Araştırma Makalesi

YAPAY ZEKA KULLANARAK YÜZDE DUYGU TANIMA

Saygı Kaan TUZCU[†], Mustafa Cem KASAPBAŞI[††]
[†] İstanbul Ticaret Üniversitesi, Siber Güvenlik, İstanbul, Türkiye
[††] İstanbul Ticaret Üniversitesi, Bilgisayar Mühendisliği, İstanbul, Türkiye
saygikaantuzcu@gmail.com , mckasapbasi@ticaret.edu.tr
0009-0003-4718-1208, 0000-0001-6444-6659

Atıf/Citation: TUZCU, S. K., KASAPBAŞI, M. C., (2025). Yapay Zeka Kullanarak Yüzde Duygu Tanıma, Journal of Technology and Applied Sciences 8(1) s.11-24, DOI: 10.56809/icujtas.1518225

ÖZET

Yapay Zeka, genellikle insan zekası gerektiren görevleri yerine getirebilen bilgisayar sistemlerinin ve algoritmaların geliştirilmesini ifade eder. Bu görevler arasında problem çözme, öğrenme, doğal dili anlama, kalıpları tanıma ve karar verme yer alır. Yapay zeka sistemleri, insanın bilişsel işlevlerini taklit edecek veya simüle edecek şekilde tasarlanmış olup, onların büyük miktarlarda veriyi işlemesine ve analiz etmesine, değişen koşullara uyum sağlamasına ve geleneksel olarak insanlara özel olan görevleri yerine getirmesine olanak tanır. Bu çalışmada, yüzleri tanımak ve duyguları sınıflandırmak için yapay zeka teknolojileri, özellikle de derin öğrenme, önceden eğitilmiş modeller kullanılmıştır. AlexNet, VGGNet, ResNet, Inception gibi popüler mimariler farklı veri setleri üzerinde eğitilerek performansları karşılaştırılmıştır. Elde edilen sonuçlar, derin evrişimli sinir ağları ile görüntü sınıflandırma performansında önemli bir iyileşme olduğunu göstermektedir. Özellikle daha derin ve karmaşık mimariye sahip ağlar daha iyi performans gösterme eğilimindedir. Ancak aşırı uyum riskini azaltmak için uygun düzenleme tekniklerinin kullanılması çok önemlidir. Çalışma Python yazılım dili ve sayısal python dilinin kısaltması olan NumPy kitaplığı kullanarak bilimsel hesaplamalar yapılmıştır. Programda çeşitli videolar kullanılarak veri seti oluşturulmuştur. Bu veri setlerinin en doğru çıktısı CNN algoritmasında performans metriklerinde 1 olarak yansımıştır.

Anahtar Kelimeler: AI, Face Detection, Haar Cascade, k-NN, YOLO, SSD, CNN, Yapay Zeka, Görüntü İşleme, Duygu Tanıma

FACIAL EMOTION RECOGNITION USING ARTIFICIAL INTELLIGENCE

ABSTRACT

Artificial Intelligence, refers to the development of computer systems and algorithms that can perform tasks that typically require human intelligence. These tasks include problem-solving, learning, understanding natural language, recognizing patterns, and making decisions. AI systems are designed to mimic or simulate human cognitive functions, enabling them to process and analyze large amounts of data, adapt to changing circumstances, and perform tasks that were traditionally exclusive to humans. In this study AI technologies specifically deep learning pre trained models are utilized to recognize faces and classify emotions. Popular architectures such as AlexNet, VGGNet, ResNet, and Inception have been trained on different datasets, and their performances are compared. The results obtained indicate a significant improvement in image classification performance with deep convolutional neural networks. Particularly, networks with deeper and more complex architectures tend to perform better. However, it is crucial to employ appropriate regularization techniques to mitigate the risk of overfitting. Scientific calculations were made using the Python programming language and the NumPy library, which is short for the numerical Python language. A data set was created using various videos in the program. The most accurate output of these data sets is reflected as 1 in the performance metrics of the CNN algorithm.

Keywords: Artificial Intelligence, AI, Image Processing, Face Detection, Emotion Recognition, Haar Cascade, k-NN, YOLO, SSD, CNN

Geliş/Received	:	**17.07.2024**
Gözden Geçirme/Revised	:	**23.08.2024**
Kabul/Accepted	:	**23.08.2024**

1. INTRODUCTION

Artificial Intelligence (AI) technologies rely on a variety of techniques, including machine learning, deep learning, natural language processing, computer vision, and reinforcement learning. Machine learning algorithms enable AI systems to learn from data and improve their performance over time. Deep learning, a subset of machine learning, involves artificial neural networks and has been particularly successful in tasks like image and speech recognition. AI finds applications in numerous fields, including healthcare, finance, autonomous vehicles, customer service, manufacturing, and more. Some common AI applications include virtual personal assistants (e.g., Siri, Alexa), recommendation systems (e.g., Netflix's movie recommendations), and autonomous vehicles (self-driving cars). AI is also used for solving complex problems in scientific research, optimizing business processes, and improving decision-making.

As AI continues to advance, it raises important ethical and societal questions, such as issues related to bias in algorithms, data privacy, automation's impact on jobs, and the responsible development and deployment of AI technologies. Researchers and policymakers are working to address these challenges while harnessing the potential benefits of AI for the betterment of society.

In this article, we will focus on facial emotion recognition. Facial emotion recognition, a vital component of affective computing, has witnessed substantial progress through the integration of artificial intelligence (AI) techniques. This abstract presents a comprehensive overview of AI-based facial emotion recognition, exploring the latest advancements, methodologies, and practical applications.

Emotions, a fundamental aspect of human communication, play a crucial role in various domains, from human-computer interaction to mental health assessment. These studies underline the significance of accurate emotion recognition and its impact on improving user experiences, healthcare, and marketing strategies.

The study outlines the core components of AI-based facial emotion recognition, focusing on the utilization of computer vision techniques, deep learning models, and neural networks. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformer architectures have shown remarkable success in extracting emotional cues from facial images, allowing for real-time emotion detection.

Various data sources, such as images, videos, and 3D facial scans, are discussed in the context of emotion recognition. The challenges associated with dealing with diverse data modalities and the importance of multimodal data fusion for robust emotion analysis are addressed in this study.

The study also explores the applications of AI-based facial emotion recognition, encompassing domains like mental health diagnostics, human-robot interaction, user experience enhancement, and sentiment analysis in social media and marketing. The potential for real-time emotion-aware systems to revolutionize industries and services is highlighted.

Ethical considerations are addressed, including concerns related to privacy, data security, and potential biases in AI models. Transparency, fairness, and interpretability in emotion recognition systems are identified as important focal points for future research and development.

2. THE INFORMATION AREA

The Information Era, also known as the Information Age, is a period in human history characterized by the widespread use of information and communication technologies (ICT), particularly computers and the internet, to create process, store, and exchange information. It represents a significant shift from the Industrial Age, which was characterized by mass production and manufacturing. It is characterized by a rapid shift from traditional industries, as established during the Industrial Revolution, to an economy centered on information technology [1].

Key features of the Information Era include:

2.1. Digital Technology

The Information Era is marked by the proliferation of digital technology, such as computers, smartphones, and the internet. These technologies have transformed the way we access and share information, conduct business, and interact with one another.

2.2. Information Accessibility

The era has seen a democratization of information, with people having easier access to vast amounts of data and knowledge. This has led to greater empowerment and opportunities for individuals and organizations.

2.3. Global Connectivity

The internet and other communication technologies have connected people and organizations across the globe. This interconnectedness has facilitated global trade, collaboration, and the exchange of ideas.

2.4. Knowledge Economy

In the Information Era, knowledge and information are valuable resources. Many economies have shifted from manufacturing and agriculture to services and information-based industries. Intellectual property and data have become central to economic growth.

2.5. Rapid Innovation

The pace of technological innovation has accelerated during the Information Era. New technologies, software, and digital platforms are continually developed and updated, leading to constant change and adaptation in various fields.

2.6. Data Analytics

The era has seen the rise of data analytics and big data, with organizations and governments using data to make informed decisions, improve efficiency, and gain insights into various aspects of society.

2.7. Social Media and Online Communities

The Information Era has given rise to social media platforms and online communities, enabling people to connect, share information, and engage in discussions on a global scale.

2.8. Digital Privacy and Security

With the increased reliance on digital technology, concerns about data privacy and cybersecurity have become prominent issues. Protecting personal and sensitive information has become a priority.

The Information Era has had a profound impact on almost every aspect of society, including communication, education, healthcare, business, and entertainment. It has transformed the way we work, learn, and interact with one another, and it continues to shape the world in new and evolving ways.

3. FACIAL EMOTION RECOGNITION

Facial emotion recognition is a technology that involves the identification and analysis of human emotions based on facial expressions. It's an area of research and application in computer vision and artificial intelligence. Here's an overview of the key aspects:

3.1. Facial Expressions and Emotions

Facial recognition technology uses machine learning and computer vision algorithms to analyze facial features and expressions. It can be used to detect emotional states by identifying patterns of muscle movement and changes in facial expressions associated with specific emotions.

3.1.1. Basic Emotions

Facial emotion recognition often revolves around the identification of basic emotions, such as happiness, sadness, anger, surprise, fear, and disgust.

3.1.2.Facial Action Coding System (FACS)

FACS is a comprehensive system that describes facial expressions in terms of anatomica9l muscle movements.

3.2. Data Collection

3.2.1.Datasets

Researchers and developers use datasets containing images or videos with labeled facial expressions to train and evaluate facial emotion recognition models. Examples include CK+, AffectNet, and FER2013.

3.2.2. Annotations

The dataset should have annotations indicating the emotion expressed in each image or video frame.

3.3. Feature Extraction

Feature extraction techniques can be used to identify key facial landmarks, such as eye, mouth, and eyebrow positions, and analyze their changes to determine emotions. These features can be extracted using image processing algorithms and then fed into machine learning models for emotion recognition.

3.3.1.Facial Landmarks

Key points on the face, such as the eyes, nose, and mouth, are identified to extract features.

3.3.2.Texture and Color Analysis

Textural patterns and color variations in facial regions are analyzed for emotion recognition [2].

3.4. Machine Learning Models

3.4.1.Traditional Approaches

Early approaches often involved manually crafted feature extraction and traditional machine learning techniques like Support Vector Machines (SVM) or Decision Trees.

3.4.2.Deep Learning

With the advent of deep learning, Convolutional Neural Networks (CNNs) and recurrent neural networks (RNNs) have become popular for facial emotion recognition due to their ability to automatically learn hierarchical features. Deep learning models, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have been effective in facial emotion recognition. These models can be trained to recognize emotional states by processing raw data and identifying emotional patterns.

3.5. Training Model

3.5.1.Supervised Learning

Models are trained on labeled data, learning to map facial features to corresponding emotions.

3.5.2.Data Augmentation

Techniques like image rotation, scaling, and flipping can be applied to artificially increase the size of the training dataset.

3.6. Applications

3.6.1. Human-Computer Interaction

Emotion recognition is used in human-computer interaction to create more intuitive and responsive interfaces.

3.6.2. Healthcare

It has potential applications in mental health monitoring and diagnosis.

3.6.3. Marketing and User Experience

Companies use emotion recognition to analyze customer reactions to products or advertisements.

3.7. Ethical Considerations

3.7.1. Privacy

The use of facial emotion recognition raises concerns about privacy, and there are debates about its ethical implications.

3.8. Future Trends

3.8.1. Multimodal Approaches

Combining facial emotion recognition with other modalities like voice and gesture for more accurate emotion analysis.

3.8.2. Explainable AI

Efforts to make facial emotion recognition models more interpretable and explainable.

Facial emotion recognition is a fascinating and evolving field with numerous potential applications. Advances in deep learning and increasing access to large, diverse datasets continue to drive progress in this area. Keep in mind that it's important to consider ethical considerations, such as privacy and potential biases in the development and deployment of facial emotion recognition technology.

The choice of method depends on the specific application, accuracy requirements, and available resources. It's important to note that while these methods can be useful for inferring emotional states, they may not always provide a perfect assessment, as emotional expression can vary widely between individuals and cultures. Additionally, ethical and privacy considerations should be taken into account when implementing facial analysis for emotion recognition, as it can raise concerns about data privacy and surveillance.

4. CLASSIFICATION OF FACE DETECTION

4.1. K-Nearest Neighbors Algorithm (k-NN)

In statistics, the k-nearest neighbors algorithm (k-NN) is a non-parametric supervised learning method first developed by Evelyn Fix and Joseph Hodges in 1951 [3]. The k-NN algorithm is a non-parametric, or an instance-based, or a lazy method, and has been regarded as one of the simplest methods in data mining and machine learning. The principle of k-NN algorithm is that the most similar samples belonging to the same class have high probability. Generally, the k-NN algorithm first finds k nearest neighbors of a query in training data set, and then predicts the query with the major class in the k nearest neighbors. Therefore, it has recently been selected as one of top 10 algorithms in data mining [4].

There is two main concepts that k-NN algorithm is used. First one is classification and second one is regression. The k-NN classification algorithm first selects k closest samples (i.e., k nearest neighbors) for a test sample from all the training samples, and then predicts the test sample with a simple classifier, e.g., majority classification rule.

The k-NN regression has been widely used and studied for many years in pattern recognition and data mining [6]. In an intuitive manner, it approximates the association between independent variables.

Mathematical description of k-NN classification is as follows: in the classification or discrimination problem with two populations, denoted by X and Y, one wishes to classify an observation z to either X or Y using only training data. The kth-nearest neighbor classification rule is arguably the simplest and most intuitively appealing nonparametric classifier. It assigns z to population X if at least ½ k of the k values in the pooled training-data set nearest to z are from X, and to population Y otherwise [5].

First thing comes to mind when algorithm is understood is choice of k and implementation of the algorithm. For more accuracy, one such algorithm uses a weighted average of the k-NN, weighted by the inverse of their distance. That way closer neighbors of k neighbors are more dominant when classification is done. There are some other approaches as weighting dimensions etc.

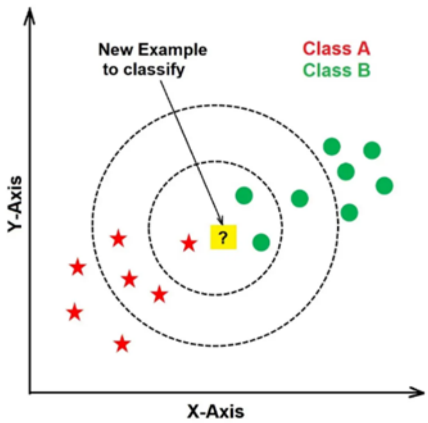

Figure 1. Choosing Different K Values Leads to Different Results.

Real challenge is choosing k value. Choosing a very low value will most likely lead to inaccurate predictions and also larger values of k reduces effect of the noise on the classification but make boundaries between classes less distinct.

In fact, T. M. Cover and P. E. Hart stick up for single-NN in “Nearest Neighbor Pattern Classification”. Their calculations result that single-NN has lower probability of error. Their article concludes as following: The single NN rule has been shown to be admissible among the class of k-NN rules for the n-sample case for any n. It has been shown that the NN probability of error R, in the M-category classification problem, is bounded below by the Bayes probability of error R* and above by R*(2 - MR*/ (M - 1)). Thus, any other decision rule based on the infinite data set can cut the probability of error by at most one half. In this sense, half of the available information in an infinite collection of classified samples is contained in the nearest neighbor [4]. Also, they add up that if the number of samples is large it makes good sense to use, instead of the single nearest neighbor, the majority vote of the nearest k neighbors. K is wished to be large in order to minimize the probability of a non-Bayes decision for the unclassified point x, but we wish k to be small (in proportion to the number of samples) in order that the points be close enough to x to give an accurate estimate of the posterior probabilities of the true class of x [4].

Another method called S-kNN d is a data-driven method for selecting the optimal k values. In this method a matrix named W is defined and each element w ij of the matrix W can be understood as the correlation between ith test sample and jth training sample. Then with an algorithm, an equation is optimized to obtain the correlation coefficient matrix W, so as to obtain an optimal k value for each test sample. Consequently, the most correlative training samples of a test sample is generated. This means that different test samples are predicted with different numbers of nearest neighbors [6]. As seen in Figure 2 it is just one of the methods among others for choosing optimal k value.

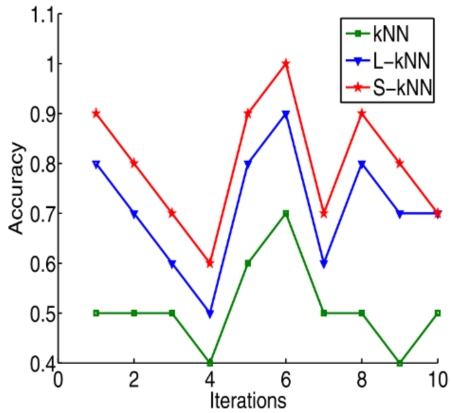

Figure 2. S-kNN has more accuracy the normal k-NN.

In conclusion k-NN is a simple-implemented learning method and no training is required before classifications. It can be cost-intensive and a lot of memory can be required while working with large data sets. As described above choosing right value of K can be tricky.

4.2. Single Shot Detectors (SSD)

Single Shot Detector (SSD) is a popular deep learning model for object detection. SSD is a fast and efficient model that can detect multiple objects in a single forward pass. Here are the key features and operating principles of SSD:

4.2.1.Model Structure

SSD starts with a pretrained base network (typically VGG (Visual Geometry Group) or ResNet) and uses multi-resolution feature maps to perform object detection. The model begins with the base network and then utilizes various feature maps at different scales for object detection.

4.2.2.Multi-Resolution Detection

SSD uses multi-resolution feature maps to detect objects of different sizes. This allows it to adapt to various object sizes and positions, providing high accuracy in detection.

4.2.3.Anchors

SSD employs reference boxes (anchors) of different sizes and aspect ratios. These anchors serve as initial points for detecting objects in the image.

4.2.4.Classification and Localization

The model performs classification (determining the object class) and localization (determining the precise boundaries of the object) for each anchor box.

4.2.5.Loss Calculation

SSD uses a loss function that balances classification and localization losses. This helps the model achieve accuracy in both object classification and localization.

4.2.6.Speed and Efficiency

SSD is fast because it performs object detection in a single forward pass, making it suitable for real-time applications.

SSD has demonstrated impressive results in various datasets and application areas. It is particularly preferred for projects that require real-time object detection.

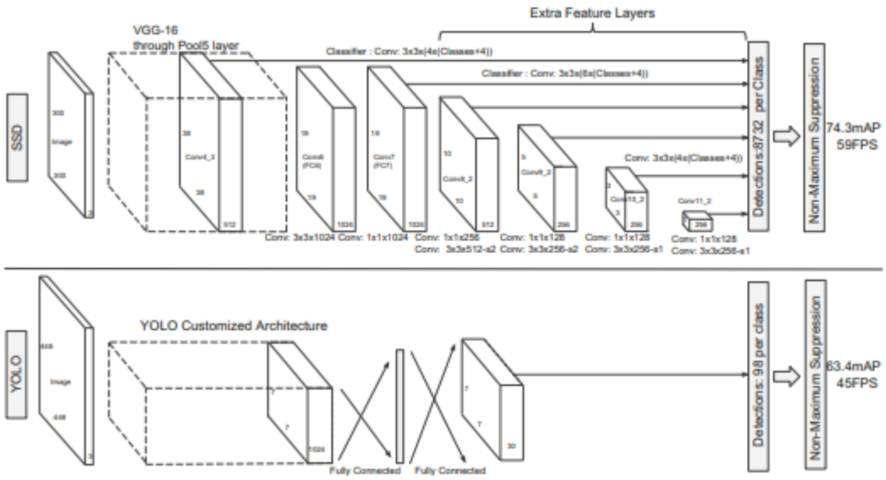

Figure 3. Defines a comparison between two single shot detection models: SSD and YOLO [16].

Figure 4. Confusion matrix of an application written in the SSD algorithm.

4.3. Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) have emerged as a pivotal advancement in artificial intelligence, particularly in computer vision applications. Renowned for their proficiency in image recognition, object detection, and image segmentation, CNNs have garnered significant attention from researchers, engineers, and practitioners alike.

Training CNNs entails utilizing extensive datasets via backpropagation. Through iterative adjustments of parameters (weights and biases), the network minimizes the disparity between predicted and actual outputs, effectively learning discriminative features from raw input data.

CNNs exhibit versatility across diverse domains, including:

- Image classification
- Object detection
- Facial recognition
- Autonomous vehicles

- Medical image analysis
- Natural language processing (when applied to text as images)

4.3.1. An Examination of the Performance of Convolutional Neural Networks (CNNs) in Image Classification

Image classification stands as a significant problem in the fields of computer vision and machine learning. With the advancements in deep learning methods in recent years, Convolutional Neural Networks (CNNs) have played a pivotal role in this domain. This study investigates various architectures of deep convolutional neural networks to evaluate their performance in image classification.

Image classification can be defined as the task of assigning a specific class to an image, such as determining whether an image belongs to a category like a dog, cat, or car. This type of classification problem is widely used in various application domains including automated vehicles, medical imaging systems, and security systems.

This study utilizes various architectures of deep convolutional neural networks to assess the performance of image classification. Popular architectures such as AlexNet, VGGNet, ResNet, and Inception have been trained on different datasets, and their performances are compared. Typically, large datasets like ImageNet are used for the training process.

The results obtained indicate a significant improvement in image classification performance with deep convolutional neural networks. Particularly, networks with deeper and more complex architectures tend to perform better. However, it is crucial to employ appropriate regularization techniques to mitigate the risk of overfitting.

This study provides an examination of the performance of deep convolutional neural networks in image classification. The findings demonstrate the effectiveness of such networks across a wide range of applications and suggest potential avenues for further improvement. [9][10][11][12]

Figure 5. Confusion matrix of an application written in the CNN algorithm.

4.4. Haar Cascade Classifier

This study, completed by Paul Viola and Michael Jones in 2001, is a machine learning approach that can quickly process images with a high detection rate to detect objects in an image [13]. Although this algorithm generally targets object detection, it is designed mostly for face detection.

The Haar Cascade Classifier is a popular method for object detection that uses simple features, known as Haar-like features, instead of directly using pixel values. This approach has several advantages:

4.4.1.Spatial Encoding

Haar-like features effectively encode spatial relationships and local patterns, which are challenging to capture with raw pixel values, especially when training data is limited.

4.4.2.Speed

Using integral images, the computation of Haar-like features is very efficient, making the detection process faster compared to methods relying on pixel values.

The algorithm detects faces in an image by using features like eyes, nose, and mouth, leveraging techniques such as line detection, edge detection, and center detection for precise feature selection and extraction.

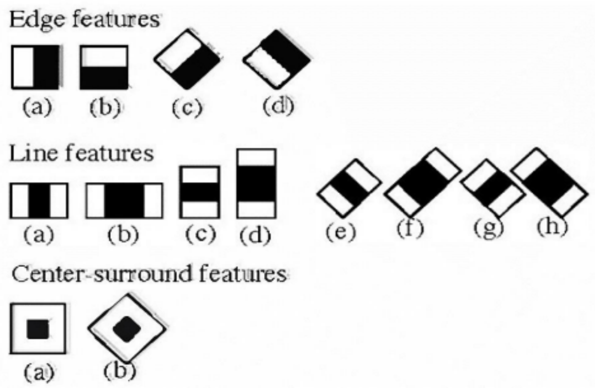

Figure 6. Haar-like features for face detection [14].

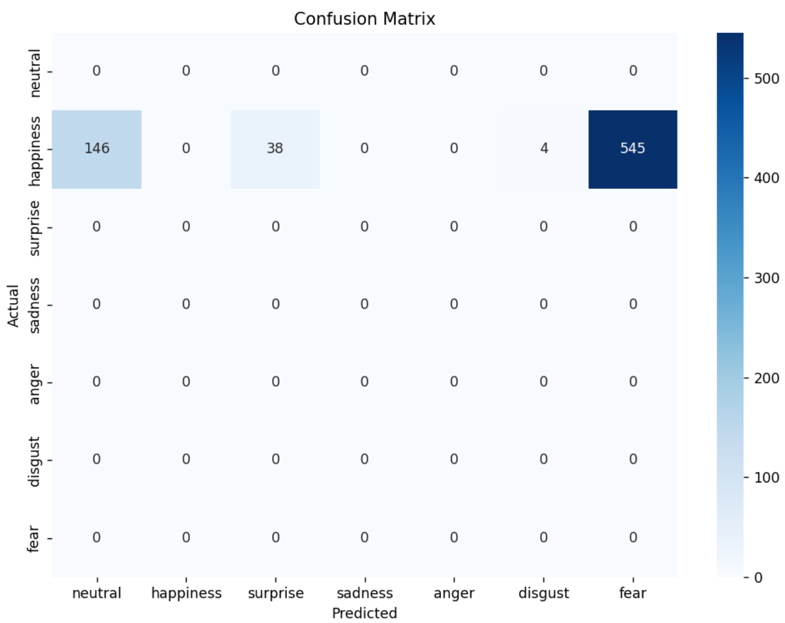

Figure 7. Confusion matrix of an application written in the Haar algorithm.

5. CONCLUSION

Face detection, like most object-detection problems, is a difficult task because of the significant pattern variations that are hard to parameterize analytically. Some common sources of pattern variations are facial appearance, expression, presence or absence of common structural features such as glasses or a moustache, and light-source distribution [15]. So, when it is about emotion recognition things get much more complicated. First of humans express their emotions in a lot of ways. As an example, when someone is angry, they furrow their eyebrows by moving them down and inward, they narrow their eyes, their skin becomes flushed, they tighten lips, they flare their nostrils, etc. In other words, there is a lot of features for recognition of that emotion. In machine learning new feature implies new dimension. That is why for comparison of these two machine learning algorithms for sure their behaviours in dimensions should definitely be taken into consideration.

Face recognition, a crucial aspect of computer vision, has been significantly enhanced by the advent of various algorithms and techniques. Among the popular methods for face recognition are You Only Look Once (YOLO), Single Shot MultiBox Detector (SSD), and k-Nearest Neighbors (k-NN). Each approach offers distinct advantages and limitations, making them suitable for different scenarios.

SSD, another popular object detection framework, offers a balance between speed and accuracy. By utilizing a series of convolutional layers with different aspect ratios, SSD efficiently detects faces at various scales while minimizing computational overhead. Its ability to handle multi-scale objects makes it well-suited for face recognition tasks where faces may appear in different sizes and orientations within an image.

On the other hand, k-Nearest Neighbors (k-NN) is a simple yet effective algorithm for face recognition. By comparing the features of a query face with those of known faces in the dataset, k-NN assigns labels based on the nearest neighbors in feature space. While k-NN may lack the sophistication of deep learning-based approaches like YOLO and SSD, it remains a viable option for small-scale face recognition applications, especially when computational resources are limited.

In the context of facial emotion recognition, a confusion matrix is used to evaluate how well the model identifies different emotions. Suppose we have a model that recognizes six emotions: Happy, Fear, Surprised, Neutral, and Disgusted.

5.1. Understanding the Confusion Matrix

5.1.1.True Positives (Diagonal Elements)

Correctly identified emotions. For example, 50 instances of "Happy" were correctly predicted as "Happy".

5.1.2.False Positives (Off-Diagonal Elements in Columns)

Correctly identified emotions. For example, 50 instances of "Happy" were correctly predicted as "Happy".

5.1.3.True Positives (Diagonal Elements)

Incorrect predictions where one emotion is mistakenly identified as another. For example, 3 instances of "Angry" were incorrectly predicted as "Happy".

5.1.4.False Negatives (Off-Diagonal Elements in Rows)

Missed predictions where an emotion is incorrectly classified as another. For example, 5 instances of "Happy" were incorrectly classified as "Neutral".

Table 1. Confusion matrix of the SSD.

	Happiness	**Neutral**	**Fear**	**Suprise**	**Disgusted**
Happiness	1372	96	0	0	0

Table 2. Confusion matrix of the CNN.

	Happiness	Neutral	Fear	Suprise	Disgusted
Happiness	0	597	0	0	0

Table 3. Confusion matrix of the Haar Cascade.

	Happiness	Neutral	Fear	Suprise	Disgusted
Happiness	0	146	545	38	4

From the confusion matrix, we can calculate various performance metrics for each emotion class:

5.2. Performance Metrics

5.2.1.Accuracy

The overall correctness of the model's predictions.

$$\text{Accuracy} = \frac{\sum \text{True Positives}}{\text{Total Instances}}$$

Formula 1.

5.2.2.Precision for Each Emotion

The ratio of correctly predicted instances of a given emotion to the total predicted instances of that emotion.

$$\text{Precision}_{\text{Happy}} = \frac{TP_{\text{Happy}}}{TP_{\text{Happy}} + FP_{\text{Happy}}}$$

Formula 2.

5.2.3.Recall (Sensitivity) for Each Emotion

The ratio of correctly predicted instances of a given emotion to the actual instances of that emotion.

$$\text{Recall}_{\text{Happy}} = \frac{TP_{\text{Happy}}}{TP_{\text{Happy}} + FN_{\text{Happy}}}$$

Formula 3.

5.2.4.F1-Score for Each Emotion

The harmonic mean of precision and recall, providing a balanced measure.

$$\text{F1-Score}_{\text{Happy}} = 2 \times \frac{\text{Precision}_{\text{Happy}} \times \text{Recall}_{\text{Happy}}}{\text{Precision}_{\text{Happy}} + \text{Recall}_{\text{Happy}}}$$

Formula 4.

Table 4. Comparisons and results of the examined algorithms are as follows.

	Accuracy	Sensitivity	Precision	F1
SSD	0,931432	0,931432	1	0,964499
CNN	1	1	1	1
Haar Cascade	0,74352	0,74352	1	0,852895

In summary, the confusion matrix is a crucial tool in facial emotion recognition, providing detailed insights into model performance, guiding improvements, and ensuring balanced accuracy across different emotion classes.

REFERENCES

1. Manuel, Castells (1996). The information age: economy, society and culture. Oxford: Blackwell. ISBN 978-0631215943. OCLC 43092627.

2. Emotion Recognition based on Texture Analysis of Facial Expression (Authors: Gyanendra K. Verma, Bhupesh Kumar Singh) (Date: November 2011)

3. Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties (Authors: Evelyn Fix and Joseph L. Hodges) (Date: 1951)

4. Nearest Neighbor Pattern Classification (Authors: T. M. Cover and P. E. Hart) (Date: January 1967) (Publisher: IEEE)

5. Choice Of Neighbor Order In Nearest-Neighbor Classification (Authors: Peter Hall, Byeong U. Park and Richard J. Samworth) (Date: 2008) (Publisher: The Annals of Statistics)

6. A novel kNN algorithm with data-driven k parameter computation (Authors: Shichao Zhanga , Debo Cheng, , Zhenyun Denga, Ming Zongc and Xuelian Deng) (Date: July 2018) (Pattern Recognition Letters)

7. A training algorithm for optimal margin classifiers (Authors: Bernhard E. Boser, Isabelle M. Guyon and Vladimir N. Vapnik) (Date: 1992)

8. Support Vector Clustering (Authors: Asa Ben-Hur, David Horn, Hava T. Siegelmann and Vladimir Vapnik) (Date: June 2008) (Journal of Machine Learning Research)

9. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

10. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556

11. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778)

12. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826)

13. Viola, P., And Jones, M., 'Rapid Object Detection Using A Boosted Cascade Of Simple Features,'Accepted Conference On Computer Vısıon And Pattern Recognıtıon 2001

14. Dwivedi, D. (2018). Face detection for beginners. Retrieved from https://towardsdatascience.com

15. Training Invariant Support Vector Machines (Authors: Dennis Decoste and Bernhard Schölkopf) (Date: January 2002) (Machine Learning) (Kluwer Academic Publishers)

16. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified,real-time object detection. In: CVPR (2016)