Multi-Class Image Classification Using Feature Fusion and Kinetic Gas Molecule Optimization

Venkataravana Kadirappa
School of Computer Science and Engineering
Presidency University, India
venkataravana.nayak@presidencyuniveristy.in

Praveen Mokshanatha
Department of Research and Innovation
Srinivas University, India
bm.praveen@yahoo.co.in

Abstract: *To manage the incremental change in the size of image data in various classification scenarios, requires methods based on appropriate visual content in a sensible measure of time. To make real-time decisions classification necessitates amended strategies, methodologies for processing, analysis, and classify on a large scale. This paper proposes an Image Classification method using feature fusion optimized by Kinetic Gas Molecule Optimization (KGMO). Multiple features such as Histogram of oriented Gradient (HoG), Compound Local Binary Pattern (CLBP), Color and Statistical features are considered. The extracted features are fused using KGMO algorithm. The proposed method utilizes KNN to find the distance between the feature vectors and is tested by comparing it with previous state-of-the-art techniques and it is found that it comparatively outperforms the state-of-the-art techniques.*

Keywords: *Accuracy, convolution neural network, feature extraction, feature fusion, image classification, kinetic gas molecule optimization, k-nearest neighbors.*

Received March 7, 2025; accepted July 31, 2025
https://doi.org/10.34028/iajit/23/1/10

1. Introduction

Image classification problems are an important part of a digital image analysis and are significant tasks of an artificial intelligence [22]. Due to the usage of internet, social media networks, and digital technologies the production of image data is increasing exponentially. Thus, recognising the correct categories of input images become difficult because of input images differ in geometrical properties and may contain background clutter [24]. To solve various types of classification problems, various simple low level features selection approaches have been designed, introduced efficient algorithms and models such as machine learning [6], deep learning, Support Vector Machine (SVM) [9], and the Bag-Of-Words (BOW) [1].

In the digital image processing, the representation, sharing and interpretation of images are increasing in real time in various scenarios. To search, retrieve and classify specific image on the basis of its visual contents in a specific situation necessitates content based classification frameworks [17]. The retrieval and classification over the comment words makes explanation complex, tedious and furthermore requires tremendous work to physically comment on the images for the large size image databases and it will not consider semantic substance.

Sketch based image classification use sketches as the input and retrieve and classify part of image, images dependent on the sketches [15]. The approach identifies key shapes rather than distinguishing over which nearby descriptors are registered and represents the structure of an image content with the combination of the bag of feature method provides an enhanced result. The method is text based but works for the large scale dataset. The search breaks down image contents instead of metadata e.g., labels, watchwords or depictions to the image. The content represents color, texture, shape or details of data obtained from the image. A framework that retrieve and classify images based on their image content provide improved outcome.

Data fusion is the way toward coordinating numerous data sources to deliver progressively predictable, precise, and helpful data than that given by any individual data source. Based on the level of handling at which fusion takes place, the procedures of data fusion are categorized in to low, intermediate and high. The Low-level data fusion combine few sources of raw-data in creating new raw-data. The combined data is much useful and synthetic compared to original information sources [14].

The pre-processing phase help to obtain robust and discriminative features with the descriptors and optimizing these using Kinetic Gas Molecule Optimization (KGMO) optimization find best way to combine the features and this feature fusion intern reduces the dimensionality and also accomplish desired result. With Genetic algorithm, the traditional, shallow layer and pre-trained Convolution Neural Network (CNN) features are fused to accumulation of errors and dimension of fused feature vectors in the existing system [25]. In the proposed method, considered the texture, color and statistical features in the feature extraction method, used KGMO algorithm for optimization and fused the features, the similarity with reference to the image dataset and the input image is computed using KNN to increase the classification accuracy and decrease time.

Objective: optimize the extracted visual features with KGMO technique to enhance classification accuracy of matching images from the database and reduce time with machine learning approach.

- Contribution: the proposed method contributions are as follows:

  1. Accomplished model classification accuracy of 4.63%, 16.7%, 11.29% and 4.41% more compared to the existing methods [5, 8, 19, 26].
  2. Obtained the hybrid optimal feature combination that increase interpretability of the image representation and helps in the process of reducing the processing complexity and time.
  3. The model gives a classification accuracy of 98% and 96% for the small-scale and large-scale datasets respectively considering less percentage of the hybrid optimal extracted feature combination. Due to this it eliminates the requirement of using feature selection methods.

The paper is organized as follows: section 2 delivers a short explanation on related works of image classification techniques. section 3 explains proposed method with implementation details. section 4 describes performance analysis and section 5 present conclusions.

2. Related Work

Different methods used for the image classification task are based on:

1. Low-level feature representation, considers image as group of low level features such as color, size, shape and texture etc.
2. Mid-level visual feature construction. Image classification in distinct domains has several benefits and applications. Table 1 show various relevant research works have been presented in this section.

Rao and Mahantesh [19] proposed a hybrid image classification model using ML and DL methods. Obtained color, texture, shape features with the corresponding descriptors and the high level features with the VGG16 convolution base and are fused. The obtained features trained using distinct ML algorithms and tested the model performance. Among the algorithms, the logistic regression gives 91.37% accuracy due to the decrease of model variance. For the feature fusion, the hybrid model gives an accuracy of 88.52%. The implemented model fails in giving stable result.

Hwang *et al*., [8] introduced a wood knot classification framework using texture and local features. The representative features are obtained using the corresponding GLCM, LBP and HoG, SIFT descriptors, trained on the ML classifiers and evaluated the model performance. Suggested texture features gives better performance than the local features. Dataset not cover variation in the defect parameters of the samples. The performance is low, so not appropriate for developing automated models to the industrial application. The performance can be enhanced with automated image acquisition modules with deep learning techniques.

Yang *et al*., [26] proposed an image classification model using local and global features. Extracted shape, texture and global features using corresponding descriptors and constructed code book using SIFT-local and GIST-global features. With Principal Component Analysis (PCA) reduced obtained and fused features dimensionality and used SVM classifier for classifying the Caltech-101 image dataset. The discriminative ability of the features is less and the model is not scalable.

Das *et al*., [3] designed a framework using low level and deep features to classify images. Extracted color, and shape features with traditional descriptors and deep features with the variants of CNN from the scaled image dataset. The features are horizontally fused and used SVM classifier for classification. The features obtained are not discriminative in nature hence the performance is less.

Table 1. Review of classification techniques and their focus.

Method	Advantage and disadvantage	Dataset	Accuracy (%)
Color+Texture+Shape+VGG19 features+ML algorithms [19]	• Feature (low level and deep features) fusion is used. • Result is not stable.	Caltech101	91.37
GLCM+LBP and HoG+SIFT+DSIFT+KNN [8]	• Performance is low. • Not appropriate to industrial application.	Wood knot	79.30
Gobal + Texture + Shape + PCA + SVM [26]	• Features discriminative ability is less. • Model not scalable.	Caltech101	84.71
HOG + CH + VGG_19 + ResNet_50 + SVM [3]	• Features are not discriminative.	Corel 5k	88.77
CNN + HoG + LBP + SVM [5]	• Works well for small scale datasets. • Model interpretability is less.	F-MNIST	91.59

Greeshma and Gripsy [5] developed a classification framework using low level features and deep features. In the first phase, extracted low level features using corresponding HoG and LBP descriptors and deep features using convolution base of the CNN. In the second phase, the SVM and the CNN classifiers are used for the image classification. The model works well for small scale datasets and its interpretability is less.

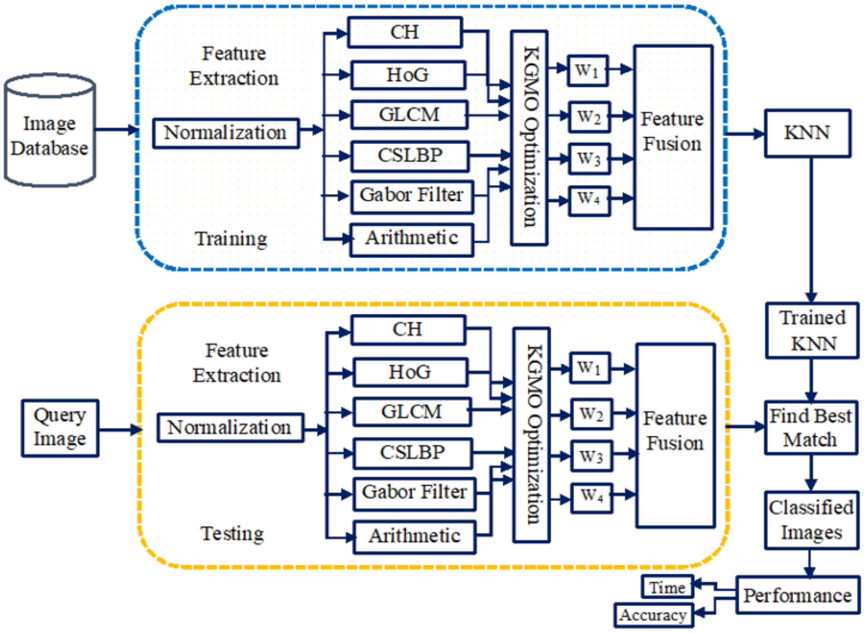

Figure 1. Proposed system (IC-FF-KGMO) framework.

This related work has carried out to understand classification performance with respect to individual feature extraction methods and fusion based approaches.

3. Proposed System

3.1. Problem Statement and Objective

The problem statement is to develop an efficient image classification model to classify images through derived image features from a large-scale database. And improve the classification model accuracy.

3.2. Architecture

The architecture of an KNN based image classification model (IC-FF-KGMO) is shown in Figure 1. It consists four phases namely:

1. Preprocessing.
2. Feature extraction.
3. Training.
4. Testing.

IC-FF-KGMO uses hybrid feature extraction technique and KGMO optimization. Initially, images from the database are resized and textural, color and arithmetic features are extracted and stored in the feature vector for performing training process. The feature vectors are fused by multiplying with a weight which is obtained after applying optimization to the extracted feature vectors. After feature fusion, used minimum distance method on database images to find best match. The KGMO parameters explore high dimensional search space and find the best way to combine the features in optimization problems that contain multiple constraints which helps to take fewer iterations to converge to global minima in the accomplishment of reaching higher stable accuracy. The parameters enhance discriminative power in the process of hybrid feature extraction, optimize parameters for fusion, and enhance image visual quality.

3.2.1. Preprocessing

The images are resized and normalized to perform feature extraction. Different images from the database will have different sizes. To give the uniformity, the images are resized to constant size.

3.2.2. Feature Extraction

In IC-FF-KGMO, multiple features are extracted from the images. Among multiple features, color features are significant over other features. By considering the color image primary parameter, the HoG and color histogram are extracted.

- **Color feature extraction**
- **Histogram oriented Gradient (HoG)**: the magnitude, and direction features are obtained to distinguish two images [4]. The extracted HoG features are discriminative and represents dependably.
- **Color Histogram (CH) analysis method**: in proposed framework, the dataset images are resized to 256∗256, the image color space are discretizing into n distinctive colors. As each image considered is 8-bit image, thus there are 256 distinct bins of colors. The summation of color bins is utilized for the computation of mean, standard deviation. A CH is a vector $<ch_1, ch_2, ..., ch_i>$ in which every bucket h_j consists count of pixels color j in the image [27]. The features of image histogram extracted in RGB color space, the pixels count of 256 distinct bins are stored in the matrix.

- **Gabor filtering**: linear orientation-sensitive filter, used for feature extraction, edge detection, texture analysis. Any frequency content in any direction in an image in a region of analysis [7].

$$g(x,y,\sigma,\theta,\lambda,\gamma,\psi) = exp\left[-\frac{(x^2+\gamma^2y^2)}{2\theta^2}\right]exp\left[i\left(2\phi\left(\frac{x}{\lambda}\right)+\psi\right)\right] \quad (1)$$

Where, x, y represents size of filter, λ is wavelength, γ is aspect ratio, ψ is phase offset, and σ is standard deviation.

- **Grey Level Co-Occurrence Matrix (GLCM)**: the images comprise of numerous pixels and each pixel consists its own intensity level. GLCM [13] is used for analysing texture by tabulating pixels with various intensity level. Here, obtained a Gray-Level Spatial Dependence Matrix (GLSDM) using the two steps:

1. First-order statistical textural analysis is used in extraction of the feature information and also at random position, measured gray-level frequencies without considering pixel neighbors.
2. The features are extracted by performing the second-order textural analysis with consideration of neighbor pixels.

It is a 2D histogram that gives spatial relation among the pixels of different gray-level values. where in $(p, q)^{th}$ elements, the frequency of occurring event p with q, is a distance function D=1, at 90◦ in vertical, 0◦ in horizontal, 45◦ with positive diagonal, and 135◦ in negative diagonal and gray-scales p, q and computes pixel frequency with intensity p, take place relative to other pixel p at a definite distance D and orientation.

Through choosing operative quantitative level for initial stages, the GLCM improve accuracy level. The textural features i.e., contrast, correlation, homogeneity, variance, energy, and entropy are obtained through HL and LL sub-bands of first 4-levels of wavelet decomposition.

$$Contrast = \sum_{x=0}^{m-1}\sum_{y=0}^{n-1}(x-y)^2ff(x,y) \quad (2)$$

$$Energy = \sqrt{\sum_{p=0}^{i-1}\sum_{q=0}^{j-1}f^2(p,q)} \quad (3)$$

$$Correlation = \frac{\sum_{p=0}^{i-1}\sum_{q=0}^{j-1}(p,q)f(p,q)-M_pM_q}{\sigma p\sigma q} \quad (4)$$

$$Homogenity = \sum_{p=0}^{i-1}\sum_{q=0}^{j-1}\frac{1}{1+(p-q)^2}f(p,q) \quad (5)$$

$$Entropy = \sum_{p=0}^{i-1}\sum_{q=0}^{j-1}f(p,q)log_2f(p,q) \quad (6)$$

- **Texture Features**
- **Local Binary Patterns (LBP)**: the image R, G, and B components individually extracted and they are stored in a vector. Then LBP operation is performed on these components independently. The pixel at the center is compared with its every neighbor pixel in a 3∗3 matrix. In estimating central pixel value of a cell, the central pixel is less than or equivalent to neighbor pixel then “1” is considered in a corresponding cell pixel value otherwise generally “0” is considered. The order in a cell from upper-left corner is considered in order to determine a binary number [23].

$$LBP_{P,R}(x_c,y_c) = \sum_{p=0}^{p-1}s(i_n,i_c)2^P \quad (7)$$

$$B(x) = \begin{cases}1 & \text{if} x \geq 0\\ 0 & \text{if} x < 0\end{cases} \quad (8)$$

- **Compound Local Binary Pattern (CLBP)**: the LBP takes in to account only sign and discard magnitude in the computing center and nieghbor pixel grey values. So the binary code generated is not appropriate. The CLBP use 2-bits, one each for both sign and magnitude values in generating appropriate binary code [21]. The codes produced by the LBP operator are inconsistent in a local neighborhood due to discarding pixel magnitude at the center and their neighbor gray-values in the image matrix. For e.g., the (11111111) 8-bit uniform LBP code related to a flat-area at the center-pixel in not correct. The LBP often fails in generating appropriate binary code due to its consideration of only the sign of the difference among two gray-values. Thus considered CLBP, an extension of LBP and it uses 2-bits for every neighbor pixel in the representation of sign and magnitude difference between the pixel at the center and their neighboring gray values. For the gray-values of a local neighborhood having P-neighbors, CLBP assigns 2P-bit code to center pixel. The first-bit indicates the sign of pixel difference between the pixel at the center and their corresponding neighboring gray-values similar to pattern of LBP and second-bit to encode the magnitude of difference with a threshold value.

$$B(i_n,i_c) = \begin{cases}00 & if i_n = i_c < 0, |i_n - i_c| \leq M_{avg}\\ 01 & if i_n = i_c < 0, |i_n - i_c| > M_{avg}\\ 10 & if i_n = i_c \geq 0, |i_n - i_c| \leq M_{avg}\\ 11 & otherwise\end{cases} \quad (9)$$

Where, i_c is center pixel value, i_n is neighbor pixel value and M_{avg} is average magnitude i.e., difference in the local neighborhood among i_n and i_c. In a local region, considers pixel magnitude through the comparison of central pixel with its neighbors. It increases texture description, variation in image orientation, less sensitive to variations in lighting conditions.

- **Arithmetic features (Malegori *et al*. [13])**
- **Mean**: computes all pixel value average of an image, by operating over the sliding-widow size of ‘m*n’

and replaces central pixel value.

$$f(x,y) = \frac{1}{mn}\sum_{(r,c)\in W} g(r,c) \tag{10}$$

- **Standard deviation**: it is used to find the amount of variation of dataset from mean. The data points are near to mean if it is less otherwise spreads out in a large range values. The change of intensity level is more at the edge of the image, due to its variability measure useful in edge sharpening.

$$\tilde{f}(x,y) = \sqrt{\frac{1}{mn-1}\sum_{(r,c)\epsilon W}\left(g(r,c) - \frac{1}{mn-1}\sum_{(r,c)\epsilon W}(g(r,c)\right)^2} \tag{11}$$

- **Variance**: it is used to find the edges positions in an image. It distinguishes distribution of moments in probability distributions.

$$\tilde{f}(x,y) = \frac{1}{mn-1}\sum_{(r,c)\epsilon W}\left(g(r,c) - \frac{1}{mn-1}\sum_{(r,c)\epsilon W}(g(r,c)\right)^2 \tag{12}$$

Where, $\tilde{f}(x, y)$-restored image, $g(r, c)$ is a noisy image, c, r-coordinates of rows and columns within W (Window size) in which operation takes place.

3.3. Kinetic Gas Molecule Optimization (KGMO)

The gas molecules considered as particles in the solution space, the kinetic energy considered for evaluating performance [16]. The particles inside the container travel until they converge that has the lowest energy and temperature. The particles attract each other depending on weak electrical inter particle Van-Der-Waal-forces, it is due to negative and positive charges in the particles. To guide the search, KGMO use kinetic theory based method instead of social behaviour imitation. Avoid trapping in local optima, explore high dimensional search space, find appropriate way to combine features and its performance is less sensitive to the adjustment of parameters. Comparatively high convergence speed, necessitate fewer computational sources so takes fewer iterations to converge to global minima.

In KGMO, every particle has: speed, location, mass and energy. Individual particle energy decides corresponding position and speed. Particle travel around the entire solution space to reach lowest temperature point. Consider a system having N particles, the i^{th} particle position is given by the Equation (13).

$$P_i = \left(P_i^1 \ldots P_i^d \ldots P_i^n\right), for\ i = 1,2,3 \ldots n \tag{13}$$

where, P_d^i is the i^{th} particle location in d^{th} dimension. The speed of the i^{th} particle is given by the Equation (14).

$$V_i = \left(V_i^1 \ldots V_i^d \ldots V_i^n\right), for\ i = 1,2,3 \ldots n \tag{14}$$

Where, P_i^d and V_i^d are the location and speed of i^{th} molecule in d^{th} dimension. The particle speed within the cylinder follows Boltzmann distribution i.e., speed is proportional to the particles exponential energy. This energy is given by the Equation (15).

$$\mathrm{K}_i^d(t) = 3/2(\mathrm{N}_b\mathrm{T}_i^d(t), \mathrm{k}_i = \left(\mathrm{K}_i^1 \ldots \mathrm{K}_i^d \ldots \mathrm{K}_i^n\right), \text{for } i = 1,2,3 \ldots n \tag{15}$$

Where, N represents molecules number, b is Boltzmann constant and $T_i^d(t)$ is i^{th} molecule temperature in d^{th} dimension at t. Updated molecule speed is given by the Equation (16).

$$\begin{aligned} V_i^d(t+1) = T_i^d(t)wv_i^d(t) + c1randi(t)\left(glob_{best}d - X_i^d(t)\right) \\ +c2randi(t)\left(Pers_{best}d(t) - X_i^d(t)\right) \end{aligned} \tag{16}$$

Where, T_i^d for the converging particles decreases exponentially related to time, it is calculated as

$$T_i^d(t) = 0.95X\left(T_i^d(t-1)\right) \tag{17}$$

The vector $(pers\_best)_i = (p\text{-}best)_i^1, (p\text{-}best)_i^2, \ldots (p\text{-}best)$ in indicate the best_previous_location of *ith* particle and $(g\text{-}best)_i = (g\text{-}best)_i^1, (g\text{-}best)_i^2, \ldots (g\text{-}best)_i^n$ is the best previous position amongst all of the particle inside the container. Initialization of each particle location and speed are carried out through random vectors in the consistent ranges. The limits of the particle speed are $[-v\_min, v\_max]$ and if $|v_i|>v\_max$, then $|v_i|=v\_max$ and w is inertia weight imitates particles resistance towards slowing down its speed. Also, $randi(t)$ is a uniform random variable in interval [0, 1] at time t for giving a randomized feature to the algorithm. C_1 and C_2 are the speed constants.

Each particle mass m is a random number in the range 0<m<1, it is once identified, in the algorithm execution remains constant because of assumption of molecule presence of only one kind inside the container at any point of time. To simulate various kinds of gases in different algorithmic iterations the random number is used. The particle location through equations of speed is given by

$$P(t+1)^i = 1/2a_i^d(t+1)t^2 + V_i^d(t+1)t + P_i^d(t) \tag{18}$$

Where, a_i^d is the i^{th} particle speed in d^{th} dimension. Through speed equation, get $a_i^d = (dV_i^d)/\mathrm{dt}$ otherwise, from laws of gas particles, obtain

$$(dk)_i^d = 1/2m(dv_i^d)^2 => dv_i^d = \sqrt{\frac{2(dK_i^d)}{m}} \tag{19}$$

Therefore, from Equations (21) and (22), the acceleration is given as:

$$a_i^d = \sqrt{\frac{\frac{2(dK_i^d)}{m}}{dt}} \tag{20}$$

In the dt time interval, Equation (20) re-written as:

$$a_i^d = \sqrt{\frac{2(dK_i^d)}{m}} \tag{21}$$

Then, from Equations (13), (18), the particle location is computed by

$$P(t+1)^i = 1/2a_i^d(t+1)(\Delta t)^2 + V_i^d(t+1)\Delta t + P_i^d(t) \tag{22}$$

$$P(t+1)^i = 1/2\sqrt{\frac{2(\Delta K_i^d)}{m}(t+1)\Delta t^2} + V_i^d(t+1)\Delta t + P_i^d(t) \quad (23)$$

Last, in each iteration of the algorithm m considered as a random number but the same for all of the particles in an iteration, for simplicity, for unit time interval location is reorganized through the Equation (24).

$$P(t+1)^i = 1/2\sqrt{\frac{2(\Delta K_i^d)}{m}(t+1)\Delta t^2} + V_i^d(t+1) + P_i^d(t) \quad (24)$$

The minimum fitness function is computed with the Equation (26)

$$pers\_best = f(x_i), iff(X_i) < f((pers_{best})_i) \quad (25)$$

$$glob\_best = f(x_i), iff(X_i) < f((glob_{best})_i) \quad (26)$$

Each of the particle try to change its location P_i^d with the distance among the current location $(pers\_best)_i$ the distance amongst the current location and $(glob\_best)_i$. The computing steps of KGMO are given below:

Algorithm 1: Pseudo code for kinetic gas molecule optimization.

1: for every gas molecule do
2: repeat initialization until it satisfy all constraints
3: end for
4: while mini error criterion or max iter not attained do
5: for every molecule do
6: Compute Fitness_Value
7: if (computed Fitness_Value is greater than prevailing best_Fitness_Value(Pers_Best) or not existing)
8: set Current_Value as the new Pers_Best
9: end if
10: end for
11: for every_molecule do
12: if (molecules Pers_Best is greater than global-best Fitness_Value (Glob_Best) in history or not exist (Glob_Best)) then
13: Set Current_Pers_Best as the new_Glob_Best
14: end if
15: Compute the kinetic energy of every gas molecule using Equation (15)
16: Update molecule_velacity using Equation (16) and position using Equation (18)
17: end for
18: end while

4. Result and Discussion

The IC-FF-KNN, Image Classification model is implemented in MATLAB, validated on distinct databases and evaluated model performance in terms of accuracy, by varying the number of features. The method is analysed with respect to the prevailing strategies and shown resultant retrieval performance.

4.1. Datasets

The proposed model is tested on the specified datasets:

- Wang: total images: 1000 with 384*256 pixel dimension, training set: 900, testing set: 100 [11].
- Cifar10: total images: 6000 with 32*32 pixel dimension, training set: 5000, testing set: 1000 [2].
- Oxfordflower: total images: 1360, training set: 1020, Testing set: 340 [18].
- Caltech101: total images: 9146, training set: 5000, Testing set: 1000 [12].
- ImageNet: total images: 14,197,122 with 384*256 dimension, training set: 8000, Testing set: 2000 [20].

4.2. Performance Analysis

The proposed IC-FF-KGMO model performance values on the five datasets are tabulated in Table 2 in terms of accuracy.

4.2.1. Accuracy

It is the ration of current predictions to the total predictions [10].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{Correctly\_Predicted}{Total\_Predictions} \quad (27)$$

True Positive (TP)-both the actual and predicted class are true, True Negative (TN)-model predict as the actual, but the actual is Negative, False Positive (FP)-model predict as actual, but the actual is negative, False Positive (FP)-model predicts as false and the actual value also negative.

Table 2. Accuracy on different datasets.

S. No.	Dataset	Accuracy (%)	Classification time (Sec)
1.	Wang	98	11.15
2.	Cifar10	98	10.13
3.	Oxfordflower	98	10.22
4.	Caltech101	96	11.67
5.	ImageNet	96	12.46

The IC-FF-KGMO method reaches global optimum with least number of iterations and accomplishes an average accuracy of 0.98, 0.98, 0.98, 0.96, and 0.96 for the standard databases and the values are tabulated in Table 2. The system produces the maximum accuracy for the small scale datasets such as 98% for Wang, Cifar10 and Oxfordflower databases next to that it produces accuracy of 96% for Caltech101 and ImageNet which are of large scale. Obtained classification accuracy of 4.63%, 16.7%, 11.29% and 4.41% more compared to the existing [5, 8, 19, 26] methods.

The presence of noise, outliers in hybrid features contribute less percentage to the model complexity, so the predictive capability of the model is high hence accuracy of small datasets is high and percentage contribution towards model complexity is more in the case of large scale dataset so model predictive power drops thus accuracy drops. Besides, the impact accuracy over the distinctive number of classes is computed, it decreases with the increase of number of classes.

4.2.2. Classification Time

It is the elapsed time between a requested image to the classification system classifies the image. The classification time for various databases is shown in

Table 2. Besides, the impact time over the distinctive number of classes is computed, it increases with the increase of number of classes.

4.2.3. Comparison of Accuracy to the Existing Approaches

The obtained result finalize that the proposed system is more accurate than the other systems. Table 3 show classification performance of the proposed system with respect to the metrics used and as well as its comparison to the existing systems [5, 19, 26].

Table 3. Accuracy and its comparison.

S. No.	Author	Method	Accuracy (%)
1	Rao and Mahantesh [19]	Color+Texture+Shape+VGG19 features+ML Algorithms	91.37
2	Hwang *et al*. [8]	GLCM+LBP and HoG+SIFT+DSIFT+KNN	79.30
3	Yang *et al.* [26]	Gobal+Texture+Shape+PCA+SVM	84.71
4	Greeshma and Gripsy [5]	CNN+HoG+LBP+SVM	91.59
5	**Proposed**	CH+GLCM+CSLBP+HoG+Gabour Filter+Arithmetic+KGMO+KNN	**96.00**

5. Conclusions

Proposed an Image Classification model (IC-FF-KGMO) using feature fusion optimized by KGMO. The three type of features are fused i.e., Color, Texture, and Arithmetic features of the test and training images. The method is tested with Wang, Oxfordflower, Cifar10, Caltech101, and ImageNet datasets. The obtained results of our proposal are better compared to the previously proposed algorithms in ([19], [26], and [5]). The outcome of the proposed approach is efficient with respect to accuracy and classification time. The proposed approach adds stability to its results for the minor changes in the training samples, so it fits well for the automatically obtained training sample. This approach makes machine learning based image classification appropriate for large scale image oriented real time applications. KGMO required hybridization with other algorithms to make convergence effective in higher dim space. The classification model accuracy can be improved through the modifications and improvements over the pre-processing phase, optimal hybrid-feature extraction phase, fusion of low level features with high level features to capture semantic information and exploring these to the classifiers for the training process and analyzing the effectiveness of machine algorithms.

Acknowledgment

We would like to thank anonymous authors for the guidance in this manuscript preparation.

References

[1] Badri N., Kboubi F., and Chaibi A., "Abusive and Hate Speech Classification in Arabic Text Using Pre-Trained Language Models and Data Augmentation," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 23, no. 11, pp. 1-28, 2024. https://doi.org/10.1145/3592615

[2] Beane R., CIFAR-10 Image Classification, https://kaggle.com/competitions/mu-cifar10, Last Visited, 2025.

[3] Das R., Kumari K., De S., Manjhi P., and Thepade S., "Hybrid Descriptor Definition for Content Based Image Classification Using Fusion of Handcrafted Features to CNN Features," *International Journal of Information Technology*, vol. 13, no. 4, pp. 1365-1374, 2021. https://doi.org/10.1007/s41870-021-00722-x

[4] Firuzi K., Vakilian M., Phung T., and Blackburn T., "Partial Discharges Pattern Recognition of Transformer Defect Model by LBP and HOG Features," *IEEE Transactions on Power Delivery*, vol. 34, no. 2, pp. 542-550. https://doi.org/10.1109/TPWRD.2018.2872820

[5] Greeshma K. and Gripsy V., "Image Classification Using HOG and LBP Feature Descriptors with SVM and CNN," *International Journal of Engineering Research and Technology*, vol. 8, no. 4, pp. 1-4, 2020. https://www.ijert.org/image-classification-using-hog-and-lbp-feature-descriptors-with-svm-and-cnn

[6] Hamza A., Javed A., Iqbal F., and et al., "Multimodal Religiously Hateful Social Media Memes Classification Based on Textual and Image Data," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 23, no. 114, pp. 1-9, 2024. https://doi.org/10.1145/3623396

[7] Hlaing C. and Zaw S., "Tomato Plant Diseases Classification using Statistical Texture Feature and Color Feature," *in Proceedings of the IEEE/ACIS 17th International Conference on Computer and Information Science*, Singapore, pp. 439-444, 2018. https://doi.org/10.1109/ICIS.2018.8466483

[8] Hwang S., Lee T., Kim H., Chung H., and et al., "Classification of Wood Knots Using Artificial Neural Networks with Texture and Local Feature-Based Image Descriptors," *Journal of Wood Research*, vol. 76, no. 1, pp. 1-13, 2022. https://doi.org/10.1515/hf-2021-0051

[9] Iqbal F., Abbasi A., Javed A., and *et al*., "Data Augmentation-Based Novel Deep Learning Method for Deepfaked Images Detection," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 20, no. 339, pp. 1-15, 2024. https://doi.org/10.1145/359261

[10] Jaradat Y., Masoud M., Manasrah A., Alia M., Suwais K., and Almanasra S., "Exploring the

Intersection of Information Theory and Machine Learning," *The International Arab Journal of Information Technology*, vol. 22, no. 5, pp. 845-858, 2025. https://doi.org/10.34028/iajit/22/5/1

[11] Khan U., Javed A., and Ashraf R., "An Effective Hybrid Framework for Content Based Image Retrieval," *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26911-26937, 2021. https://doi.org/10.1007/s11042-021-10530-x

[12] Kumar M. and Kumar M., "Performance Comparison of Various Feature Extraction Methods for Object Recognition on Caltech-101 Image Dataset," *in Proceedings of the Applications of Artificial Intelligence and Machine Learning*, Singapore, pp. 289-303, 2021. https://doi.org/10.1007/978-981-16-3067-5_22

[13] Malegori C., Franzetti L., Guidetti R., and Casiraghi E., "GLCM, an Image Analysis Technique for Early Detection of Biofilm," *Journal of Food Engineering*, vol. 185, pp. 48-55, 2016. https://doi.org/10.1016/j.jfoodeng.2016.04.001

[14] Mansourian L., Abdullah M., Abdullah L., Azman A., and Mustaffa M., "An Effective Fusion Model for Image Retrieval," *Journal of Multimedia Tools and Applications*, vol. 77, no. 13, pp. 16131-16154, 2018. https://doi.org/10.1007/s11042-017-5192-x

[15] Matsui Y., Ito K., Aramaki Y., Yamasaki T., and Aizawa K., "Sketch-based Manga Retrieval Using Manga109 Dataset," *Journal of Multimedia Tools and Applications*, vol. 76, no. 20, pp. 21811-21838, 2017. https://doi.org/10.1007/s11042-016-4020-z

[16] Moein S. and Logeswaran R., "KGMO: A Swarm Optimization Algorithm based on the Kinetic Energy of Gas Molecules," *Journal of Information Sciences*, vol. 275, pp. 127-144, 2014. https://doi.org/10.1016/j.ins.2014.02.026

[17] Nayak V., Arunalatha J., Vasanthakumar G., and Venugopal K., "Design of Deep Convolution Features Extraction for Multimedia Information Retrieval," *International Journal of Intelligent Unmanned Systems*, vol.11, no. 1, pp. 5-19, 2021. https://doi.org/10.1108/IJIUS-11-2021-0126

[18] Nilsback M. and Zisserman A., "Automated Flower Classification over a Large Number of Classes," *in Proceedings of the 6th Indian Conference on Computer Vision, Graphics and Image Processing*, Bhubaneswar, pp. 722-729, 2008. https://doi.org/10.1109/ICVGIP.2008.47

[19] Rao S. and Mahantesh K., "Hybrid Model for Image Classification: Bridging the Gap with Global Feature Descriptor," *Journal of Data Acquisition and Processing*, vol. 38, no. 3, pp. 2874-2883, 2023. DOI: 2023-03_2874.pdf

[20] Russakovsky O., Deng J., Su H., and et al., "ImageNet Large Scale Visual Recognition Challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211-252, 2015. DOI https://doi.org/10.1007/s11263-015-0816-y

[21] Sliti O., Hamam H., and Amiri H., "CLBP for Scale and Orientation Adaptive Mean Shift Tracking," *Journal of King Saud University Computer and Information Sciences*, vol. 30, no. 3, pp. 416-429, 2018. https://doi.org/10.1016/j.jksuci.2017.05.003

[22] Toennies K., *An Introduction to Image Classification: From Designed Models to End-to-End Learning*, Springer Singapore, 2024. https://doi.org/10.1007/978-981-99-7882-3

[23] Vatamanu O., Frandes M., Ionescu M., and Apostol S., "Content based Image Retrieval Using Local Binary Pattern, Intensity Histogram and Color Coherence Vector," *in Proceedings of the E-Health and Bioengineering Conference*, Iasi, pp. 1-6, https://doi.org/10.1109/EHB.2013.6707396

[24] Wang P., Fan E., and Wang P., "Comparative Analysis of Image Classification Algorithms based on Traditional Machine Learning and Deep Learning," *Pattern Recognition Letters*, vol. 141, pp. 61-67, 2021. https://doi.org/10.1016/j.patrec.2020.07.042

[25] Wang Y., Song B., Zhang P., Xin N., and Cao G., "A Fast Feature Fusion Algorithm in Image Classification for Cyber Physical Systems," *IEEE Access*, vol. 5, pp. 9089-9098, 2017. https://doi.org/10.1109/ACCESS.2017.2705798

[26] Yang F., Ma Z., and Xie M., "Image Classification with Superpixels and Feature Fusion Method," *Journal of Electronic Science and Technology*, vol. 19, no. 1, pp. 100096, 2021. https://doi.org/10.1016/j.jnlest.2021.100096

[27] Zhou J., Liu X., Xu T., Gan J., and Liu W., "A New Fusion Approach for Content Based Image Retrieval with Color Histogram and Local Directional Pattern," *International Journal of Machine Learning and Cybernetics*, vol. 9, no. 4, pp. 677-689, 2018. https://doi.org/10.1007/s13042-016-0597-9

Venkataravana Kadirappa is an Assistant Professor in the Department of Computer Science and Engineering, School of Computer Science and Engineering, Presidency University, Bangalore, India. Received his Master Degree in Engineering and Ph.D. Degree in Computer Science and Engineering from University Visvesvarya College of Engineering, Bangalore University, Bangalore. Currently Post-Doctoral research fellow at the Institute of Engineering and Technology, Srinivas University, Mangalore, India. His research interests primarily revolve around Digital Image Processing, Big Data and Artificial Intelligence.

Praveen Mokshanatha is working as Director, Research and Innovation Council, Department of Research, Srinivas University, Mangalore, India. Received P.D.F from Indian Institute of Science, Bangalore, India. His research areas are in the field of Nanocomposites, Electrodeposition, Nano Multilayers, and Corrosion.