Computer Vision of Smile Detection Based on Machine and Deep Learning Approach

Huda Lafta Majeed[1], Oday Ali Hassen[2,*], Dhyeauldeen A. Farhan[3], Yu Yu Gromov[4], Kavita Sheoran[5], Geetika Dhand[5]

[1]Computer Science and Information Technology, University of Wasit, Al Kut 52001, Iraq

[2]Ministry of Education, Wasit Education Directorate, Kut 52001, Iraq.

[2]Computer Department, College of Education for Pure Sciences, Wasit University, Iraq

[3]Department of Computer Science, University of Al Maarif, Al-Anbar, 31001, Iraq

[4]Director of the Institute of Automation and information technology, Tambov State technical university, Rusia

[5]Maharaja Surajmal Institute of Technology, New Delhi, India

Emails: hulafta@uowasit.edu.iq; odayali@uowasit.edu.iq; dheyauldeen.farhan@uoa.edu.iq; gromovtambov@yandex.ru; kavita.sheoran@gmail.com; geetika.dhand@gmail.com

Abstract

Smile detection and recognition have been a key component of sentiment analysis, social robotics, human-computer interaction, and mental health monitoring before the advent of deep learning. Understanding and accurately identifying smiles can provide deep insights into human behavior, strengthen communication systems, and enhance adaptive responses in AI interfaces. This paper is a comprehensive review of algorithms developed for smile detection and recognition, and categorizes their main approaches into three traditional computer vision techniques: feature-based, machine learning-based, and deep learning-based. These techniques rely on handcrafted features such as edges, geometric features of the face, and texture, which give interpretability and limited adaptability. This paper explores feature extraction methods such as geometric and histogram-based features (e.g., histograms of directed gradients). In addition, this paper evaluates the effectiveness of traditional classifiers, including support vector machines that use machine learning-based methods, leveraging algorithms such as support vector machines (SVMs), extracted features to classify smiles with improved accuracy. Deep learning techniques, especially convolutional neural networks (CNNs) and hybrid methods provide end-to-end learning capabilities, extracting features directly from raw pixel data and enabling real-time performance. These frameworks, including recurrent neural networks (RNNs) for temporal analysis, generative adversarial networks (GANs) for data augmentation, and graph neural networks (GNNs) for structural analysis, have also pushed the boundaries of smile detection in dynamic and challenging environments. It also aims to provide a comprehensive overview of these classical methods, and analyze their strengths, limitations, drawbacks, and performance across diverse datasets of the proposed databases by focusing on describing these datasets and researchers' methods of working on them as benchmarks for their research, and highlighting their importance in the environments and their contributions to the development of smile detection algorithms in the field of computer vision. Among these datasets are datasets such as CK+, FER2013, AffectNet, and Jaffe in developing, training, and evaluating smile detection and recognition algorithm models. By comparing these methodologies, our paper recommends directing future research towards more efficient, robust, and scalable solutions for smile detection and recognition in diverse applications.

Keywords: Smile detection algorithm; Facial expressions; Deep and machine learning of facial expressions; Computer vision real-time smile detection

1. Introduction

Facial expression analysis involves the process of detecting and recognizing smiles. Identifying and interpreting smiles from visual data such as images or videos plays a critical role in various fields, including human-computer

DOI: https://doi.org/10.54216/JCIM.160115

Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

interaction, recognition and detection, where it enhances user experience through emotion-aware systems; healthcare in monitoring mental health and treatment progress; and marketing by helping to analyze consumer engagement and sentiment. As a universal indicator of positive emotions, smiles are essential for understanding human behavior and building empathetic and responsive systems. Computer vision is a branch of artificial intelligence that aims to enable computers to understand and analyze images and videos in a way similar to humans. One common application of computer vision is face detection and expression analysis, including smile detection. This type of application is used in various fields such as security, marketing, human-machine interaction, and even mental health applications. Smile detection is a classification task that determines whether a person in an image is smiling or not. This can be achieved using machine learning and deep learning techniques. In machine learning, algorithms such as SVM (Support Vector Machine) or Random Forests are used. This is done by manually extracting features from images, such as facial contours, mouth shape, or other features associated with a smile. The model is then trained on these features to determine whether a smile is present or not. [1]

The use of machine learning (ML) approaches involves training pattern recognition algorithms using handcrafted features extracted from images. These include random forests and k-nearest neighbors (KNNs). This is why they are used in future research: due to their superior feature optimization and refinement, they enable better smile detection and classification with additional training data. Consequently, this technique significantly improves performance compared to traditional methods. [2]. in deep learning, deep neural networks (DNNs) and convolutional neural networks (RNNs) are used to automatically detect smiles without the need for manual feature extraction. The most popular models used in this field are convolutional neural networks (CNNs), which are very effective in image processing. These networks are trained on large datasets containing images of smiling and non-smiling faces, where the network automatically learns the distinctive features of a smile. Deep learning (DL) is an approach to automatically extracting features and performing classification. This is why they are used in future research due to their high accuracy. Deep learning models consistently outperform traditional methods and machine learning in smile detection and recognition tasks. Especially with large datasets and complex scenarios involving diverse facial expressions and varying lighting conditions, deep models achieve advanced results in smile recognition benchmarks. [3]. Traditional methods often require redesigning feature extraction rules, are poor at achieving satisfactory results, and are less effective in dynamic, realistic environments with diverse faces and expressions. By leveraging machine learning and deep learning, future research can achieve unprecedented accuracy, adaptability, and scalability in smile detection and recognition, paving the way for innovative practical applications [4].

This review is motivated by the growing interest in smile detection and recognition, driven by advances in artificial intelligence algorithms as well as the increasing availability of large-scale standard datasets. Although various algorithms have been proposed over the years, they vary greatly in complexity, accuracy, and applicability. This diversity creates a need for comprehensive analysis to guide researchers in selecting appropriate methods for their specific use cases. The aim of this paper is to provide a systematic review of smile detection and recognition algorithms. It classifies the techniques into feature-based, machine learning, and deep learning approaches, compares their application strengths and limitations, and highlights recent and recent hybrid approaches. While exploring the evolution of traditional computer vision techniques to these techniques, which is the aim of writing this paper, challenges are identified, the impact of algorithmic quality on datasets is evaluated, and insights into future trends and a call for research and development in this important area are provided. We can summarize the contribution of this comprehensive study of smile detection algorithms based on three main approaches mentioned in the abstract (1) traditional computer vision, (2) machine learning, and (3) deep learning) as follows:

1. A systematic classification of the algorithmic approaches as to how they work strengths, weaknesses, and challenges. The classification of algorithms based on these algorithms was used into three distinct categories: feature-based, machine learning, and deep learning, such as using features and designed classifiers such as SVM or decision trees and extracting features automatically through neural network layers.

2. This work highlights the development of smile detection and recognition methodologies, providing a detailed assessment of the advantages and limitations of each approach.

3. An analysis of the techniques in terms of comparison between accuracy, performance, speed, and complexity of traditional methods, machine learning, and deep learning.

4. The most important contribution of this work is to demonstrate how modern AI methods outperform traditional methods in terms of capability, robustness, and capacity.

5. This research also focused on the balance of all the approaches used, as it was clear that using the feature-based approach as a basic feature is fast, but its speed is limited to simpler data, on the one hand, and on the other hand, the machine learning-based approach has better accuracy, but requires manual feature engineering intervention, and a third aspect depends on deep learning and is highly accurate, but its computation process is very expensive, which in turn made our paper focus on finding scenarios to solve each method appropriately.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

6. In terms of research gaps, our research highlighted smile detection and recognition algorithms that lack the ability to deal with diverse lighting conditions and face orientation and improve performance in real time.

It highlights unsolved challenges such as variation in lighting, posture, and facial expressions.

7. As for directions for future research, our approach was proposed and suggestions and paths were explained, especially hybrids that combine machine and deep learning, as well as traditional ones.

2. Related Work

A review and exploration of algorithms in smile detection and recognition was conducted with the aim of unifying the current knowledge and comparing their methodologies in this evolving field. The fragmented nature of research across multiple fields necessitated a comprehensive evaluation. This related work section provides an overview and insight into the study of important methods that have contributed to the development of smile detection systems.
Recently, smile detection systems have relied on feature-based methods, which are based on manually extracting facial features such as mouth curvature, facial shape, and eye contours. For example, developed a new method that uses mouth motion as geometric features to detect smiles. Similarly, used local binary patterns (LBP) as texture features to extract facial images to recognize different emotions, as well as smiles. These methods, although effective, often suffer from lighting variations and facial expression poses, making them less robust in real-world applications. To improve the accuracy of smile detection, researchers began applying classifiers such as support vector machines (SVMs) and k-nearest neighbors (k-NNs) with the advent of machine learning and deep learning. They used support vector machines trained to extract features from facial landmarks to detect expressions such as smiles and others, which greatly improved the robustness compared to previous methods. Similarly, they used machine-learning models with facial feature extraction techniques such as histogram of oriented gradients (HOG) and Gabor filters to enhance the accuracy of smile detection. These methods were a significant step forward, as they performed better in a variety of different conditions compared to traditional feature-based methods.

The emergence of deep learning has brought about a radical and major shift in smile detection and recognition, as convolutional neural networks (CNNs) have made significant advances in facial expression recognition, including smile detection. CNNs have been applied to the standard AffectNet dataset, and have shown remarkable success in facial expression and smile detection. VGG-Face, a deep CNN model fine-tuned for smile detection, has been introduced, demonstrating that deep learning can effectively learn complex features directly from raw image data. [4]

This work refers to techniques for improving the power and hybridization of multiple models and contributes to improving the accuracy and power of smile detection further. Proposed an integrated model that combines recurrent neural networks (RNNs), CNNs, and recurrent neural networks (RNNs) and gave positive and recent results for smile detection and recognition. In [5], this research referred to long short-term memory (LSTM) units to analyze smile dynamics from video clips. All these models are able to recognize smiles over time and deal with sequential dependencies between frames. In [6], synthetic smile data was generated to augment the training datasets using generative adversarial networks (GANs), which allowed the models to become more flexible to deal with differences in smile expressions and different lighting conditions.

In [7] Graph Neural Networks (GNNs) have also been explored for smile detection by representing faces as graphs where facial landmarks are nodes, and the relationships between them are edges. Used GNNs to improve facial expression recognition by analyzing the structural changes during smiles, offering more interpretability and accuracy. Decision Tree (DT) [8] as a classifier is a flowchart represented as a tree model. A DT splits the database into smaller sets of data until no more splits can be made, and the resulting leaves are the classes used for classification. The strengths of this classifier lie on the potential to learn nonlinear relationships of data, on handling high-dimensional data, and on the simple implementation. However, the main disadvantage of this classifier is overfitting, since it can keep branching until it memorizes the data during the training step. Random Forest (RF) [9] is essentially an ensemble classifier, consisting in a group of DTs. Each DT outputs a prediction, and the final prediction is based on majority voting, meaning that the most predicted class will be the last prediction. It has the advantage of reducing overfitting over just one DT, since it reduces the bias by averaging the predictions of the ensemble. However, it has the disadvantage of becoming slower when increasing its complexity (e.g., by adding more DTs to the ensemble). [10] Real-time smile detection is crucial for applications such as emotion-aware systems and human-robot interaction. Recent methods focus on improving both the speed and accuracy of smile detection in real-time environments. [11] explored the use of YOLO (You Only Look Once) for real-time facial expression recognition, including smiles, achieving a balance between accuracy and computational efficiency. Similarly, MediaPipe by Google [12] provides a highly efficient solution for real-time facial detection and expression recognition, including smiles, making it a popular tool in both research and industry applications. The development of datasets has been fundamental to training and evaluating smile detection models. [13] CK+ (Cohn-Kanade Extended Dataset) and [14] FER2013 are two widely used datasets for facial expression recognition, including smiles. [15] AffectNet, Larger datasets enable more accurate modeling by providing a larger set of facial expression images with more detailed smile labels. In [16], in the JAFFE

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

(Japanese Female Facial Expressions) dataset, the JAFFE dataset consists of 213 images of different facial expressions from 10 different Japanese females. With 7 facial expressions (6 basic and 1 neutral), the images were annotated with average semantic ratings for each facial expression by 60 annotators. These datasets provide diverse examples of smiles, from subtle to broad, and are instrumental in improving model generalization and performance. However, there is a common discrepancy of accuracy when testing in controlled environment databases and in wild environment databases. This shows a clear difficulty in translating the good results in controlled environments (such as CK+ and JAFFE) to uncontrolled environments (such as FER-2013 and SFEW) [17]. Every work that tested its algorithms with various databases obtained a significantly worse result on the uncontrolled environment ones.

The field of smile detection and recognition has evolved significantly, with early methods based on handcrafted features being surpassed by machine learning and deep learning techniques. Today, hybrid approaches combining different algorithms and advanced techniques like GANs, RNNs, and GNNs are pushing the boundaries of smile detection. The continuous evolution of these algorithms, along with the development of large-scale datasets, has made smile detection more accurate, robust, and applicable in real-world scenarios such as human-computer interaction, healthcare, and entertainment.

Because the number of databases used is very large and we cannot list them all, we have created a table that shows their details. Below is a comprehensive Table 1 summarizing the relevant literature on smile detection and recognition algorithms: researcher, year, publisher, dataset, and study objective.

Table 1: summarizing the relevant literature on smile detection and recognition algorithms: researcher, year, publisher, dataset, and study objective

Dataset	**Characteristics**	**Usage**
GENKI-4K, VGG-Face, GENKI-4K	4,000 images with smile and non-smile labels, diverse poses, lighting, and backgrounds.	Training and evaluating smile detection models.
FER2013	35,000 grayscale images, 48x48 resolution, labeled with seven emotions, including smiles.	Deep learning-based emotion recognition tasks.
JAFFE	213 images of Japanese female subjects with various facial expressions, including smiles.	Smile and facial expression recognition.
VGG-Face	2.6 million Images of 2,600 individuals for face and expression recognition tasks.	Fine-tuning deep learning models for smile detection.
EmotionNet	1 million images annotated for action units (AUs) related to emotions and smiles.	AU-based analysis for fine-grained smile detection.
YouTube Faces	Videos of faces, some annotated for smiles and other expressions.	Smile detection in dynamic environments.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

We also found during research and writing that the algorithms that were discussed in our research were based on other databases that were not mentioned in our research. Because of their importance, we mentioned them in a Table 2. That shows the algorithm, databases, year of publication, and publishers.

Table 2: that shows the algorithm, databases, year of publication, and publishers

Algorithm/Approach	**Researchers**	**Year**	**Dataset**	**Aim of Work**
Graph Neural Networks(GNNs)[18]	Jiang et al.	2021	FERPlus	Propsed GNNmodels for expression classification, including smile detection, emphasizing graph-based feature for robust recognition.
Histogram of Oriented Gradients(HOG)[19]	Dalal and triggs	2015	INRIA Pedestrian Dataset	Developed HOG features for object detection by capturing gradient orientation, later adapted for facial feature and smile recognition tasks.
Haar Cascades [20]	Viola and Jones	2015	Custom datasets	Introduced an AdaBoost framework with Haar-like features, initially for face detection, which was adapted for smile detection in facial regions.
Random Forest(RF) [21]	Breiman	2020	Various image datasets	Applied ensemble methods for robust classification, often combined with HOG for improving smile detection accuracy.
Decision Tree(DT)[22]	Quinlan	2021	FEROlus,CK+	Used for decision- making processes in machine learning pipelines, often forming the base of other ensemble techniques like RF.
YOLO(You Onal Look Once) [23]	Redmon et al.	2016	MS-COCO,FER datasets	Real-time object detection framework optimized for face and smile recognition through high-speed and accuracy trade-offs.

3. Traditional Computer Vision Techniques

3.1 Feature-Based Algorithms: They are computational techniques that rely on identifying and analyzing specific characteristics or attributes (features) of data to perform tasks such as classification, recognition, or detection. In the context of computer vision, features typically refer to distinct patterns or structures in images, such as edges, corners, textures, or keypoints that are used to describe and distinguish objects or regions. Feature-based algorithms were widely used before the advent of deep learning and remain useful in scenarios where computational efficiency and interpretability are critical [15]. The consist of:

a. **Haar Cascades**: based on the work of Paul Viola and Michael Jones in 2001, is a machine learning approach to object detection, including smile detection. It relies on Haar-like features, which are patterns similar to edge, line, and rectangle configurations used to detect specific objects or expressions within an image. It works on several feature extractions including: (1) Haar-like features analyze pixel intensity in rectangular regions, (2) Training with AdaBoost: A classifier is trained sequentially using labeled positive (images with smiles) and negative

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

(images without smiles) samples, (3) Detection: It scans the image at different scales and locations to detect smiles, using multiple stages (or layers) of classifiers in the early stages quickly detecting non-smiling regions, and later stages performing more precise analysis. The advantages of Haar Cascades for smile detection are: (1) Efficiency: Fast computing makes it suitable for real-time applications; (2) It struggles with variations in lighting or head, making it less accurate than modern deep learning techniques. Haar Cascades is Effective for detecting smiles but limited in generalizability. This Fig.1. illustrates Haar Cascades.[16]

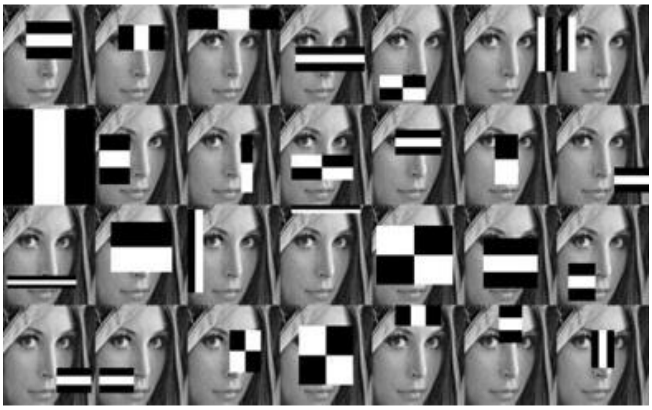

Figure 1. Haar Cascades [32].

b. Histogram of Oriented Gradients (HOG):is a powerful feature descriptor used in computer vision and image processing, designed to capture the structure and shape of objects within an image. To detect a smile, a gradient feature identification mechanism identifies the characteristic patterns associated with smiles, such as the curvature of the lips or the elevation of the cheek areas. The first main mechanism is: (1) Gradient calculations: HOG calculates the gradient (i.e. change in intensity) for each pixel in the image. Gradients emphasize edges, making them ideal for detecting realistic smile features. (2) Direction classifications: Gradients are divided into cells (small rectangular regions) and classified into direction bins based on their direction to detect a smile. This step highlights the typical curved regions of the lips for a smile. (3) Block Normalization: Groups of adjacent cells are normalized to make the feature descriptor invariant to lighting variations. This ensures that smiles are detected consistently under different conditions. (4) Feature Vector Creation: The gradient information from all cells is combined into a single feature vector, representing the smile pattern in the image. the second is classification with Support Vector Machines (SVM) is: (1) HOG is often paired with SVM for classification: HOG Features: Extracted from labeled training images (smiles vs. non-smiles): (2) SVM Training: Learns a decision boundary to classify new images based on the HOG feature vector: (3) Detection: When presented with new input, the SVM predicts whether a smile is present. Active Shape Models (ASM) and Active Appearance Models (AAM) Used to locate and track facial landmarks. Smiles can be detected by analyzing the shape and movement of the lips.[17] See Fig. 2.

Input image **Histogram of Oriented Gradient**

Figure 2. Explain Histogram of Oriented Gradients (HOG).

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

Advantages to (HOG): are (1) Robustness: It is effective in capturing edge and texture information. (2) Efficiency: It can work with relatively small datasets compared to deep learning models. (3) Stability: It is resilient to variations in lighting and slight changes in pose due to normalization. Disadvantages: Although deep learning has now overtaken HOG + SVM, it remains a staple method, especially for resource-constrained applications. To gain deeper insights. Applications in Smile Detection: HOG with SVM is widely used for tasks such as real-time emotion detection in controlled environments and smile recognition for surveillance systems.

c. **Edge Detection and Contour Analysis:** Edge detection and contour analysis for smile detection Edge detection is a very important technique in image processing that is used to identify the boundaries of objects by detecting large changes in pixel density. Edge detection can highlight the contours of facial features, especially the lips, which are key indicators of a smile. The main techniques are: (1) Canny edge detector: It is one of the most popular edge detection algorithms due to its ability to detect edges clearly while reducing noise. It is done by applying Gaussian smoothing to find the density gradients and to reduce the noise, which leads to identifying the areas of large changes in density. (2) Smile detection: is the use of a Kani detector to highlight the edges around the lips, where a smile is characterized by the curvature of the lips, so the edge detection algorithm can detect them effectively. Then, line analysis is used to trace the shape of the lines that reveal the distinctive shape of the lips as they curve upward. Its advantages are to detect certain features and work well on them to detect the lip borders, and what is very necessary for smile recognition is: (1) Light sensitivity: as poor lighting can cause the edge to be detected incompletely and incorrectly. (2) Face position: If the face is not directly in front of the camera, the detection may miss the edge of the lip border or distort it due to inaccurate positioning, and noise and complex backgrounds can interfere with the accuracy of lip edge detection. (3) Simple and effective: Edge detection is computationally efficient and can work in real-time applications. Its most important applications are: (1) Real-time smile recognition. (2) Healthcare to track therapeutic or diagnostic emotional conditions. [18] See Fig. 3.

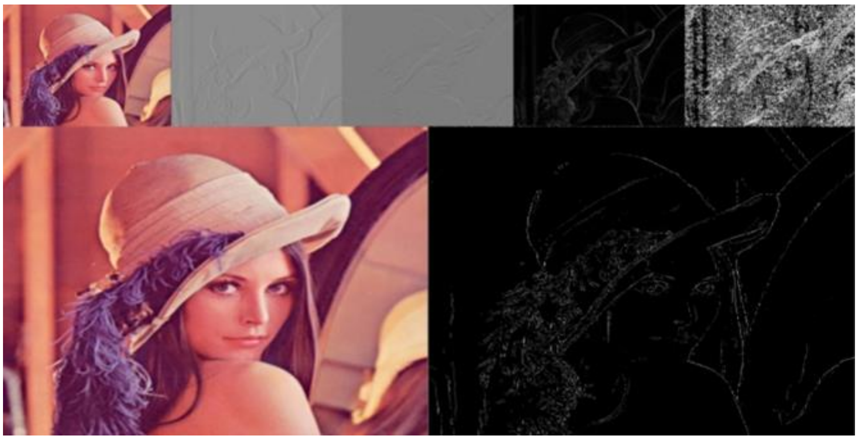

Figure 3. Edge Detection and Contour Analysis. [36]

3.2 Machine Learning-Based Approaches

Machine learning-based approaches for smile detection and recognition involve using algorithms that learn patterns from labeled data to identify and classify smiles in images or videos. These methods rely on training a model with examples of smiling and non-smiling faces, enabling the system to generalize and accurately detect smiles in new, unseen data. Also is a supervised machine-learning algorithm that classifies data by finding an optimal line or hyperplane that maximizes the distance between each class in an N-dimensional space.

a. **Support Vector Machines (SVM):** Support vector machines (SVMs) are one of the most widely used machine-learning algorithms for smile detection and recognition. They are a supervised learning model that finds the optimal level of separation between classes (e.g., smiling vs. non-smiling). SVMs are used in applications like handwriting recognition, intrusion detection, face detection, email classification, gene classification, and in web pages. This is one of the reasons we use SVMs in machine learning. It can handle both classification and regression on linear and non-linear data. When applied to smile detection, SVMs for smile detection work in the following stages: (1) Feature extraction: Key facial features, such as lips and cheeks, are extracted, and these features often contain valuable information for distinguishing smiling from non-smiling faces. (2) Feature extraction: SVMs are often combined with different feature extraction methods to improve their performance. (a) HOG (Histogram of Oriented Gradients): This technique extracts gradient-based features that highlight the edges and shapes of facial regions (especially lips). (b) (Principal Component Analysis) PCA: Principal

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

component analysis is used to reduce dimensionality, select the most relevant features for smile detection and ensure that the SVM focuses on the most informative parts of the image (3) Model training: The SVM classifier is trained using labeled datasets (images of smiling and non-smiling faces). It learns to distinguish between the two classes by identifying patterns in the extracted features. (a) Positive class: Images that contain smiles. (b) Negative class: Images that do not contain smiles (neutral or non-smiling expressions). (4) Classification: When a new image is presented, the SVM uses the learned hyperparameter to classify the image as a smile or not. The decision-making process of the SVM depends on how well the image features match the features learned from the training data.[18] See fig. 4.

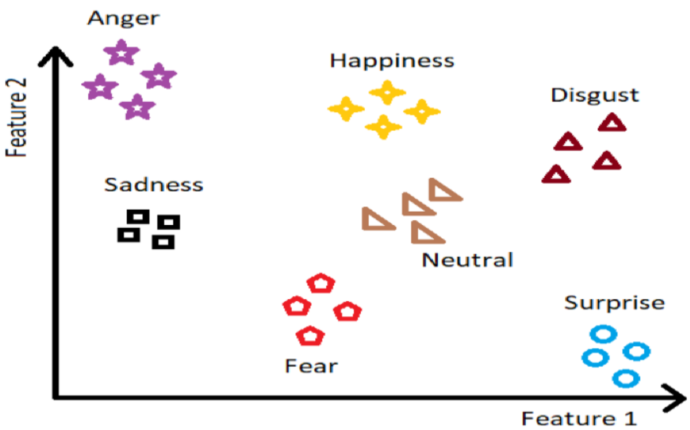

Figure 4. SVM algorithm for face detection and classification.

The advantages of SVM in smile detection are: (1) Effective with small datasets as SVM can perform well with limited amounts of data compared to deep learning methods. (2) High accuracy: Using appropriate feature extraction methods, SVM can achieve high accuracy in detecting smiles, especially in controlled environments. (3) Robustness: It is less prone to overfitting when tuned with regularization, which is critical for generalizing smile detection across different images. The performance of SVM is highly dependent on the quality of the features and feature extraction methods such as HOG or PCA must be chosen carefully. While SVM is effective for smaller datasets, it becomes computationally expensive as the dataset grows, especially when used with high-dimensional features such as HOG. SVM has difficulty performing well on images with different lighting, head poses, or occlusions, unless combined with more complex techniques such as deep learning. Some applications of SVM are: (1) Emotion recognition systems: SVM is used in systems that require real-time smile detection, such as customer feedback systems or virtual assistants. (2) Healthcare: SVM-based smile detection is used in therapeutic settings to monitor emotional states or interactions between patients. [19]

SVM can be of two types: (1) Linear SVM: Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.

(2) Non-linear SVM: Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.

By integrating feature extraction methods such as HOG and PCA with SVM, smile detection systems can achieve effective performance, especially in well-controlled environments. However, for more challenging, real-world scenarios, hybrid models combining deep learning and machine learning may be necessary.

(b) Decision Trees and Random Forests: Decision trees and random forests are supervised learning algorithms that partition data into subsets based on feature values, creating a tree-like structure where each node represents a feature and each branch represents a decision rule. They are also machine-learning algorithms that can be effectively applied to smile detection by learning patterns in facial features to classify smiles from neutral expressions [20]. These algorithms are popular because of their ability to handle complex classification and decision-making tasks in a clear and interpretable manner. They use features such as lip curvature, mouth position, or cheek contour. Based on these features, the tree decides whether an image represents a smile or a neutral expression. The decision tree is trained using a dataset containing labeled images (smiles and neutral faces). It repeatedly splits the data at each node based on the feature that best separates the data (using metrics such as Gini impurity or information gain). See Fig. 5.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

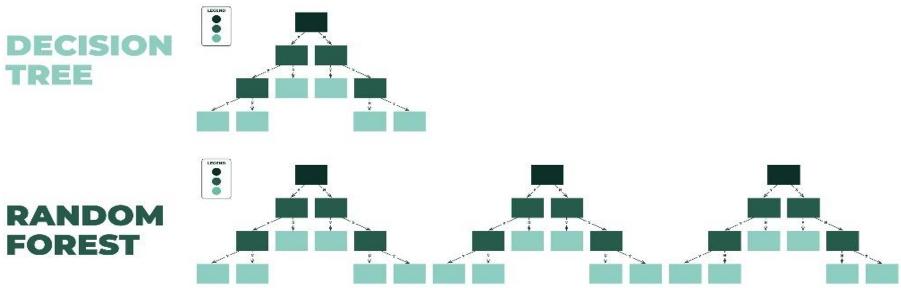

Figure 5. Decision Trees and Random Forests [49].

Its advantages are: (1) its interpretability is one of its main strengths, as it is highly interpretable and each decision in the tree can be traced back to a specific feature, making the model easier to understand and debug. (2) Efficiency on small datasets, as decision trees can be trained on smaller datasets with relatively lower computational requirements. Decision trees are prone to overfitting, especially when the tree becomes very deep. Pruning techniques are often required to avoid overfitting A single wrong or noisy data point can significantly impact the performance of the tree. Random forest is an ensemble learning method that combines multiple decision trees to improve performance. It creates an ensemble (forest) of decision trees by training each tree on a random subset of the data and using random features in each split. Each individual tree in the random forest makes its prediction (smiles or not), and the final prediction is made by averaging the results of all the trees (for regression tasks) or by majority voting (for classification tasks). When detecting a smile, random forests can use multiple facial features (e.g., lip curvature, mouth width, and cheek height) where complex patterns are learned from different subsets of features. This helps improve the generalization and accuracy of the smile detection model. [21]

Some of the advantages include: (1) Improved accuracy: By combining multiple decision trees, random forests reduce the risk of overfitting and can improve the generalization of the model to new, unseen data. (2) Robustness: Random forests are more robust to noise and outliers than individual decision trees. Random forests can provide insight into the importance of different features in detecting smiles, allowing for the selection of the best features for the features. (3) Random forests outperform individual decision trees in accuracy, but are less interpretable due to the combination of multiple trees.

applications of smile detection are: (1) Emotion recognition systems: Decision trees and random forests are often used in real-time smile detection systems, where facial features such as lip curvature and cheek height are key indicators. (2) Human-computer interaction: These algorithms can be used to recognize smiles in gaming applications and video calls to respond interactively. (3) In treatment or diagnostic and care settings, facial expression and smile detection can help monitor emotional states. Both decision trees and random forests provide effective methods for smile detection by learning facial feature patterns. While decision trees provide clear interpretability, they also provide better accuracy and robustness by aggregating multiple decision trees. These algorithms can be particularly useful when combined with other feature extraction methods (e.g., HOG and PCA) to improve performance in both controlled and real-world scenarios. [22]

4. Deep Learning-Based Approaches:

Deep learning-based smile detection and recognition methods rely on advanced machine learning techniques, especially neural networks, to automatically interpret and analyze smile patterns in facial videos or images. These methods use hierarchical feature extraction, and deep neural networks, such as convolutional neural networks (CNNs), learn complex representations of data directly from raw inputs such as pixel values. This enables high accuracy and efficiency in smile detection and recognition.

These methods can be scaled and improved using large datasets, making them suitable for a variety of real-world applications, including social robotics, user experience enhancement, and sentiment analysis. These key techniques include convolutional neural networks, transfer learning, and region-based object detection models, which together form a powerful framework for automatic smile analysis. They consist of: (1) Convolutional Neural Networks (CNNs): (a) Convolutional neural networks are widely used in smile detection and recognition due to their ability to automatically learn features directly from pixel data, eliminating the need for handcrafted features. (b) The main

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

layers in CNNs (e.g., convolutional layers, pooling layers, and fully connected layers) extract hierarchical features such as edges, textures, and patterns that are critical for identifying smiles. Examples of architectures are (AlexNet, VGG, and ResNet). (2) Transfer learning: pre-trained models (e.g., VGGFace, MobileNet, and ResNet) are adapted to smile detection by fine-tuning on specific smile datasets. [23] See Fig. 6.

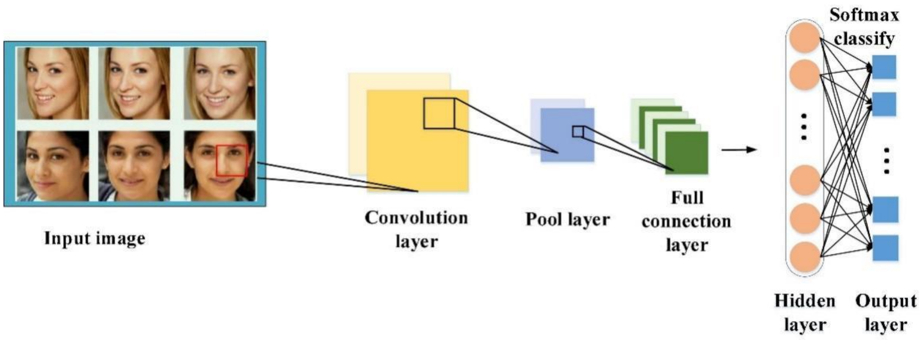

Figure 6. Here is a conceptual illustration of a deep learning-based approach for smile detection and recognition, showing the flow from input (image of a face) through a Convolutional Neural Network (CNN) to the output.

The advantages include: (a) Time and resource efficiency: Reusing knowledge from large-scale datasets such as ImageNet. (b) Improved performance: Pre-trained models provide robust feature representations. (c) Fine-tuning typically involves freezing lower-level layers and training higher-level layers to specialize in smile detection tasks. (d) Focusing specifically on facial regions, reducing noise from irrelevant parts of the image. (e) Improving accuracy by targeting regions most relevant for smile analysis. (3) Region-based convolutional neural networks (R-CNN). R-CNN and its variants (Fast R-CNN, Faster R-CNN, and Mask R-CNN) integrate object detection and classification. Its operation involves the following steps: (a) Regions of interest (ROI) are proposed from the image to isolate facial regions. (b) The extracted regions are then classified as smiling or non-smiling. [22]

5. Hybrid and Advanced Techniques:

Hybrid and advanced smile detection techniques combine multiple deep learning methodologies and introduce new architectural frameworks to enhance the accuracy, efficiency, and adaptability of smile analysis systems. These approaches integrate temporal and structural modeling to address challenges such as micro expression recognition, data overload, and complex facial dynamics. Especially in challenging scenarios such as dynamic environments, micro expressions, or low-data-demanding situations, these hybrid and advanced techniques leverage temporal, generative, and structural insights to enhance the accuracy and robustness of smile detection and recognition systems. Leveraging innovations such as recurrent neural networks (RNNs) with long short-term memory (LSTM) modules, generative adversarial networks (GANs), and graph neural networks (GNNs), these approaches go beyond traditional frameworks. They focus on capturing sequential dependencies in video-based smiles, synthesizing diverse training data, and representing facial features as graph structures, enabling more robust and accurate smile detection under diverse conditions. (1) Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) modules: Recurrent neural network descriptions, especially those enhanced with Long Short-Term Memory (LSTM) modules, are effective in analyzing sequential data by preserving information over time. Its advantages include: (a) It is suitable for recognizing smiles, especially video-based ones. (b) It deals with variations in the duration of a smile since its onset. (2) Generative Adversarial Networks (GANs): GANs consist of two networks - (a) a generator that generates synthetic data, and (b) a discriminator that evaluates its validity, to recognize subtle differences in the expression of any smile while generating various training examples. Its advantages include: (a) It improves the robustness of models in different conditions. (b) It enables the detection of subtle differences in smiles, such as polite or sincere smiles. (3) Graph Neural Networks (GNNs): Graph neural networks process data organized in the form of graphs, where nodes and edges represent entities and their relationships, respectively, and analyze structural changes in the graph during the formation of a smile. Its advantages include: (a) Captures spatial relationships between facial features. (b) Effective in modeling complex geometric transformations during smile expressions. (c) Particularly useful for detecting asymmetric or subtle smiles. [23] See Fig. 7.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

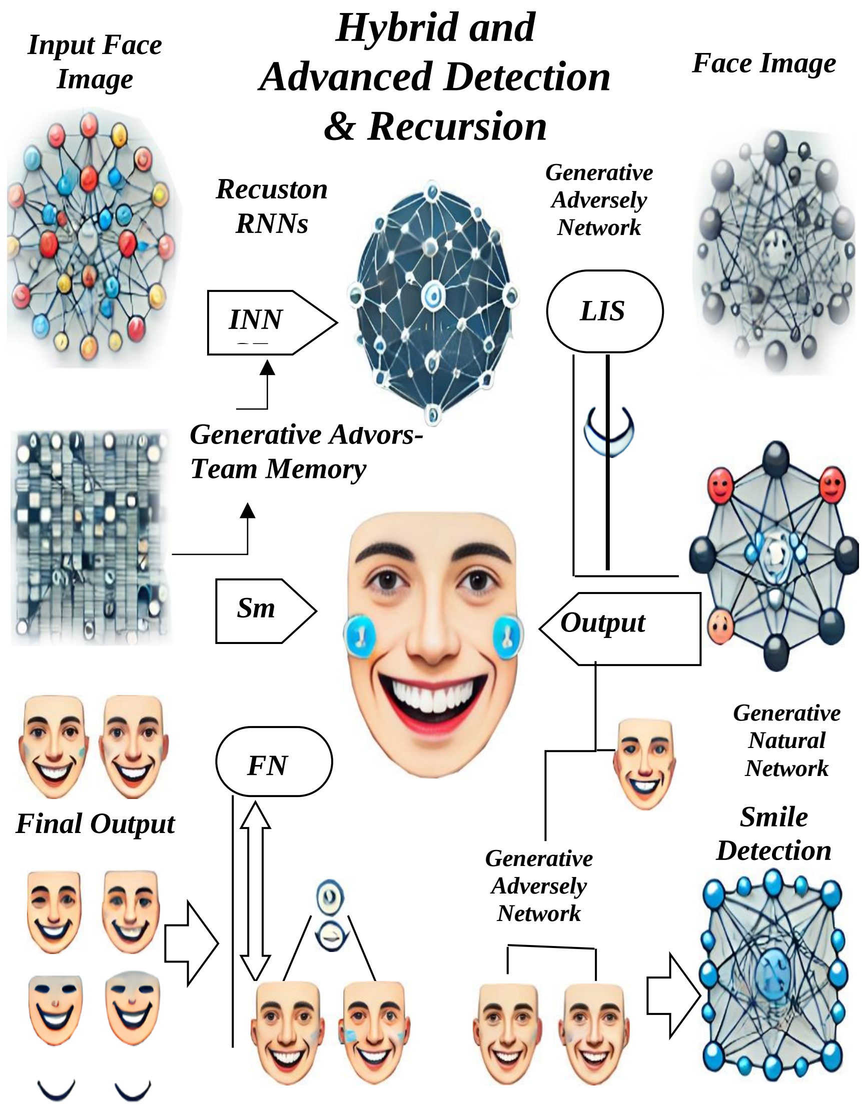

Figure 7. Here is a conceptual diagram illustrating hybrid and advanced techniques for smile detection and recognition. It includes components such as RNNs with LSTMs, GANs for synthetic data, and GNNs analyzing facial landmarks.

6. AI Frameworks for Smile Detection:

AI frameworks for smile detection and recognition are libraries and software tools designed to facilitate the development, training, and deployment of models for analyzing and identifying smiles in images or videos. These frameworks provide pre-trained models, algorithms, and auxiliaries for tasks such as facial feature detection, expression recognition, and real-time analysis.

These frameworks provide a comprehensive set of tools and libraries for performing smile detection and recognition, and cater to a variety of needs ranging from lightweight solutions to advanced deep learning-based analysis. These frameworks include:

(1) OpenCV: which provides pre-trained models for fast smile detection, and is a popular library for computer vision tasks (open-source computer vision library), providing tools for image and video analysis. One of its applications in smile detection is that it provides pre-trained Haar Cascade classifiers for smile detection, enabling fast and efficient implementation. It supports real-time smile detection using lightweight models. It also has the advantages of (a) Easy to use and integrate into applications. (b) Open source and highly customizable [24]. See Fig. 8.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

Figure 8. Opencv python [71][72]

(2) Dlib: which specializes in facial feature detection for smile dynamics analysis, is a powerful machine-learning library with computer vision tools, including face recognition and feature detection. Its applications in smile detection include: (a) Detects facial features (such as mouth corners and eyes) for smile dynamics analysis. (b) Tracks facial regions to ensure accurate smile analysis even in videos. It also has the following advantages: (a) accurate facial feature detection. (b) Lightweight and works well with deep learning frameworks [25]. See Fig. 9

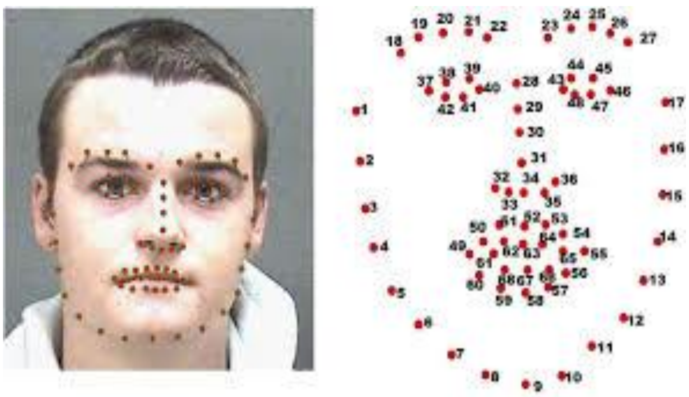

Figure 9. Identification of facial landmarks using Dlib. a Facial landmark.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

(3) DeepFace: A deep learning-based framework that supports advanced facial analysis such as recognition, age estimation, and expression detection. These frameworks simplify the implementation of smile detection systems by providing accessible, efficient, and scalable solutions. Its applications in smile detection include: (a) Recognizes and analyzes facial expressions, including smiles, with high accuracy. (b) Efficiently integrates pre-trained models for smile recognition tasks. Its advantages include: (a) advanced accuracy in expression detection. (b) Support for multiple backends, including TensorFlow and PyTorch, for scalability [26]. See Fig. 10.

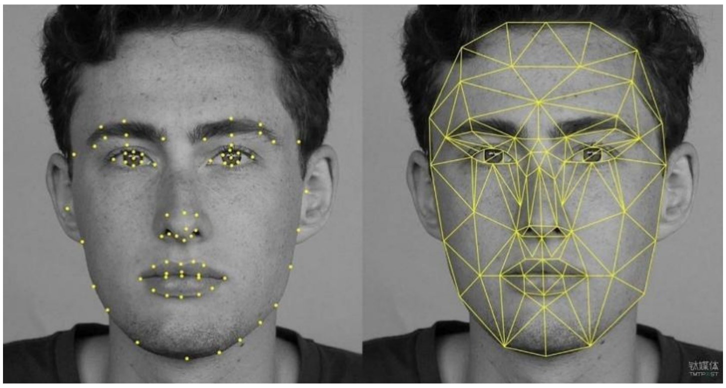

Figure 10. Deepface model [79].

7. Real-Time Detection and Recognition Techniques:

Real-time smile detection and recognition techniques are computational methods and frameworks designed to analyze and identify smiles instantly from live video streams or dynamic environments. These methods prioritize speed and efficiency while maintaining accuracy, enabling seamless smile detection in interactive and time-sensitive applications. It consists of: (1) YOLO (You Only Look Once): It is a real-time facial expression-based object detection framework known for its balance of accuracy and speed. Its applications include: (a) it is dedicated to facial expression recognition tasks exclusively, including smile detection. (b) It classifies expressions in a single forward pass of the network, enabling real-time performance. Its advantages include: (a) High-speed detection suitable for live applications. (b) It maintains accuracy even in dynamic environments [27]. See Fig. 11.

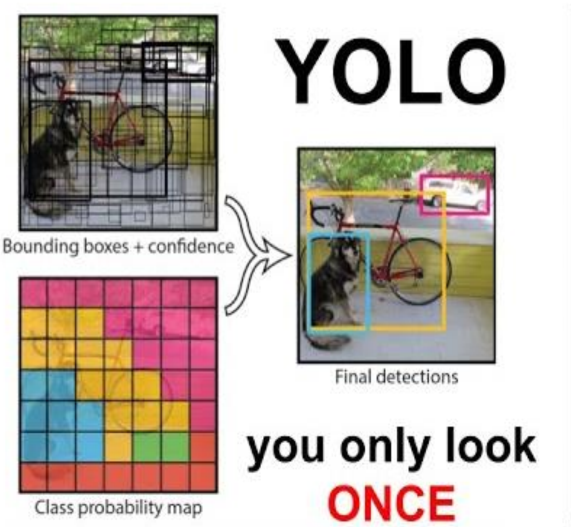

Figure 11. YOLO running on sample artwork and natural images from the internet. It is mostly accurate although it does think one person is an airplane.

(2) MediaPipe by Google: A cross-platform framework developed by Google for building multimodal machine learning pipelines, including face detection and tracking. Its applications include: (a) It provides efficient real-time face and smile detection across different platforms, (b) It uses lightweight models optimized for mobile and embedded devices. Its advantages include: (a) High efficiency and accuracy, (b) It supports cross-platform deployment on mobile, web, and desktop [28]. See Fig. 12.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

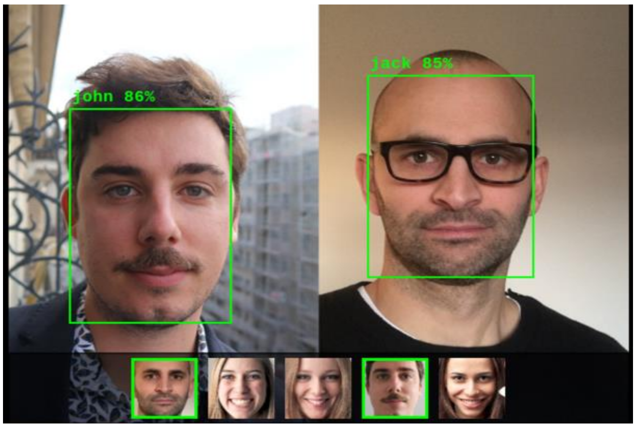

Figure 12. MediaPipe by Google [82].

These technologies enable and are widely used to seamlessly detect and recognize smiles in real-time, making them ideal for applications such as interactive systems, live video analysis, and social robotics.

8. Datasets for Training Models

Smile detection and recognition training datasets are collections of annotated images, videos, or facial landmarks that are used to teach machine-learning models to recognize smiles and other facial expressions. These datasets typically include diverse samples representing different individuals, ethnicities, lighting conditions, and facial movements. Using such datasets, models learn to detect patterns associated with smiles, such as mouth movement and surrounding facial muscles.

These datasets are essential for training robust and accurate models that can recognize smiles in a variety of real-world scenarios. Notable datasets for training smile detection models include (1) **CK+:** (Extended Cohn-Kanade Dataset): This is a widely used dataset for facial expression analysis, containing annotated images and videos of facial expressions. It includes labels for facial expressions, including smiles. It is (1) ideal for training and validating models for detecting expressions and smiles, (2) high-quality images, and (3) includes still frames and video sequences . See Fig. 13. Dataset illustrating.

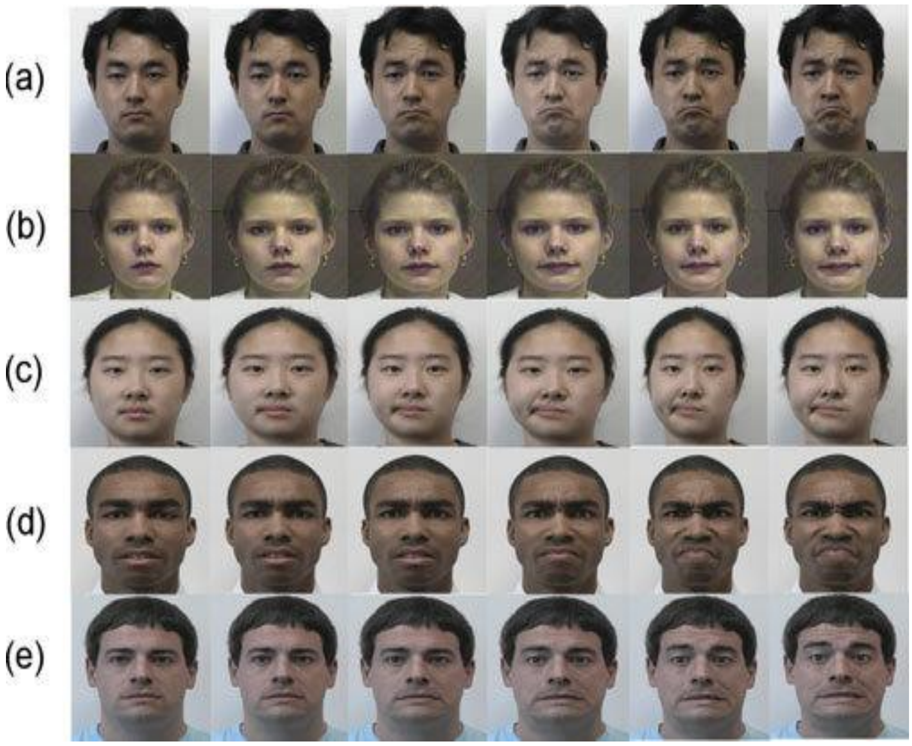

Figure. 13. Examples from the CK+ dataset illustrating the strong temporal links present within neighboring frames among different expressions, (a) sadness, (b) happiness, (c) contempt, (d) anger and (e) fear.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

(2) **FER2013:** (Facial Expression Recognition 2013) is a Kaggle-hosted dataset of grayscale facial images for emotion recognition. It contains expressions, including happiness (smiles), making it suitable for training smile recognition models. Its advantages include (1) useful for comparison due to its inclusion in Kaggle challenges, (2) large dataset with over 35,887 images, (3) diverse facial expressions in challenging real-world conditions. See Fig. 14. Fer2013 used for facial expression recognition. Grayscale facial images containing 7 different emotions (anger, disgust, fear, happiness, neutral, sad, and surprise).

Figure 14. Fer2013 dataset is a common dataset used for facial expression recognition. The dataset contains 35,887 grayscale facial images containing 7 different emotions (anger, disgust, fear, happiness, neutral, sad, and surprise).

(3) **AffectNet:** A large-scale, large-scale dataset and the largest facial expression recognition dataset with a focus on emotion detection, covering a wide range of expressions including smiles. Its advantages include (1) annotated images for basic and complex expressions, including happiness (smiles) and variations in pose, lighting, and occlusions, (2) well-suited for training deep learning models that require large-scale data, (3) over a million images annotated with facial expression tags [29].

Figure 14. Samples of queried image from the web and their annotated tags. The queried expression is written in parentheses.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

(4) JAFFE: (Japanese Female Facial Expressions): The JAFFE dataset consists of 213 images of different facial expressions from 10 different Japanese female subjects. Each subject was asked to do 7 facial expressions (6 basic facial expressions and neutral) and the images were annotated with average semantic ratings on each facial expression by 60 annotators. JAFFE contains facial expression samples from Japanese female subjects, designed for sentiment analysis and expression dataset containing facial images that capture and provide annotated samples for various emotions, including smiles, making it valuable for targeted smile detection tasks. Its advantages include: (1) high-resolution images, (2) small but well-annotated dataset, ideal for initial experiments. These datasets are useful for training and evaluating machine-learning models for smile detection and recognition, as they provide a variety of expressions, image quality, and scenarios to improve model performance [30]. See Fig. 15.

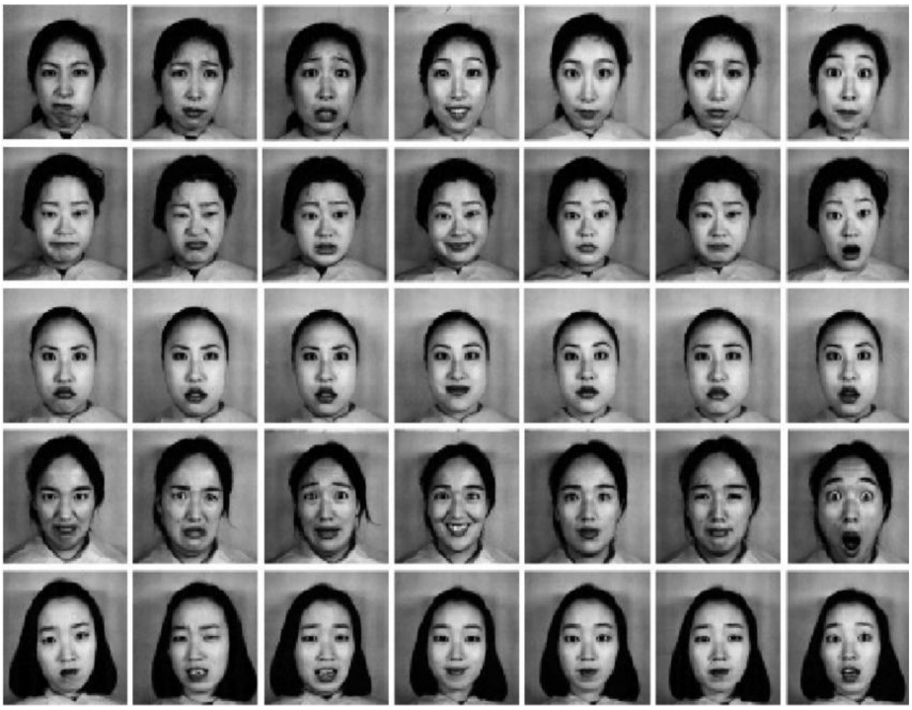

Figure 15. Examples of facial expression images from the JAFFE database.

9. Evaluation Metrics

Evaluation metrics are quantitative measures used to evaluate the performance, effectiveness, and reliability of an algorithm, model, or system. These metrics provide insights into how well the algorithm achieves its intended goals, allowing for comparison, improvement, and verification. Their purpose is to ensure that the model meets the required performance criteria in specific applications or scenarios. They consist of several types that vary depending on the context, such as classification accuracy, error rates, and computational efficiency. Their importance is that they guide improvements by identifying the strengths and weaknesses of the algorithm or system. Evaluation metrics consist of:

1. **Performance metrics:** Measure the ability of an algorithm to perform the required task (e.g., accuracy, precision, recall) because accuracy alone can be misleading for imbalanced datasets, so accuracy and recall must be analyzed together to gain accurate insights.
2. **Efficiency metrics:** Evaluate resource usage such as time and computational power (e.g., processing time, memory consumption) as computational efficiency affects scalability and usability in resource-constrained systems.
3. **Robustness metrics:** Robustness ensures reliability in realistic scenarios with incomplete inputs. Where the stability of the algorithm is evaluated under different conditions (e.g., lighting changes, occlusions). Evaluation metrics are chosen based on the specific application and trade-offs between factors such as speed, accuracy, and real-world application. The F1 score is critical to determining overall performance in scenarios that require trade-offs.

Here is a Table 3 listing and explaining common evaluation metrics for algorithms, including accuracy, precision, recall, F1-score, computational efficiency, and robustness.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

Table 3: listing and explaining common evaluation metrics for algorithms.

Metric	**Description**	**Formula/Measurement**	**When to Use**	**Importance in Smile Detection and Recognition**
Accuracy [30]	Measures the proportion of correct predictions to the total predictions.	Accuracy= $\frac{\text{True Positives}+\text{True Negatives}}{Total}$	Use when classes are balanced, as it does not differentiate between types of errors.	- Provides an overall sense of how well the algorithm performs. - Does not account for class imbalances or the cost of misclassifications.
Precision [31]	Indicates the proportion of true positive predictions among all positive predictions.	Precision= $\frac{\text{True Positives}}{\text{True Positives}+\text{False Positives}}$	Use when false positives are costly (e.g., spam detection).	- Ensures the algorithm is not falsely detecting smiles (important for applications requiring high specificity, like healthcare diagnostics).
Recall [32]	Measures the proportion of actual positives correctly identified by the algorithm.	Recall= $\frac{\text{True Positives}}{\text{True Positives}+\text{False Nagatives}}$	Use when false negatives are costly (e.g., disease detection).	- Measures the algorithm's ability to detect all smiles, even subtle or hard-to-detect ones. - Crucial for real-time or assistive applications.
F1-Score [33]	Harmonic mean of precision and recall, balancing the two metrics.	F1-Score=2 $\times \frac{Precision \times recall}{Precision + recall}$	Use when there's a need to balance precision and recall, especially with imbalanced classes.	- Useful in scenarios where false positives and false negatives carry different implications, such as customer feedback analysis.
Robustness [34]	Evaluates how well the algorithm performs under varying conditions like lighting, angles, or occlusions.	Empirical testing with datasets containing variations.	Crucial for tasks like smile detection under diverse environmental conditions.	
False Positive Rate (FPR) [35]	The percentage of non-smile cases incorrectly classified as smiles.	- Helps identify over-sensitivity in detection, which could lead to irrelevant outputs in marketing or user feedback systems.	False Positive Rate (FPR)	The percentage of non-smile cases incorrectly classified as smiles.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

10. Comparison of Techniques

Technique comparison refers to the evaluation and differentiation of algorithms or models based on specific criteria to determine the most appropriate approach for a particular problem or application. It highlights the strengths, weaknesses, and optimal use of each technique (e.g., supervised or unsupervised deep learning). The data used refers to the type or criteria used to train and evaluate the technique and is important in choosing the most appropriate method for specific tasks. For example, comparing a convolutional neural network (CNN) to a support vector machine (SVM) in image classification can reveal that CNNs are better at extracting features in complex image data, while SVMs excel on small datasets with fewer features. By systematically comparing techniques, researchers and practitioners can make informed decisions tailored to specific goals and constraints. Here is a detailed Table 4. Comparing algorithms based on the specified criteria:

Table 4: Comparing Algorithms Based on the Specified Criteria.

Algorithm	**Category**	**Performance**	**Dataset Used**	**Applications**	**Advantages**	**Limitations**
Support Vector Machines (SVM) [28]	Supervised Learning	High accuracy for small to medium-sized datasets, especially for binary classification tasks.	UCI Machine Learning Repository, MNIST	Image classification, Text categorization	Effective in high-dimensional spaces; robust to overfitting.	Computationally expensive for large datasets; kernel selection is critical.
Random Forest [29]	Ensemble Learning	High accuracy for classification and regression; less prone to overfitting than decision trees.	Kaggle Datasets, UCI Repository	Fraud detection, Healthcare predictions	Handles large datasets well; interpretable through feature importance.	Not suitable for real-time applications due to slower predictions.
Convolutional Neural Networks (CNN) [26]	Deep Learning	Excellent performance for image and video data; state-of-the-art for tasks like object detection.	ImageNet, CIFAR-10, COCO	Facial recognition, Autonomous vehicles	Highly accurate for image-related tasks; automatic feature extraction.	Requires large datasets and computational resources; prone to overfitting on small datasets.
K-Nearest Neighbors (KNN)\[22]	Instance-Based Learning	Moderate performance, dependent on dataset size and choice of k.	UCI ML Repository, OpenML	Recommendation systems, Pattern recognition	Simple and interpretable; no training phase required.	Computationally expensive during inference; sensitive to irrelevant features
K-Means Clustering [20]	Unsupervised Learning	Good for partitioning data into groups, but dependent on initialization and number of clusters.	Customer datasets, Transaction data	Customer segmentation, Market analysis	Simple and fast; scales to large datasets.	Struggles with non-spherical clusters; sensitive to outliers.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

Recurrent Neural Networks (RNN) [30]	Deep Learning	Effective for sequence data, but may face vanishing gradient issues.	IMDB Sentiment Analysis, Time-series datasets	Language modeling, Time-series predictions	Captures temporal dependencies; suitable for sequential tasks.	Training is computationally expensive; struggles with long-term dependencies.
Naïve Bayes [19]	Probabilistic Learning	Performs well with small datasets and text data; fast training and inference.	UCI ML Repository, Spam datasets	Spam detection, Sentiment analysis	Simple, fast, and effective for text-based problems.	Assumes feature independence; may not perform well with correlated features.

11. Applications

The different areas where smile detection and recognition are being applied are particularly focused on sentiment analysis, marketing, customer feedback, assistive technologies, healthcare (therapy and diagnosis), enhancing accessibility and driving innovation in understanding human emotions. Table 5 summarizes the areas where smile detection and recognition are being applied, with explanations for each:

Table 5: summarizes the areas where smile detection and recognition are being applied.

Domain	Description/Application	Examples	Benefits	Challenges
Emotion Analysis	Smile detection is used to infer emotional states of individuals by analyzing facial expressions. It helps in understanding user reactions and emotional responses in real-time.	-Sentiment analysis in social media -Enhancing human-computer interactions -Real-time feedback in virtual meetings	-Improved user experience through adaptive interfaces -Enhanced understanding of user emotions -Facilitates empathetic AI interactions -Improves human-computer interaction; -Enhances user experience in gaming, virtual reality, and online meetings.	-Cultural variations in expressing emotions -Subtles smiles vs.genuine -Privacy concerns
Marketing and customer feedback	Businesses utilize smile recognition to gauge customer satisfaction and engagement. By analyzing customer expressions companies can assess the effectiveness of products services, and marketing campaigns.	-Real stor analytic -Advertising effectiveness studies -Customer satisfaction surveys	-Objective measurement of customer emotions -Data-driven insights for improving products and services -Enhanced targeting of marketing strategies	-High variability in individual expressions -Environmental factors affecting accuracy -integration with existing systems
Assistive Technologies	Smile deduction aids in creating more intuitive and responsive assistive	-Smart home devices	-Increases accessibility and ease of use	-Ensuring reliability and accuracy

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

	devices for individuals with disabilities. It enhances communication and interaction between users and technology by recognizing and responding to facial cues.	responding to user emotions -interactive tools for individuals with speech impairments -Emotion aware robotics	-Enhanced user-device interaction - improves usability of devices for individuals with physical or sensory limitations.	-addressing diverse user needs and expressions -Maintaining user privacy
Healthcare (Therapy and Diagnostics)	In healthcare, smile recognition assists in monitoring patient well-being, diagnosing mental health conditions and providing feedback during therapy seeions.It enables changes and patient progress objectively.	-Mental health monitoring tools -Therapeutic feedback systems -Diagnostic tools for neurological conditions	-Objective assessment of the patient's feelings increases the effectiveness of treatment - Early detection of emotional and psychological problems	-Ensuring data security and patient privacy -Handling sensitive emotional data -Integrating with clinical workflows

Smile detection and recognition technologies have diverse applications across multiple domains, each benefiting from enhanced emotional understanding and interaction capabilities. While the advantages are significant, addressing the associated challenges is crucial for the effective and ethical implementation of these technologies.

12. Future Directions for Smile Detection and Recognition

The field of smile detection and recognition is evolving rapidly, with opportunities for significant advancements to improve accuracy, robustness, and ethical deployment. Below are some potential improvements and emerging trends.

1- Combining multiple types of data (e.g., audio, video, and physiological signals) can enhance the accuracy and reliability of smile detection systems.

2- Combining facial expressions with audio cues, such as laughter or tone of voice, provides richer emotional context. For example, detecting a laugh alongside a smile can improve differentiation between genuine and polite smiles.

3- Incorporating data from wearable devices (e.g., heart rate variability, skin conductance) assesses emotional states more holistically.

4- In scenarios like video conferencing, integrating spoken or written text with visual expressions can add valuable context to smile interpretation.

5- Providing users with clear explanations of how smile recognition systems work and how their data is used. Here is a detailed Table 6 summarizing the future directions for smile detection and recognition:

Table 6: summarizing the future directions.

Future Direction	**Description**	**Benefits**	**Challenges**
Use of Mulita-Model Data	Combining multiple data types (e.g., audio, video, physiological signals) to improve smile detection.	-Improved accuracy for ambiguous or subtle smiles. -Holistic understanding of emotions.	-Increased computational complexity. -Synchronizing data from diver's sources. -Privacy concerns with sensitive modalities.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

More Robust Algorithms for Real-World Setting	Adapting algorithms to handle diverse a challenging environments.	-Enhanced reliability in real-world applications. -Broader adoption in resource-constrained environments.	-Balancing computational efficiency with accuracy. -Managing biases due to environmental variability.
Inclusion of Ethical considerations	Embedding ethics into development and deployment to ensure fairness, privacy, and transparency.	-Builds user trust and compliance with regulations. -Reduces risks of misuses or biased outcomes.	-Balancing ethical goals with business technological constraints. -Addressing global regulatory and cultural differences.

Future directions in this field aim to detect and recognize smiles and create accurate, adaptable, and ethically sound systems. Together, these directions lead to the development of smile detection technologies for comprehensive applications. By leveraging multimodal data and integrating ethical frameworks, these technologies can achieve wider adoption and impactful applications while respecting user rights and societal norms.

13. Conclusion

This review of traditional computer vision techniques for smile detection and recognition highlights the fundamental role of these methods in understanding, analyzing, and accurately detecting and recognizing smiles from facial expressions. Despite the development and progress made in machine and deep learning, traditional methods remain important due to their interpretability, efficiency, and historical contributions to the field. these algorithmic approaches often rely on hand-crafted feature extraction techniques, such as HOG (Histogram of Oriented Gradients), Haar Cascades, and classifiers such as Support Vector Machines (SVM), k-nearest neighbors (k-NN), decision trees, and random forests, which are essential for classifying smile and non-smile expressions. The advantages of these methods are numerous, including efficiency, as these methods are less computationally intensive compared to deep learning, making them better suited for resource-constrained environments, and interpretability, as these handcrafted features provide clear insights into the aspects of the image that contribute to smile detection. With fewer parameters to modify, traditional algorithms are easier to implement and debug. Traditional methods suffer from variations in lighting and angles across diverse datasets and complex real-world scenarios, and these methods perform imperfectly with these datasets. Their influence continues in hybrid systems that combine traditional features with machine learning models or as benchmarks for newer data-driven approaches. Moreover, their efficiency and simplicity make them applicable to edge applications and low-resource environments. The future prospects for these methods are to combine traditional feature extractors with lightweight neural networks to improve performance in diverse conditions, extend traditional methods to unseen domains or in combination with audio and text analysis, and advantage the transparency of traditional methods to ensure smile detection systems. Our conclusion is that traditional computer vision techniques for smile detection and recognition remain the cornerstone of the field. They provide valuable insights into feature design and algorithmic simplicity, providing a strong foundation for continued innovation. As the field evolves, these methods continue to serve as complementary standards and tools in the broader landscape of smile detection and recognition techniques.

References

[1] Saeed, V. A. (2024). A Framework for Recognition of Facial Expression Using HOG Features. International Journal of Mathematics, Statistics, and Computer Science, 2, 1–8. https://doi.org/10.59543/ijmscs.v2i.7815

[2] Huang, Z., Zheng, H., Li, C., & Che, C. (2024). Application of machine learning-based k-means clustering for financial fraud detection. Academic Journal of Science and Technology, 10(1), 33-39.

[3] Passos, L. A., Jodas, D., Costa, K. A., Souza Júnior, L. A., Rodrigues, D., Del Ser, J., ... & Papa, J. P. (2024). A review of deep learning-based approaches for deepfake content detection. Expert Systems, 41(8), e13570.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

[3] Sedaghatjoo, Z., Hosseinzadeh, H., & Bigham, B. S. (2024). Local Binary Pattern (LBP) Optimization for Feature Extraction. arXiv preprint arXiv:2407.18665.

[4] Aydın, Muhammed Taha, Oğuzhan Menemencioğlu, and İlhami Muharrem Orak. "FACE RECOGNITION APPROACH BY USING DLIB AND K-NN." Current Trends in Computing 1.2 (2024): 93-103.

[5] Fard, Ali Pourramezan, et al. "AffectNet+: A Database for Enhancing Facial Expression Recognition with Soft-Labels." arXiv preprint arXiv:2410.22506 (2024).

[5] Rabea, Mahmoud, et al. "IdentiFace: A VGG based multimodal facial biometric system." arXiv preprint arXiv:2401.01227 (2024).

[6] Etemad, Shervin, et al. "SP25. Design And Implementation Of A Hybrid Paper-electronic Patient Data Tool For Operation Smile: Best Practices And Lessons Learned Across 20 Countries." Plastic and Reconstructive Surgery–Global Open 12.S4 (2024): 96-97.

[7] He, Ying, et al. "De novo Design of Biocompatible Nanomaterials Using Quasi-SMILES and Recurrent Neural Networks." ACS Applied Materials & Interfaces (2024).

[8] Rao, K. Venkateswara, Kunjam Nageswara Rao, and G. Sita Ratnam. "Accelerating Drug Safety Assessment using Bidirectional-LSTM for SMILES Data." arXiv preprint arXiv:2407.18919 (2024).

[9] Cherian, A. K., Vaidhehi, M., Arshey, M., Briskilal, J., & Simpson, S. V. (2024). Generative adversarial networks with stochastic gradient descent with momentum algorithm for video-based facial expression. International Journal of Information Technology, 16(6), 3703-3722.

[10] De Carlo, Alessandro, et al. "Predicting ADMET Properties from Molecule SMILE: A Bottom-Up Approach Using Attention-Based Graph Neural Networks." Pharmaceutics 16.6 (2024): 776.

[11] Costa, V. G., Pérez-Aracil, J., Salcedo-Sanz, S., & Pedreira, C. E. (2024). Evolving interpretable decision trees for reinforcement learning. Artificial Intelligence, 327, 104057.

[12] Yakar, Murat, et al. "An Innovative Fusion of New Smile Random Forest and Fuzzy Analysis for Comprehensive Seismic hazard Assessment: A Case Study in Mersin, Türkiye." (2024).

[13] Cherian, Aswathy K., et al. "A Robust Design of Real-Time Resilient Smile Recognition System using Hybrid Deep Learning Principles." 2024 10th International Conference on Communication and Signal Processing (ICCSP). IEEE, 2024.

[14] Vanamoju, Sri Vaibhav Mohan Dev, et al. "Facial Emotion Recognition using YOLO based Deep Learning Classifier." 2024 First International Conference on Electronics, Communication and Signal Processing (ICECSP). IEEE, 2024.

[15] Thomas, B., Bhatt, A., & Singh, S. N. (2024, August). Recognition of Facial Emotions Using CNN Architecture and FER2013. In 2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT) (Vol. 1, pp. 1-6). IEEE.

[16] Fard, Ali Pourramezan, et al. "AffectNet+: A Database for Enhancing Facial Expression Recognition with Soft-Labels." arXiv preprint arXiv:2410.22506 (2024).

[217] Abu, N. A., Abidin, Z. Z., & Darwish, S. M. (2021). A new descriptor for smile classification based on cascade classifier in unconstrained scenarios. Symmetry, 13(5), 805.

[18] Guijo-Rubio, David, et al. "Unsupervised feature-based algorithms for time series extrinsic regression." Data Mining and Knowledge Discovery (2024): 1-45.

[19] Antipona, Clarence A., et al. "An Enhancement of Haar Cascade Algorithm Applied to Face Recognition for Gate Pass Security." arXiv preprint arXiv:2411.03831 (2024).

[20] Sofronievski, Bojan, et al. "Efficient Real-time On-the-edge Facial Expression Recognition using Optomyography Smart Glasses." 2024 International Conference on Intelligent Environments (IE). IEEE, 2024.

[21] Akrom, Muhamad, Supriadi Rustad, and Hermawan Kresno Dipojono. "SMILES-based machine learning enables the prediction of corrosion inhibition capacity." MRS Communications (2024): 1-9.

[22] Patil, Nagabhushan. "An enhanced segmentation technique and improved support vector machine classifier for facial image recognition." International Journal of Intelligent Computing and Cybernetics 15.2 (2021): 302-317.

[23] Chen, Pai-Hsuen, Rong-En Fan, and Chih-Jen Lin. "A study on SMO-type decomposition methods for support

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025

vector machines." IEEE transactions on neural networks 17.4 (2024): 893-908.

[25] Mohammed, Hisham, Ben K. Daniel, and Mauro Farella. "Smile analysis in dentistry and orthodontics–a review." Journal of the Royal Society of New Zealand (2024): 1-14.

[26] Hassen. A., Abu, N. A.., & Darwish, S. M. (2022). Realistic Smile Expression Recognition Approach Using Ensemble Classifier with Enhanced Bagging. Computers, Materials & Continua, 70(2).

[27] Rajesh, T. R., Kuchipudi, L. N., Dammalapati, A., & Surendran, R. (2024, July). Attendance System Based on Facial Recognition Using OpenCV. In 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS) (pp. 1529-1535).

[28] Pradeep, V., Sumukha, B. S., Richards, G. R., & Prashant, S. P. (2024, April). Facial Emotion Detection using CNN and OpenCV. In 2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS) (pp. 1-6).

[29] Winarno, S., Nugraha, A., Muttaqin, A. N. I., Zarifa, Y., Salsabila, P. M., & Mumtaz, G. F. (2024). Comparison of ArcFace and Dlib Performance in Face Recognition with Detection Using YOLOv8. INOVTEK Polbeng-Seri Informatika, 9(2), 890-903.

[30] Dai, Y., Itai, T., Pei, G., Yan, F., Chu, Y., Jiang, X., ... & Zhao, Z. (2024). DeepFace: Deep learning-based framework to contextualize orofacial cleft-related variants during human embryonic craniofacial development. Human Genetics and Genomics Advances.

[31] Fard, A. P., Hosseini, M. M., Sweeny, T. D., & Mahoor, M. H. (2024). AffectNet+: A Database for Enhancing Facial Expression Recognition with Soft-Labels. arXiv preprint arXiv:2410.22506.

[32] Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). Affectnet: A database for facial expression, valence, and arousal computing in the wild. IEEE Transactions on Affective Computing, 10(1), 18-31.

[33] Cherian, A. K., Devipriya, S., Saoji, B. P., Mallikeswari, B., Thiagarajan, R., & Krishnamoorthy, R. (2024, April). A Robust Design of Real-Time Resilient Smile Recognition System using Hybrid Deep Learning Principles. In 2024 10th International Conference on Communication and Signal Processing (ICCSP) (pp. 592-596). IEEE.

[34] Saif, F. M., Rubayat, J., Pranto, F. S., Al Walid, M. H., & Ahammed, K. M. S. Facial Expressions Recognition Using Bezier Curve and Performance Analysis Using Ann, Naïve-Bayes and SVM Classifiers.

[35] Mattioli, M., & Cabitza, F. (2024). Not in My Face: Challenges and Ethical Considerations in Automatic Face Emotion Recognition Technology. Machine Learning and Knowledge Extraction, 6(4), 2201-2231.

DOI: https://doi.org/10.54216/JCIM.160115
Received: November 14, 2024 Revised: January 21, 2025 Accepted: February 21, 2025