Journal of Intelligent Systems and Internet of Things
Vol. 18, No. 02, PP. 130-141, 2026
DOI: https://doi.org/10.54216/JISIoT.180210

Features Extraction Improvement for Facial Expression Recognition Using HOG and Machine Learning Techniques

Dhiaa M. Abed[1], Awab Qasim Karamanj[1], Thura J. Mohammed[2] , Saja B. Attallah[1,4] , Abusnina M. Mukhtar[3,4,*]

[1]College of Biomedical Engineering, University of Technology, Baghdad, Iraq

[2] School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysia

[3]Department of Biomedical Engineering, Alasmarya Islamic University, Libya

[4]Department of Biomedical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

Emails: Dhiaa.M.Alfyadh@uotechnology.edu.iq; awab.q.abdulrasool@uotechnology.edu.iq; thurajamal@student.usm.my; 060191@uotechnology.edu.iq; m.abosninah@asmarya.edu.ly

Abstract

Facial Expression Recognition (FER) is a vital aspect of human-computer interaction with applications in healthcare, education security, and affective computing. Even with the success of deep learning, generalizability, interpretability, and efficiency of most systems, especially in uncontrolled settings, are still problematic. In this study, we propose an enhanced feature extraction technique based on Histograms of Oriented Gradient (HOG) where the central difference operator, not the conventional forward difference, used for gradient estimation. The modification enhances the accuracy of gradients, reduces truncation error, and leads to more stable facial feature descriptors. The enhanced HOG is tested on five popular datasets, CK+, JAFFE, MMI, ExpW, and AffectNet, using three traditional Machine Learning (ML) classifiers: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF). Experimental results indicate uniform accuracy enhancements across all the classifiers and datasets, with improvements spiking to 7%–10% and recall and F1-score also witnessing marked increases. In this study, RF registered the maximum accuracy, 97.94%, on CK+ and 95.48% on AffectNet, hence solidifying its stability and dependability. This study shows how well mathematical optimization works with classical ML for FER. The approach we suggest provides an easy-to-understand, small, and quick alternative to deep models, making it perfect for real-time and resource-limited applications.

Received: April 02, 2025 Revised: June 05, 2025 Accepted: July 29, 2025

Keywords: Facial Expression Recognition; Improved HOG; Central Difference Gradient; Machine Learning; Feature Extraction; Affective Computing

1. Introduction

Image recognition is the use of computers to process, scrutinize, and comprehend images in order to differentiate different patterns of targets and objects [1]. It's the primary research direction in the field of computer vision, and has an important impact on the acquisition and processing of intelligent data based on images [2]. Image recognition technology is capable of efficiently recognizing and completing the detection of specific target objects (like handwritten characters, products or faces), the classification of images and the evaluation of subjectivity, all of which are specific to the target object [3]. Facial Expression Recognition (FER) has experienced significant growth in the computer vision field over the past few decades, and has become an intriguing field of study [4]. It's hard to create a machine-based system for recognizing human expressions [5]. The objective of recognizing facial expressions is to recognize the human emotional states associated with biological traits. Facial expressions are considered the most effective tool for recognizing human emotion and intent, this is because it provides specific information that is crucial to the design of a FER system [6].

Facial expressions are the most effective and natural means of expressing oneself, communicating thoughts, and carrying out actions, all of which are useful in managing relationships and communication with others. Humans typically communicate in multiple ways, including through body language and expressive facial actions. For a more in-depth understanding of human behavior, a precise analysis and interpretation of the components of people's facial expressions that are associated with love are necessary [7]. The problem for FER is to develop accurate and efficient computational methods for automatically detecting and classifying human emotions because not all emotions are clearly expressed through distinct facial expressions [8]. Subtle or nuanced emotions, such as micro-expressions, can be difficult to detect accurately, necessitating the development of algorithms capable of detecting even the smallest emotional signals [9]. Certain facial expressions can convey multiple emotions at the same time or be ambiguous. A significant research issue is developing techniques that can handle and interpret these complex expressions correctly. In this study, the main contribution to solving these problems is the development of an algorithm with a high ability to extract facial features with high accuracy for enabling ML techniques to distinguish between converging facial expressions. Such advancements are not only relevant for academic research but also have broad applicability in real-world systems, including smart surveillance cameras, affective IoT devices, healthcare monitoring (e.g., detecting patient stress or discomfort), and human–computer interaction in intelligent tutoring or driver-assistance systems.

2. Related Work

In the last years, the problem of FER has been studied very extensively, and many methods based on ML have been proposed to increase the performance of recognition [10]–[11]. Main goals of the following researches were the increase of the accuracy of FER systems and their robustness in different conditions.

In 2021, the study [11] proposed model using a Convolutional Neural Network (CNN) that is derived from the MobileNet model. The method of transfer learning was also employed in this research to increase the performance of the model. The experimental results demonstrated that the MobileNet model with a transfer learning approach had a successful performance in the recognition task, having the highest validation accuracy of 89% and the test accuracy of 87%. The F1 score and precision value also advocated the method's reliability by achieving a high score of 87% for both metrics. Additionally, in 2021, the study [12] developed a fuzzy-optimized CNN-RNN hybrid model to solve problems related to feature vagueness and the loss of information during aggregation. It combined fuzzy logic with CNN-based feature extraction and included affine transformation and bilinear interpolation for data augmentation, and this model achieved state-of-the-art results on datasets like CK (99.22%), JAFFE (96.64%), and FER2013 (72.81%).

In 2022, the study [13] demonstrated a mental health evaluation system that is real-time and employs a lightweight CNN algorithm to recognize the behavioural psychology of patients. The proposed method utilizes a shortcut connection that facilitates a more effective propagation of the gradient through the network, the dataset FER-2013 and the laboratory-created dataset CK+ have been employed to achieve 68.93 and 96.12% accuracy in the experiment.

In 2023, a notable study utilized transfer-learning approaches using pre-trained deep networks ResNet50, VGG19, Inception V3, and MobileNet by replacing the final fully connected layers with task-specific trainable layers. When evaluated on the CK+ dataset, the system reached an average accuracy of 96%, highlighting the benefit of reusing robust feature representations from large-scale image datasets [10].

In 2024, a recent study aimed to improve the performance of FER in real, low-resolution images using deep learning techniques. This approach improved recognition accuracy by a large margin, particularly in the presence of challenges that mimic nature within an uncontrolled setting. Extensive experiments conducted provided evidence with ample support regarding the ability of the system to generalize well, reaching an accuracy level of 97.7% while keeping the process real-time. Results from manual and classical techniques set side by side with it in laboratory as well as in-field conditions confirmed their much lower performance, thus validating the superiority of the deep learning-based method proposed over it [14].

Much as there has been much improvement in the use of ML to recognize facial expressions, there are still major real-world, generalization, and interpretability limitations in the systems that have been developed so far [4], [11], [14]. Apparent variations in lighting, resolution, occlusion, and the problem of unbalanced classes of emotions tend to make most models generalized poorly. In the end, MobileNet-like lightweight architectures may provide fast inference but at the cost of a lower accuracy, especially in recognizing fine or similar emotions such as fear and surprise. Most existing systems are mere 'black boxes' and, therefore, provide little, if any, transparency or visibility, thus limiting their application in such sensitive areas as health, education, and human-computer interaction. Moreover, such systems are not suitable for small, real-time applications due to their dependence on extensive data and deep learning models. This research will address the above problems by coming up with an improved Histograms of Oriented Gradient (HOG) feature extraction method which uses central difference operators, meant to reduce errors in gradient estimation and make facial emotion discriminability better. This proposed technique will be checked on multiple benchmark datasets and ML models to try to deliver an interpretable, efficient, and robust alternative to the deep learning-heavy systems for FER.

3. Methodology

The proposed methodology for FER framework is divided into three main phases, as presented in Figure 1. In the first phase, multiple benchmark datasets should be used so that the system sees different types of facial structures, ethnicities, lighting conditions, and intensity of expressions, which supports generalizability and reduces dataset bias. The second phase should introduce a HOG feature descriptor mathematically enhanced by a central difference gradient estimation for precision in feature extraction and the maintenance of fine facial details. This will lift the cons of both conventional handcrafted methods and deep intensive models computationally. Finally, the third phase should include building an FER model pipeline using classifiers that are interpretable and efficient, particularly Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF) and evaluated the results using accuracy, precision, recall, and F1-score metrics.

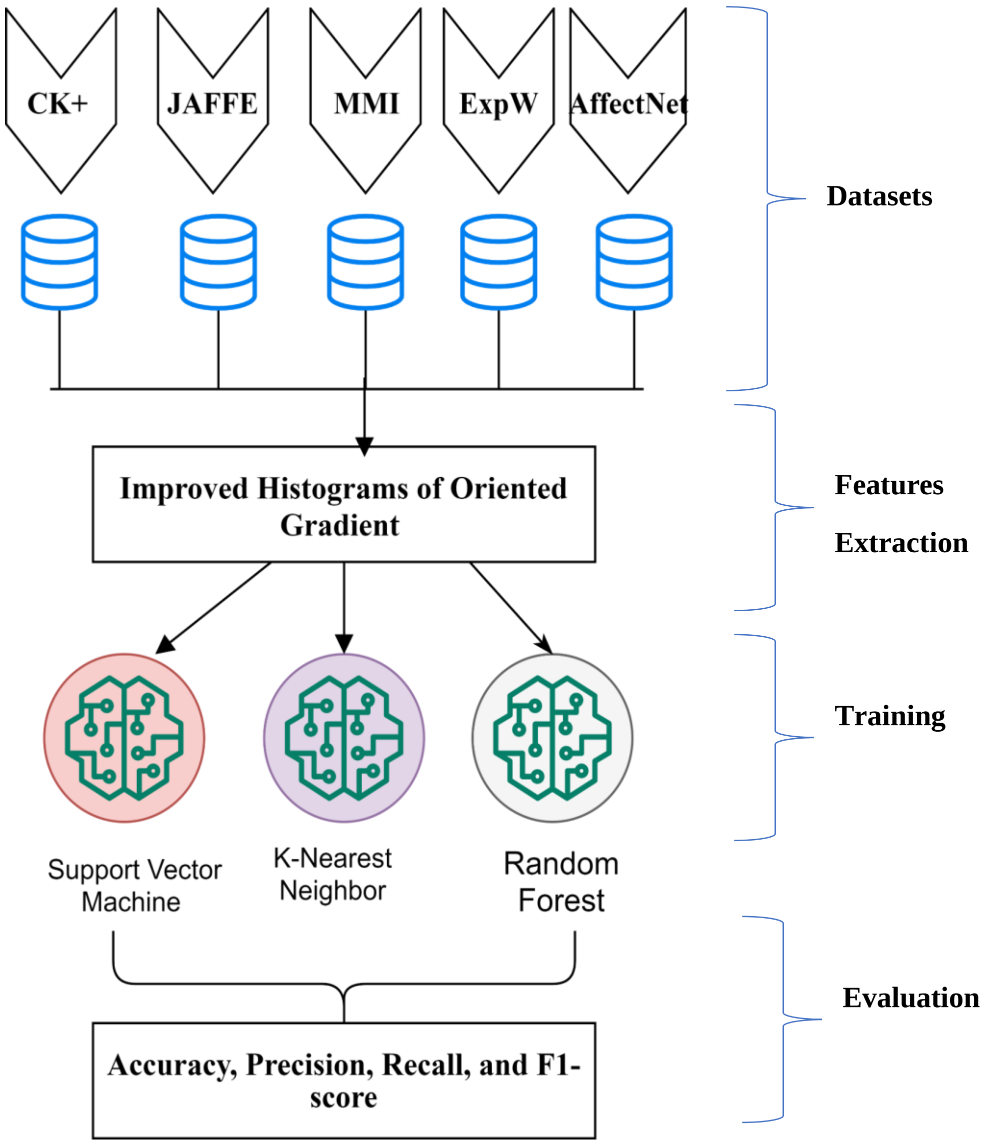

Figure 1. The Proposed Methodology Phases Diagram

3.1 Phase 1: Dataset Selection and Identification

To guarantee that the FER framework recommended is both wide and can be applied in five very well-known benchmark datasets were chosen. These datasets were picked to show both controlled laboratory environments and real, unconstrained situations [15]. This variety helps check the model's robustness over changes in facial pose, intensity of expression, lighting conditions, occlusion, gender, and ethnicity. Table 1 presents the pertinent information pertaining to the most widely utilized datasets within the domain of FER.

3.1.1 Dataset 1: Extended Cohn–Kanade Dataset (CK+)

The CK+ dataset is an extension of the CK dataset, comprising 593 video sequences and static images depicting seven distinct face expressions, including one expression of contempt and six basic emotions [16]. The static images are displayed within a laboratory setting, whereas the videos are recorded under comparable conditions. The age range of the 123 participants falls between the brackets of 18 to 30 years. The images have a resolution of 640 pixels × 480 pixels and 640 pixels × 490 pixels, with a grey value precision of 8 bits.

3.1.2 Dataset 2: AffectNet Dataset

AffectNet comprises a dataset of over one million images sourced from the Internet. These images were obtained by querying several search engines using expression-related labels. The dataset under consideration may be regarded as one of the most extensive collections of facial expressions. These expressions are categorized into two distinct emotion models, namely the dimensional model and the category model. Notably, this dataset has 450,000 images that have been manually labelled with comments, specifically targeting eight essential emotions [17].

3.1.3 Dataset 3: Expression in-the-Wild Database (ExpW)

The ExpW database comprises 91,793 facial images that were obtained by using the Google image search function. Each facial image was subject to manual analysis and categorized according to one of the seven fundamental emotion categories. The elimination of non-face images was implemented throughout the comment process [18].

3.1.4 Dataset 4: MMI Facial Expression Dataset

The MMI Facial Expression dataset comprises a collection of over 2900 videos and high-resolution images featuring 75 subjects. Each of the elements depicted in the videos has been thoroughly annotated. The dimensions of the original facial images are 720× 576 pixels [19].

3.1.5 Dataset 5: Japanese Female Facial Expressions (JAFFE)

The JAFFE dataset contains a total of 213 photos that represent seven different facial expressions, including one expression that is neutral and six basic emotions [20]. These images were obtained from 10 different Japanese females that posed for the collection. Each photograph is evaluated by six emotional words using a sample of 60 Japanese participants. The original images' dimensions are 256 × 256 pixels.

Table 1: FER Datasets

Dataset	**Resolution**	**Samples**	**Subjects**	**Access Source**	**Emotions Number**
JAFFE	256 × 256	213 Images	10	http://www.kasrl.org/jaffe.html	6 Basic Emotions + Neutral
MMI	720 × 576	740 Images and 2900 video	75	http://mmifacedb.eu/	6 Basic Emotions + Neutral + Action Units.
ExpW	Original web images	91,793 Images	N/A	http://mmlab.ie.cuhk.edu.hk/projects/socialrelation/ind	6 Basic Emotions + Neutral
AffectNet	N/A	450,000 images	N/A	http://mohammadmahoor.com/databases-codes/	6 Basic Emotions + Neutral
CK+	640 × 480, 640 × 490	593 images and video	123	http://www.pitt.edu/~emotion/ck-spread.htm	6 Basic Emotions + Neutral + contempt

By including both controlled (CK+, JAFFE, MMI) and in-the-wild (AffectNet, ExpW) datasets, this study ensures a balanced evaluation that captures both idealized and real-world conditions. The dataset diversity enables thorough testing of the proposed feature extraction method's robustness, especially in handling noisy backgrounds, facial occlusion, and non-frontal head poses.

3.2 Phase 2: Improved HOG-Based Features Extraction

The second phase of the methodology proposed is to increase feature extraction using HOG. The feature extraction stage is very critical to FER as it is one of the stages that determine the model's ability to differentiate very subtle emotional cues while maintaining robustness in challenging imaging conditions [21] . The traditional HOG is well applied to computer vision for the powerful description of local edge structures; however, its performance may be even worse in noisy or low-contrast environments since it depends on forward difference gradients [22]. To overcome these limitations, this study presents a modified HOG descriptor based on the central difference gradient operator. Improved HOG technique will be developed in extracting the features of images, where equations will be developed by use central difference instead of forward difference in gradient calculation step, and to prove why the central difference is more accurate than the forward difference for numerical differentiation, analyze the truncation error using Taylor series expansion. Let $f(x)$ be a smooth function, and expand $f(x+h)$ and $f(x-h)$ using Taylor series around x:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + O(h^4) \dots\dots(1)$$

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(x) + O(h^4) \dots\dots(2)$$

The forward difference formula is:

$$f'(x) = \frac{f(x+h) - f(x)}{h} \dots\dots\dots(3)$$

Substitute the Taylor expression with forward difference formula of $f(x+h)$:

$$f'(x) = \frac{(f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + O(h^4)) - f(x)}{h} \dots..(4)$$

Now simplify the equations:

$$f'(x) = f'(x) + \frac{h}{2}f''(x) + \frac{h^2}{6}f'''(x) + O(h^3) \dots\dots(5)$$

Compute error truncation for the forward difference formula:

$$Forward\ Difference\ Error = \frac{h}{2}f''(x) + + \frac{h^2}{6}f'''(x) + O(h^2) \dots\dots(6)$$

For the central difference formula, substitute with Taylor expression, and simplify it:

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} \dots\dots..(7)$$

$$f'(x) = \frac{(f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + O(h^4)) - (f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(x) + O(h^4))}{2h} \dots(8)$$

$$f'(x) = \left(f'(x) + \frac{h^2}{6}f'''(x)\right) \dots\dots(9)$$

Compute error truncation for the central difference formula:

$$Central\ Difference\ Error = \frac{h^2}{6}f'''(x) \ \dots\dots(10)$$

When comparison of accuracy between forward difference and central difference that the error of forward difference is proportional to h and the linear dependence on step size h, while the central difference error is proportional to h^2 and the quadratic dependence on h which makes it significantly smaller for small h. Many advantages can be obtained when applying the central difference equation, which are as follows:

A. Better Stability in Numerical Computations: Central difference is less sensitive to noise or rounding errors in numerical differentiation because of its higher-order accuracy [7].

B. Symmetry in Gradient Estimation: The central difference considers information from both sides of the point of interest ($x + h$ and $x - h$), leading to a more balanced and robust gradient estimation. In contrast, forward difference only considers one side ($x + h$), which can lead to biased estimates, especially near boundaries [23].

C. Improved Performance in PDE Solvers: Central difference is widely used in Partial Differential Equation (PDE) solvers [24]. The higher accuracy directly translates to better numerical solutions for PDEs. It allows for larger grid spacing (h) while maintaining acceptable accuracy, reducing computational cost. In this work, the central difference operator was implemented with a step size of h=1, which provided a good balance between numerical precision and computational efficiency. This choice introduces negligible overhead, adding less than 2% to the overall feature extraction runtime. Compared with forward or backward differences, the central difference method yields higher accuracy in gradient estimation, thereby improving the robustness of extracted HOG features without significantly affecting processing speed.

When apply the previous derivative on HOG can improve the gradient equations. Given a grayscale image, $I(x, y)$, gradients can be computed using central differences:

$$G_x = \frac{I(x+1, y) - I(x-1, y)}{2} \dots \dots (11)$$

$$G_y = \frac{I(x, y+1) - I(x, y-1)}{2} \dots \dots (12)$$

Central difference produces smoother, more accurate gradient maps. Forward difference results in noisier maps, especially in images with subtle intensity variations or noise [21]. Many edge detection algorithms, such as Sobel, Prewitt, and Canny, rely on gradient calculations. Using central difference results in sharper and more consistent edges compared to forward difference, which can blur edges or miss weak ones [25], [26]. For example, central difference enhances the accuracy of detecting edges in low-contrast regions, improving performance in images with subtle texture variations. In additional noisy images, central difference is less sensitive to random fluctuations in pixel intensity because it takes information from both sides of a pixel into account. Tasks such of emotion face recognition rely on precise edge and gradient computations. Central difference enhances these computations, improving the reliability of machine vision systems [27]

In this study, improve the standard HOG (Histogram of Oriented Gradients) feature extraction by using a central difference operator instead of the traditional forward difference. Unlike Sobel or Prewitt edge detectors, which compute image gradients using fixed convolution kernels that emphasize edge intensity and direction, the central difference HOG calculates the gradient at each pixel using the average change across neighboring pixels, providing a more precise and smoother estimation of intensity changes. This approach reduces errors that can arise from abrupt changes or noise in the image and allows HOG to capture subtle facial features and fine variations in expression that Sobel or Prewitt detectors may overlook. In short: Sobel/Prewitt: Focus on general edge detection with predefined kernels; sensitive to noise and may lose fine details. Central Difference HOG: Computes gradient using pixel differences on both sides, leading to more accurate and detailed gradient maps.

3.3 Phase 3: FER Model Development and Evaluation

The final phase of the proposed framework is to make and check ML models for spotting facial expressions. Use the better HOG feature descriptions made in Phase 2. This step tries to see the real power of the taken features and how well typical classifiers can put facial expressions into groups with many different sets of data.

3.3.1 Classifier Descriptions

A majority of deep learning models require large datasets, high computational resources, and extensive training time. Therefore, this study chooses classical ML classifiers SVM, KNN, and RF were chosen due to their good performance in prior FER research, and for their capacity to work well in the generalization of structured such as those that will be derived from the enhanced HOG technique. SVM is a powerful supervised learning model used for classification tasks. It works by identifying a hyperplane that maximally separates different emotion classes in a high-dimensional feature space [17] , as presented in Figure 2. In this study, a Radial Basis Function (RBF) kernel is used to handle potential non-linear boundaries between emotions.

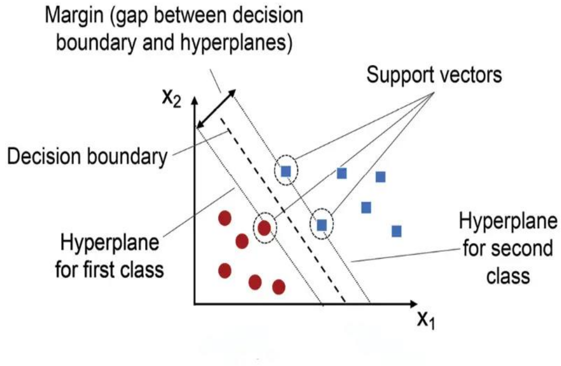

Figure 2. SVM Work Idea

KNN being non-parametric, labels of class of test sample are determined by majority labels of class of its k-nearest training samples in the feature space (Figure 3). Since it is simple and depends on distance measures, KNN works well when the descriptors of features are compact and clear such as in HOG features.

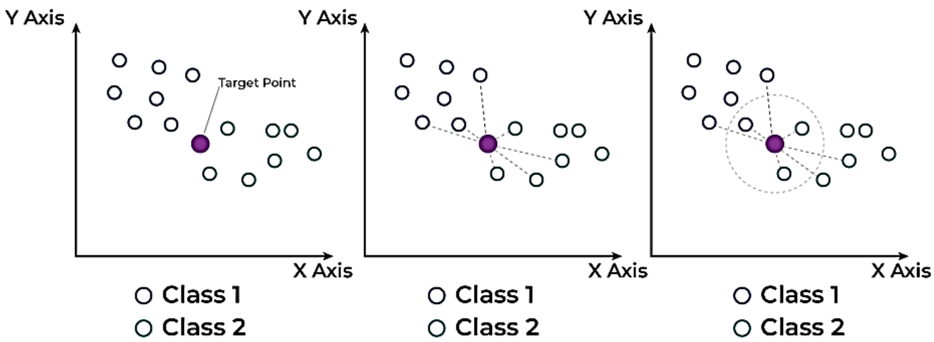

Figure 3. KNN Work Idea

RF is a learning method that is made up of many decision trees, as shown in Figure 4. It gives good results, does not become too attached to the training data, and shows strength in front of dirty data. Using bagging and feature randomness lets, it show strong aspects of how different emotions relate and interact, even when they are complex.

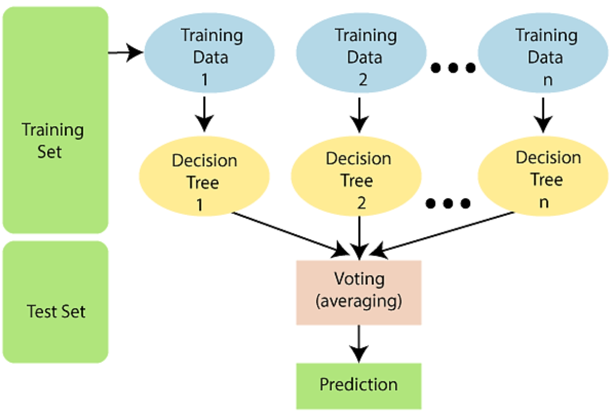

Figure 4. RF Work Idea

3.3.2 Model Training and Validation

Each classifier was trained independently on the HOG feature vectors extracted from the five datasets: CK+, JAFFE, MMI, ExpW, and AffectNet. To check that no random errors occurred in the evaluation, the training process used a stratified 10-fold cross-validation procedure. This also helps in controlling the performance estimation, making it strong against different emotion class distributions. Both AffectNet and ExpW datasets exhibit uneven emotion distributions, where certain emotions (e.g., “happy” or “neutral”) dominate over others (e.g., “disgust” or “fear”). To mitigate this issue, we adopted a class-weighting strategy during training, assigning higher weights to underrepresented emotion categories in the loss function. This approach reduced bias toward majority classes and resulted in improved recognition of minority classes, thereby enhancing the overall fairness and robustness of the system. Model performance is evaluated through four standard and effective evaluation metrics were used: Accuracy, Precision, Recall, and F1-score. Table 2 provides detailed formulations for these metrics.

Table 2: The Employed Performance Metrics

Evaluation metrics	Mathematical Equation	Explanation
Accuracy	$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$	Measures overall accuracy in classifying facial expressions under all seven-emotion categories; in other words, how well the model performs on both controlled and in-the-wild datasets.
Precision	$Precision = \frac{TP}{TP + FP}$	It is the capability of the model to rightly identify a particular emotion without mistakenly labelling other expressions as that emotion; high precision results in fewer false detections between visually similar expressions, for example, fear versus surprise.
Recall (Sensitivity)	$Recall = \frac{TP}{TP + FN}$	Indicates the degree to which the model detects all actual instances of an emotion. High recall ensures that true expressions of emotion are not missed, especially in real-time or assistive applications.
F1-score	$F1\text{-score} = \frac{2 * TP}{2 * TP + FP + FN}$	Provides a balanced measure of the model’s accuracy in both detecting and not missing emotional classes. It is especially useful when dealing with class imbalance in datasets like FER2013 or JAFFE.

TP: True Positive, TN: True Negative, FP: False Positive, TN: True Negative, FN: False Negative

4. Results and Discussion

This section gives a full analysis of the planned FER system using the enhanced HOG descriptor over five benchmark sets and three classical ML schemes: SVM, KNN, and RF. Figure of merits used is accuracy, precision, recall, and F1-score; comparisons made between standard HOG and proposed improved HOG approach. The algorithms will be taken separately in subsections below.

4.1 SVM with Improved HOG

The SVM classifier is seen to have major enhancements in its performance when made use of improved HOG, with all five datasets showing every evaluation metric to be improved. For example, in the CK+ dataset (Table 3), the accuracy goes to 96.15%, and 0.8908 to 0.9636 in the F1-score. The AffectNet dataset has recalled it to 0.9815, showing how well the model can learn increased subtle facial cues from real, unconstrained images. These results verify that a better representation of features, based on gradients, allows the SVM to create more accurate decision boundaries between emotional categories.

Table 3: SVM-Improved HOG Results

Dataset	HOG				Improved HOG			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
JAFFE	0.8400	0.8620	0.8620	0.8620	0.9346	0.9107	0.9623	0.9358
MMI	0.8065	0.8500	0.7727	0.8095	0.8929	0.8793	0.9107	0.8947
ExpW	0.8197	0.8226	0.8226	0.8226	0.9009	0.8644	0.9444	0.9027
AffectNet	0.8264	0.8669	0.8030	0.8346	0.9174	0.8689	0.9815	0.9217
CK+	0.8850	0.8983	0.8833	0.8908	0.9615	0.9815	0.9464	0.9636

4.2 KNN with Improved HOG

It is simple algorithm: greatly bene-fit from improved HOG features. In JA-FFE dataset, accuracy rose to 86.49%, and recall to 95.95%. For ExpW dataset, accuracy better by more than 10 percentage points. These improve highlight the advantage that the improved HOG descriptor provides KNN with more dimensional local features. In addition, the features help over-come the typical challenges of KNN in small datasets and subtle expressions. Table 4 presented the results of using KNN model with the improved HOG technique.

Table 4: KNN-Improved HOG Results

Dataset	HOG				Improved HOG			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
JAFFE	0.7576	0.7571	0.7794	0.7681	0.8649	0.8068	0.9595	0.8765
MMI	0.7442	0.7553	0.7717	0.7634	0.8421	0.8659	0.8452	0.8554
ExpW	0.7574	0.7717	0.7602	0.7760	0.8590	0.8987	0.8452	0.8712
AffectNet	0.7831	0.7604	0.8488	0.8022	0.8610	0.8588	0.8902	0.8743
CK+	0.8176	0.8588	0.8111	0.8343	0.8904	0.9125	0.8902	0.9012

4.3 RF with Improved HOG

RF consistently achieved the highest scores among the three classifiers. On the CK+ dataset, accuracy jumped from 89.62% to 97.94%, and the F1-score improved from 0.9000 to 0.9802. Significant gains were also observed on the ExpW and AffectNet datasets, with accuracy increases of over 8% and F1-score improvements of over 6% (Table 5). The ensemble mechanism of RF is particularly effective when provided with high quality, noise-resilient features, precisely what the improved HOG delivers through better gradient computation.

Table 5: RF-Improved HOG Results

Dataset	HOG				Improved HOG			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
JAFFE	0.8584	0.8655	0.8583	0.8620	0.9479	0.9626	0.9364	0.9493
MMI	0.8475	0.8655	0.8374	0.8512	0.9174	0.8655	0.9810	0.9196
ExpW	0.8507	0.8584	0.8509	0.8546	0.9353	0.9417	0.9327	0.9372
AffectNet	0.8664	0.9327	0.8151	0.8700	0.9548	0.9340	0.9802	0.9565
CK+	0.8962	0.8840	0.9167	0.9000	0.9794	0.9706	0.9900	0.9802

4.4 Comparative Analysis and Summary of Improvements

In all experiments, the new HOG descriptor proved to be better than the baseline in all evaluation setups. The gains in accuracy that were registered with the new HOG went from 7% to 10%, most major improvements being visible in wild datasets (AffectNet and ExpW) as presented in Figure 5. These are very challenging due to variability from multiple factors in facial expression, with performance improvement here being most welcome. Recall has also improved impressively in the detection of more challenging emotions, such as fear and surprise. For instance, in the detection of these two hard-to-recognize expressions, Recall has reached up to 0.9815 in AffectNet with SVM, and up to 0.9900 in CK+ with RF. Such improvements are required for applications in real-world FER; otherwise, they may fail due to minor unnoticeable emotional expressions. Moreover, there was consistent improvement in the F1-score of all classifiers. It means that not only one part (precision or recall) is improved but also the balance is enhanced between both. In other words, this improved HOG will enhance the accuracy of the system and reduce false positives as well as false negatives. In simpler terms, these findings congregate to say that employing the HOG descriptor with the computation of the gradient at the central difference would make much stronger the ability of the system to extract precise, robust, and generalizable facial features, leading to more reliable and interpretable models for recognizing facial expressions.

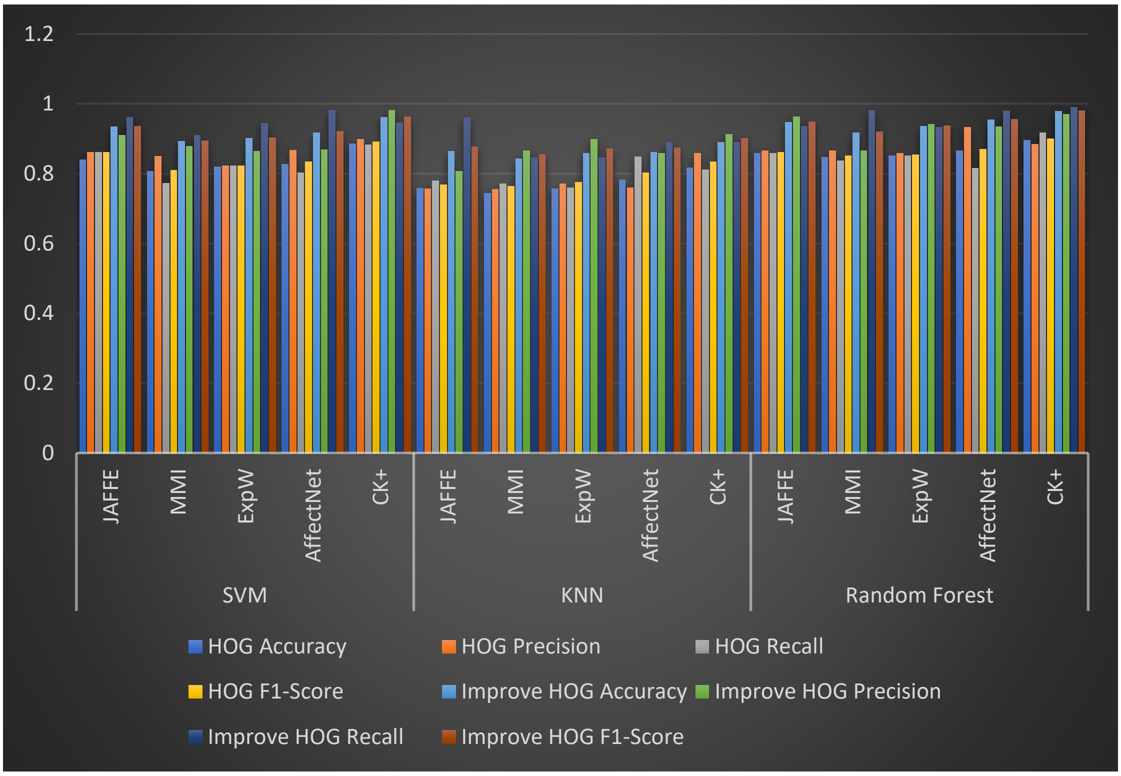

Figure 5. Comparative Performance of Standard HOG vs. Improved HOG across Five FER Datasets Using SVM, KNN, and RF Classifiers

5. Conclusions

This study introduced and tested an improved face look recognition system based on a changed HOG feature extraction methodology. By using a traditional forward difference operator with a central one, for the calculation of gradients, the improved HOG method allowed for smoother and more accurate gradient maps, and significantly reduced error. This mathematical refinement allowed the descriptor to better capture facial details critical for the distinction of overlapping and subtle emotions. A series of comprehensive experiments were executed with three classical classifiers, SVM, KNN, and RF, on five benchmark datasets: CK+, JAFFE, MMI, ExpW, and AffectNet. Results proved that an enhanced version of HOG performed consistently and significantly better in terms of accuracy, precision, recall, and F1-score with all classifiers and all datasets. The highest accuracy was noted with the RF, reaching 97.94% on CK+ and 95.48% on AffectNet, hence further proving robustness and generalizability due to enhancements made in the proposed methodology. Comparative analysis further proved that the advanced

HOG performs best in real-life, unconstrained data such as ExpW and AffectNet, where facial expressions are variable, noisy, and occluded. Moreover, gains in recall and F1-score inform that the system is not only more sensitive to true emotional states but also more balanced towards reducing both false positives and false negatives. This work confirms that competitive results against deep learning approaches can be delivered in low resource or real-time settings when handcrafted features are combined with optimized mathematical operators. The improved HOG technique is a lightweight, interpretable, and efficient alternative for FER, making it ideal for deployment in embedded systems, mobile applications, and edge AI scenarios. Proposed FER system is well suited for deployment in IoT and smart environments. Specifically, the lightweight nature of the HOG-based feature extraction combined with central difference enhancements enables execution on edge devices with limited computational resources. This makes the framework practical for real-time monitoring scenarios, such as healthcare (continuous patient emotion tracking), smart classrooms (engagement monitoring), and surveillance (suspicious behavior detection). The system's modular design allows integration with IoT gateways and cloud platforms, ensuring low-latency emotion recognition while maintaining privacy by processing sensitive data locally at the edge.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] M. Shokry, A. I. Awad, M. K. Abd-Ellah, and A. A. M. Khalaf, "Systematic survey of advanced metering infrastructure security: Vulnerabilities, attacks, countermeasures, and future vision," *Future Gener. Comput. Syst.*, vol. 136, pp. 358–377, 2022, doi: 10.1016/j.future.2022.06.013.

[2] M. Sajjad et al., "Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities," *Future Gener. Comput. Syst.*, vol. 108, pp. 995–1007, 2020, doi: 10.1016/j.future.2017.11.013.

[3] X. Lv, M. Su, and Z. Wang, "Application of Face Recognition Method Under Deep Learning Algorithm in Embedded Systems," *Microprocess. Microsyst.*, p. 104034, 2021, doi: 10.1016/j.micpro.2021.104034.

[4] S. Barra, S. Hossain, C. Pero, and S. Umer, "A Facial Expression Recognition Approach for Social IoT Frameworks," *Big Data Res.*, vol. 30, p. 100353, 2022, doi: 10.1016/j.bdr.2022.100353.

[5] M. A. P. Chamikara, P. Bertok, I. Khalil, D. Liu, and S. Camtepe, "Privacy Preserving Face Recognition Utilizing Differential Privacy," *Comput. Secur.*, vol. 97, p. 101951, 2020, doi: 10.1016/j.cose.2020.101951.

[6] X. Zhao, J. Li, W. Liu, J. Zhang, and Y. Li, "Design of the sleeping aid system based on face recognition," *Ad Hoc Netw.*, vol. 99, p. 102070, 2020, doi: 10.1016/j.adhoc.2019.102070.

[7] S. Demir, S. Key, T. Tuncer, and S. Dogan, "An exemplar pyramid feature extraction based humerus fracture classification method," *Med. Hypotheses*, vol. 140, p. 109663, 2020, doi: 10.1016/j.mehy.2020.109663.

[8] A. Amirkhani, M. P. Karimi, and A. Banitalebi-Dehkordi, "A survey on adversarial attacks and defenses for object detection and their applications in autonomous vehicles," *Vis. Comput.*, vol. 39, no. 11, pp. 5293–5307, 2023, doi: 10.1007/s00371-022-02660-6.

[9] S. M. Imran, S. M. M. Rahman, and D. Hatzinakos, "Differential components of discriminative 2D Gaussian–Hermite moments for recognition of facial expressions," *Pattern Recognit.*, vol. 56, pp. 100–115, 2016, doi: 10.1016/j.patcog.2016.03.006.

[10] M. K. Chowdary, T. N. Nguyen, and D. J. Hemanth, "Deep learning-based facial emotion recognition for human–computer interaction applications," *Neural Comput. Appl.*, vol. 35, no. 32, pp. 23311–23328, 2023, doi: 10.1007/s00521-021-06012-8.

[11] R. Sadik, S. Anwar, and L. Reza, "AutismNet: Recognition of Autism Spectrum Disorder from Facial Expressions using MobileNet Architecture," in *Proc. Int. J. Adv. Trends Comput. Sci. Eng.*, 2021, doi: 10.30534/ijatcse/2021/471012021.

[12] D. Zhang and Q. Tian, "A Novel Fuzzy Optimized CNN-RNN Method for Facial Expression Recognition," 2021, *arXiv: 2112.14513*.

[13] T. Podder, D. Bhattacharya, and A. Majumdar, "Dew Computing-Inspired Mental Health Monitoring System Framework Powered by a Lightweight CNN," in *Disruptive Technologies for Big Data and Cloud Applications*. Singapore: Springer, 2022, pp. 309–319.

[14] H. B. U. Haq, W. Akram, M. N. Irshad, A. Kosar, and M. Abid, "Enhanced Real-Time Facial Expression Recognition Using Deep Learning," *Acadlore Trans. AI Mach. Learn.*, vol. 3, no. 1, pp. 24–35, 2024, doi: 10.56578/ataiml030103.

[15] M. Ahmed, S. Byreddy, A. Nutakki, L. F. Sikos, and P. Haskell-Dowland, "ECU-IoHT: A dataset for analyzing cyberattacks in Internet of Health Things," *Ad Hoc Netw.*, vol. 122, p. 102621, 2021, doi: 10.1016/j.adhoc.2021.102621.

[16] M. M. Bapat, S. M. Mali, and V. Karad, "Database Development and Recognition of Facial Expression using Deep Learning," *Int. J. Comput.*, vol. 23, no. 4, 2023, doi: 10.47839/ijc.23.4.3760.

[17] L. Boussaad and A. Boucetta, "An effective component-based age-invariant face recognition using Discriminant Correlation Analysis," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 5, pp. 1739–1747, 2022, doi: 10.1016/j.jksuci.2020.08.009.

[18] I. S. Badr et al., "Cancellable face recognition based on fractional-order Lorenz chaotic system and Haar wavelet fusion," *Digit. Signal Process.*, vol. 116, p. 103103, 2021, doi: 10.1016/j.dsp.2021.103103.

[19] I. S. Na et al., "FacialNet: facial emotion recognition for mental health analysis using UNet segmentation with transfer learning model," *Front. Comput. Neurosci.*, vol. 18, 2024, doi: 10.3389/fncom.2024.1485121.

[20] S. Minaee, M. Minaei, and A. Abdolrashidi, "Deep-emotion: Facial expression recognition using attentional convolutional network," *Sensors*, vol. 21, no. 9, p. 3046, 2021, doi: 10.3390/s21093046.

[21] S. Karanwal, "Robust face descriptor in unconstrained environments," *Expert Syst. Appl.*, vol. 247, p. 123302, 2024, doi: 10.1016/j.eswa.2024.123302.

[22] A. K. Sharma et al., "HOG transformation based feature extraction framework in modified Resnet50 model for brain tumor detection," *Biomed. Signal Process. Control*, vol. 84, p. 104737, 2023, doi: 10.1016/j.bspc.2023.104737.

[23] Z. Ding et al., "Multi-resolution 3D-HOG feature learning method for Alzheimer's Disease diagnosis," *Comput. Methods Programs Biomed.*, vol. 214, p. 106574, 2022, doi: 10.1016/j.cmpb.2021.106574.

[24] M. Mebarkia, A. Meraoumia, L. Houam, and S. Khemaissia, "X-ray image analysis for osteoporosis diagnosis: From shallow to deep analysis," *Displays*, vol. 76, p. 102343, 2023, doi: 10.1016/j.displa.2022.102343.

[25] A. Sharma, D. P. Yadav, H. Garg, M. Kumar, B. Sharma, and D. Koundal, "Bone Cancer Detection Using Feature Extraction Based Machine Learning Model," *Comput. Math. Methods Med.*, vol. 2021, p. 7433186, 2021, doi: 10.1155/2021/7433186.

[26] S. Ramachandra and S. Ramachandran, "Region specific and subimage based neighbour gradient feature extraction for robust periocular recognition," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, pp. 7961–7973, 2022, doi: 10.1016/j.jksuci.2022.07.013.

[27] D. M. Abed, S. Abdul-Rahman, and S. Mutalib, "Dental segmentation via enhanced YOLOv8 and image processing techniques," *Mesopotamian J. CyberSecurity*, vol. 4, no. 3, pp. 189–202, 2024, doi: 10.58496/MJCS/2024/022.