inter-

actions. Emotions are expressed through a variety of channels, including facial expressions, body language, vocal tone, and other nonverbal signals. Scientific research has identified six universal emotions: joy, sadness, fear, anger, surprise, and disgust [4]. These emotions are expressed and recognized similarly across all cultures, highlighting their fundamental importance to human experience.

2.2.2 Application of Emotion Recognition

The automotive industry is leveraging Facial Expression Recognition (FER) technology to enhance driver safety and comfort. Emotion recognition systems in vehicles can detect a driver's mood and alertness, thereby adjusting safety mechanisms accordingly. For instance, these systems can alert drivers if they appear drowsy or distracted, potentially preventing accidents.

Companies are also utilizing emotion detection tools to analyze consumer sentiments about products. For example, virtual try-on apps can benefit from emotion recognition by capturing a user's response to different makeup looks or hairstyles. This feedback is invaluable for brands, as it allows them to recommend products that customers are likely to enjoy and purchase, thus improving customer satisfaction and sales.

In the field of market research, emotion detection tools can be employed for real-time analysis of viewer reactions to advertisements. This enables advertisers to optimize content for emotional engagement and tailor advertising strategies to the viewer's emotional state. By understanding and leveraging emotional responses, advertisers can create more impactful and effective marketing campaigns. [5].

2.3 Machine Learning

Machine Learning (ML) is a subfield of Artificial Intelligence (AI) focused on enabling computers to learn from data and make predictions without explicit programming. The term “machine learning” was coined by Arthur Samuel in the 1950s, who developed algorithms that allowed computers to improve performance through experience.

Neural networks, a key component of ML, are computational models inspired by the human brain. They consist of layers of interconnected neurons that process and learn from data. Early breakthroughs, including backpropagation, by researchers like Geoffrey Hinton and Yann LeCun, allowed neural networks to effectively handle tasks like pattern recognition.

Today, ML is widely applied across various fields, including supervised learning for tasks like medical diagnosis and unsupervised learning for clustering and recommendation systems. It is also central to advancements in natural language processing, image and speech recognition, and autonomous systems such as self-driving cars. The continued growth of data and computational power drives the increasing integration of ML into modern technologies. [6]

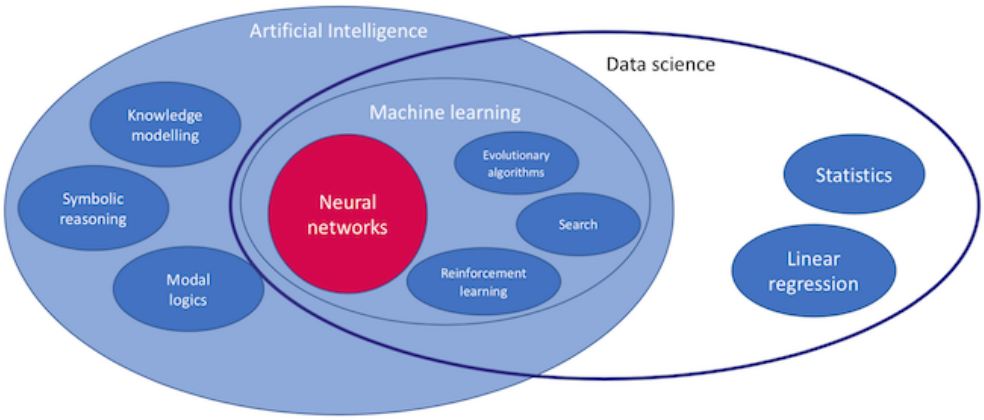

Figure 1 — Machine Learning Field

This image was found in [7].

2.4 Neural Networks

A Neural Network (NN) is a general framework of interconnected nodes, or "neurons," organized in layers. Typically, it consists of an input layer, one or more hidden layers, and an output layer. In standard neural networks, each neuron in one layer connects to every neuron in the next, creating what is called a "fully connected" structure. This setup is suitable for tasks like classification, regression, and prediction. However, it is less efficient when handling large image data, as it does not inherently recognize spatial patterns, such as edges or textures, within images.

That is for a Convolutional Nerual Network (CNN). It is a specialized type of neural network designed to process image and video data effectively. CNNs use **convolutional layers** to apply filters to small sections of the image, allowing the network to extract features like edges, shapes, and textures. This convolutional process preserves the spatial relationships within the image, making CNNs exceptionally effective for tasks like image recognition and classification.

The unique structure of CNNs often includes:
- **Convolutional Layers** that extract essential features from the image,
- **Pooling Layers** that reduce data volume while emphasizing significant features, and
- **Fully Connected Layers** that integrate these features and produce the final classification.

This architecture makes CNNs more efficient than standard neural networks by reducing the number of parameters and focusing on image-specific patterns, allowing CNNs to process visual data with a high degree of accuracy.

[8], [9]

3

Methodology

- Data Collection and Preprocessing Techniques (e.g., dataset selection, annotation strategies, data sanitization)
- Selection of Model Architecture and Algorithms (e.g., Convolutional Neural Networks, Recurrent Neural Networks)
- Model Training Protocols and Performance Metrics for Evaluation

Review of Current Research and Technologies

- Systematic Overview of Contemporary Methods for face Expression Recognition
- Comparative Analysis of Existing Emotion Recognition Systems
- Identification of Gaps, Challenges, and Limitations in Current Technologies

3.1 Wie hilf Tensorflow uns bei der Entwicklung

3.2 Wozu dient OpenCV

4

Implementation of the Model

5

Implementation of the Website

6

Experimental Setup and Results

- Design of Experiments and Test Environment
- Quantitative Evaluation of System Accuracy (e.g., accuracy, precision, recall, F1 score)
- Comprehensive Analysis and Interpretation of Experimental Findings

7

Conclusion

A References

[1] Thales, “Was ist ein Convolutional Neural Network?.” Accessed: Nov. 14, 2024. [Online]. Available: https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/inspired/where-face-recognition-used

[2] N. Braun Binder, E. Kunz, and L. Obrecht, “Maschinelle Gesichsterkennung im öffentlichen Raum.” Accessed: Nov. 14, 2024. [Online]. Available: https://edoc.unibas.ch/88208/

[3] R. Campillo, “face Recognition History.” Accessed: Nov. 14, 2024. [Online]. Available: https://www.mobbeel.com/en/blog/face-recognition-history/#:~:text=The%20initial%20steps%20in%20face,RAND%20tablets%20in%20the%201960s.

[4] P. Ekman, “Facial expression and emotion,” 1993. doi: 10.1037/0003-066X.48.4.384.

[5] D. Mekinec, “face Recognition History.” Accessed: Nov. 14, 2024. [Online]. Available: https://visagetechnologies.com/face-emotion-recognition-guide/

[6] S. J. Bigelow, “Machine learning vs. neural networks: What`s the difference?.” Accessed: Nov. 14, 2024. [Online]. Available: https://www.techtarget.com/searchenterpriseai/answer/Machine-learning-vs-neural-networks-Whats-the-difference

[7] S. Van Otterloo, “AI, Machine Learning and neural networks explained.” Accessed: Nov. 14, 2024. [Online]. Available: https://ictinstitute.nl/ai-machine-learning-and-neural-networks-explained/

[8] The MathWorks Inc., “What is a neural network?.” Accessed: Nov. 14, 2024. [Online]. Available: https://de.mathworks.com/discovery/convolutional-neural-network.html#:~:text=Ein%20Convolutional%20Neural%20Network%20(CNN,Klassen%20und%20Kategorien%20zu%20erkennen.

[9] IBM, “Was ist ein Convolutional Neural Network?.” Accessed: Nov. 14, 2024. [Online]. Available: https://www.ibm.com/topics/neural-networks

B Acronyms

AI	Artificial Intelligence
API	Application Programming Interface
CNN	Convolutional Nerual Network
FER	Facial Expression Recognition
HTTP	Hypertext Transfer Protocol
LOL	Last One Laughing
ML	Machine Learning
NN	Neural Network
REST	Representational State Transfer

C Glossary

Komponente Ein Architekturbaustein. Zusammengesetzte Komponenten bestehen aus weiteren Subkomponenten. Einfache Komponenten sind nicht weiter unterteilt.

Softwareschnittstelle Ein logischer Berührungspunkt in einem Softwaresystem: Sie ermöglicht und regelt den Austausch von Kommandos und Daten zwischen verschiedenen Prozessen und Komponenten.

Declaration of Authorship

Gemäß Ziffer 1.1.13 der Anlage 1 zu §§ 3, 4 und 5 der Studien- und Prüfungsordnung für die Bachelorstudiengänge im Studienbereich Technik der Dualen Hochschule Baden- Württemberg vom 29.09.2017. Wir versichern hiermit, dass wir unsere Arbeit mit dem Thema:

Implementation on a smile tracking AI model

selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt haben. Wir versichern zudem, dass alle eingereichten Fassungen übereinstimmen.

Karlsruhe, 31.01.2025

Jeremie Bents

Marvin Lindner