Article

Deep Feature Fusion and Optimization-Based Approach for Stomach Disease Classification

Farah Mohammad [1],[*] and Muna Al-Razgan [2]

1 Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11451, Saudi Arabia
2 Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11451, Saudi Arabia; malrazgan@ksu.edu.sa
* Correspondence: fsheikh@ksu.edu.sa

Abstract: Cancer is the deadliest disease among all the diseases and the main cause of human mortality. Several types of cancer sicken the human body and affect organs. Among all the types of cancer, stomach cancer is the most dangerous disease that spreads rapidly and needs to be diagnosed at an early stage. The early diagnosis of stomach cancer is essential to reduce the mortality rate. The manual diagnosis process is time-consuming, requires many tests, and the availability of an expert doctor. Therefore, automated techniques are required to diagnose stomach infections from endoscopic images. Many computerized techniques have been introduced in the literature but due to a few challenges (i.e., high similarity among the healthy and infected regions, irrelevant features extraction, and so on), there is much room to improve the accuracy and reduce the computational time. In this paper, a deep-learning-based stomach disease classification method employing deep feature extraction, fusion, and optimization using WCE images is proposed. The proposed method comprises several phases: data augmentation performed to increase the dataset images, deep transfer learning adopted for deep features extraction, feature fusion performed on deep extracted features, fused feature matrix optimized with a modified dragonfly optimization method, and final classification of the stomach disease was performed. The features extraction phase employed two pre-trained deep CNN models (Inception v3 and DenseNet-201) performing activation on feature derivation layers. Later, the parallel concatenation was performed on deep-derived features and optimized using the meta-heuristic method named the dragonfly algorithm. The optimized feature matrix was classified by employing machine-learning algorithms and achieved an accuracy of 99.8% on the combined stomach disease dataset. A comparison has been conducted with state-of-the-art techniques and shows improved accuracy.

Keywords: stomach disease; deep features; features optimization; disease classification

1. Introduction

In medical imaging, gastric lesion identification is an active research domain. Different gastric lesions include bleeding, esophagitis, ulcer, and polyps. Bleeding and ulcer are the most common among all abnormalities [1]. These stomach lesions have become a leading cause of mortalities in humans [2]. Globally, stomach cancer is the third major cause of death among all cancer deaths [3]. In the global cancer cases, esophageal cancer is at the seventh position [4]. In 2021 in the United States, Siegel et al. [5] estimated 26,560 new cases of stomach cancer of which 16,160 cases were in males and 10,400 were in females. The approximate deaths were 11,180, including 6740 males and 4440 females. In the United States, the American Cancer Society estimated there would be 26,380 new patients with stomach cancer for the year 2022. The estimated cases in men are 15,900 and 10,480 in women. The estimated deaths from these cases are 11,090, including 6690 males and 4400 females [6].

Wireless capsule endoscopy (WCE) [7] is the medical imaging method used to analyze the gastrointestinal (GI) tract. This imaging technique is widely utilized in clinics for the examination of the GI tract. In this technique, a tiny camera is sent into the GI tract to record the images. During the WCE procedure, about 50,000 images are generated. The gastroenterologists examine these images, which is a time-consuming process that is neither efficient nor very accurate. An average of two hours are required for an expert to analyze the images [8]. This issue can be resolved by developing automated computer-aided diagnostic (CAD) systems. These CAD models extract features from the WCE images to diagnose diseases. The fundamental steps of CAD systems include data acquisition, pre-processing, feature extraction, feature selection or feature optimization, and finally, classification.

In recent years, several computer vision (CV) researchers have introduced automated CAD models for the recognition of GI abnormalities [9–12]. The CAD models utilize hand-crafted features and deep convolutional neural network (CNN) features for the recognition of WCE images. Different studies have utilized color features [13,14], texture features [15], point features [16], and histogram of oriented gradients (HOG) features [17]. With the recent developments in the deep-learning area, some researchers have utilized deep features to identify GI diseases.

Stomach disease recognition is a challenging task due to the presence of several diseases and the availability of disease databases. The propound method of stomach disease classification's main goal is to accurately predict the disease with less computation time. Deep-learning models inception v3 and DenseNet-201 are utilized in this paper for the deep feature extraction from stomach endoscopic images. Features fusion is performed on deep extracted features to attain rich feature space. The presence of redundant features which impacts the accuracy and prediction time came under consideration. The meta-heuristic approach was adopted to optimize the features for the accurate classification of stomach infections with less processing time. The robust selected features classification was performed using several machine algorithms. The major contribution of our propounded technique is as follows:

- Fine-tuning of pre-trained deep-learning models was performed with parameter modification for the derivation of high-level features by implementing the activation function at the average pooling layer.
- Deep features were concatenated using the parallel maximum covariance (PMC) method, and the resultant feature map was optimized using the nature-inspired modified binary dragonfly algorithm to reduce feature redundancy.

The remaining paper is organized as follows: Section 2 comprises related work, and Section 3 contains a brief description of the proposed methodology. Section 4 illustrates the experiments and results in a detailed description, and the conclusion of the paper is presented in Section 5.

2. Related Work

Several models have been designed for the detection and identification of stomach abnormalities. The presented techniques aim to identify stomach diseases in less computation time with higher accuracy. Different deep-learning techniques [18] equipped with feature optimization approaches for robust disease recognition have been utilized by the researcher [19]. In [20], the authors introduced a method for the recognition of bleeding and ulcers in WCE images. In the first step, they performed preprocessing techniques on the dataset including 3D-median filtering, 3D-box filtering, and HSV color variation. In the second phase, extraction of geometric features was performed to obtain binary images, and masks were generated and implemented to further improve the data. After that, the researchers extracted shape, and color, and speeded up robust features (SURF), and a serial-based fusion technique was applied to obtain the feature set. In the third step, feature selection was performed using the correlation coefficient and principal component analysis (PCA) approaches. Finally, in the classification phase, the researchers achieved a good classification accuracy on the SVM classifier. Zhao et al. [21] introduced TriZ, a rotation

tolerant image feature for the recognition of gastric infections. Researchers compared the TriZ with the HOG method and observed a better recognition rate. Only 126 image features of TriZ were utilized for the experiment to compare with the HOG method. They achieved an accuracy of 87% using their model.

Majid et al. [22] classified stomach infections using handcrafted and deep features. They computed discrete cosine transform (DCT), color, discrete wavelet transforms (DWT), and VGG16 features in the feature extraction step. After extraction of the features, the feature set was given to the genetic algorithm (GA) for the selection of the best features. The maximum classification accuracy achieved on deep and classical optimized features was 96.5% on the ensemble classifier. Khan et al. [2] introduced an automated deep-learning technique based on deep feature classification for the recognition of stomach infections. First, researchers utilized a saliency-based technique for lesion detection. After lesion recognition, transfer learning was performed on the pre-trained VGG16 model, and features were computed from the fully connected layer. Then particle swarm optimization (PSO) was applied to the feature set for optimization. Finally, the selected features were fed to the classifiers for recognition. Using this model, they obtained the best accuracy on the Cubic SVM.

Researchers in [23] recognized stomach diseases using deep features. Researchers extract features from the Inception V3 model and implemented PSO and the crow search algorithm (CSA) for feature optimization. The results of both optimization techniques were fused and fed into a multi-layer perceptron for classification. Their technique outperformed as compared to the other methods. Bora et al. [24] designed a model for the recognition of polyps. They utilized a generic Fourier descriptor (GFD) for shape features extraction, while color and texture features were collected using the non-subsampled contourlet transform (NSCT). After feature extraction, they evaluated the significance of features using the analysis of variance (ANOVA). In the feature selection phase, researchers utilized the fuzzy entropy-based method. Finally, the selected features were classified using a multi-layer perceptron, and least square SVM. Ayyaz et al. [25] utilized deep CNN models for the recognition of gastric abnormalities. First, they implemented the transfer-learning technique on AlexNet and VGG19. Then researchers extracted features from both CNN models and gave them to the GA for feature optimization. After that, the selected features were combined using the serial-based technique and given to the multiple classifiers for final recognition. Their model obtained a promising recognition rate on Cubic SVM. In [19], researchers extracted ResNet101 [26] features to classify stomach infections. Urban et al. [27] utilized features from VGG16 [28], VGG19 [29], and ResNet50 [26] models for the classification of polyps. Some researchers observed that the combination of the handcrafted and CNN features enhances the performance of the CAD model. Different classical and deep-learning-based approaches have been proposed by authors, but the challenges such as features redundancy, computational cost, and model robustness persist [30,31].

3. Materials and Methods

A detailed description of the proposed method is presented in this section. The method comprises several steps for stomach disease classification. In the first phase, data augmentation is employed to balance and increase the per-class images using flip and transpose operations. At the second phase, fine-tuned deep CNN models, namely inception v3 and DenseNet-201, are employed for deep features derivation by utilizing transfer learning and features fusion performed on deep derived features maps. The concatenated feature space has redundant features, which were removed using a metaheuristic approach. In the last phase, a features optimization binary dragonfly algorithm was implemented on the fused feature vector to remove redundant features, and classification was performed to classify the stomach disease. The detailed explanation of our proposed model for stomach disease classification is illustrated in Figure 1.
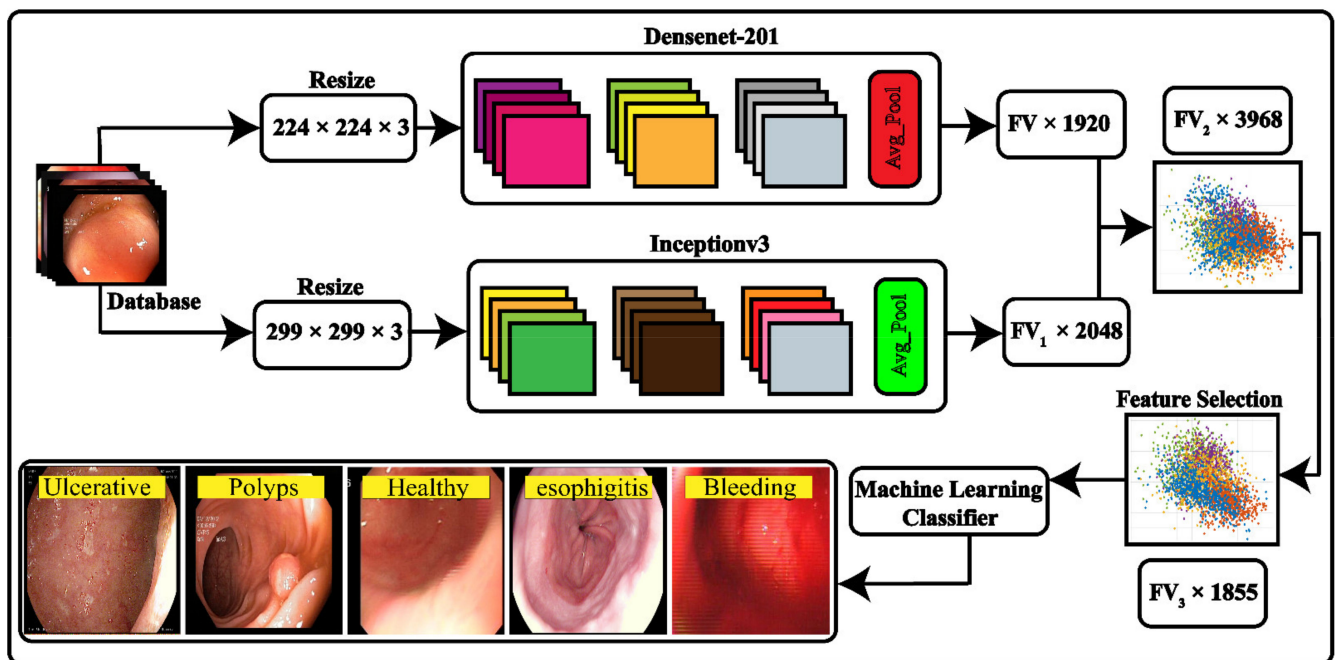
Figure 1. Proposed technique for stomach disease classification.

3.1. Data Acquisition and Preparation

The proposed stomach disease recognition method is utilized to classify the five major diseases of the stomach. We acquired two classes of bleeding and healthy from datasets [9] containing 3000 images per class. Esophagitis, polyps, and ulcerative colitis were extracted from publicly available challenging stomach gastrointestinal tract Kvasirv1 having 500 images per class and KvasirV2 having 1000 images per class [32]. The imbalance in the division of images per class and the varied sizes affected the performance of the proposed classification techniques. Data augmentation is a prominent step toward deep-learning performance. The inclusion of more data for the training of deep-learning models increases their performance. The maximum number of images per class is 3000, and the minimum images per class are 500. We performed different flip and transpose operations, which increased the number of images without losing the information and the features of images. The applied flip operation increased the dataset images and equaled out the images per class. The sample images after the augmentation process are shown in Figure 2. The mathematical notation of the flip operation on images is as follows:
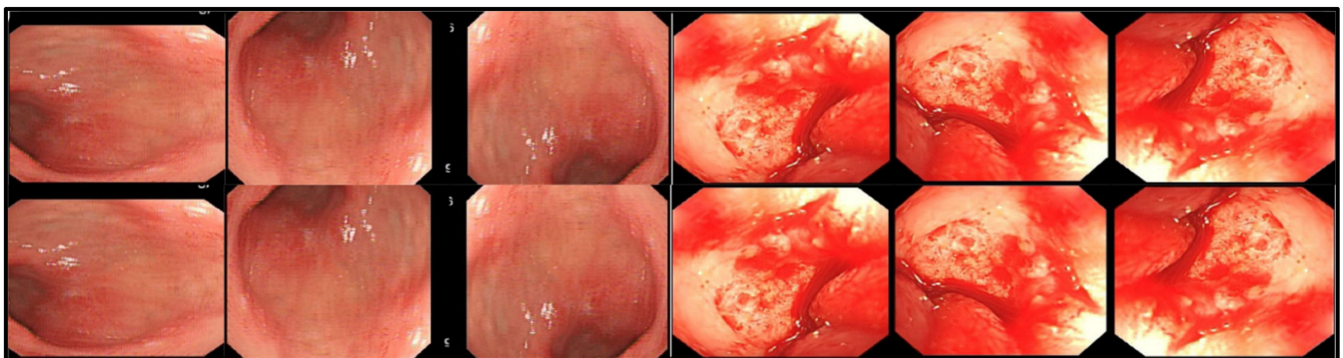

Figure 2. Sample Images after data augmentation.

Let us have an input size of the image matrix 256×256 expressed as $\widetilde{N}_{l,m}$ having lth rows and mth column and $\widetilde{N}_{l,m} \in S^{l \times m}$. The matrix row $l = \{1, 2, 3 \ldots \widetilde{o}\}$ and column

$m = \{1, 2, 3 \ldots \widetilde{p}\}$ where the number of channels is 3. The orientation of the RGB image was processed using three data augmentation operations.

$$\widetilde{N}^{T} = \widetilde{N}_{l,m} \tag{1}$$

The transposition of the original image is expressed using $\widetilde{N}^{T}$, the operation which alters the indices of images.

$$\widetilde{N}^{H} = \widetilde{N}_{l}(\widetilde{p} + 1 - m) \tag{2}$$

The horizontal flip of images is expressed as $\widetilde{N}^{H}$.

$$\widetilde{N}^{V} = \widetilde{N}_{(\widetilde{o}+1-l)} \tag{3}$$

$\widetilde{N}^{V}$ illustrates the vertical image flip. The mentioned augmentation operation is repeated to equal the number of images per class and reach the 3500 number of images.

3.2. Deep Feature Extraction

Pattern recognition and computer vision tasks are performed using features derived from images. The features present an object according to its color, shape, and positioning. In the computer vision and medical imaging domain, the inclusion of deep learning significantly improves the diagnosis of diseases [33]. A deep convolutional neural network (CNN) normally has different layers such as the input layer, the convolutional layer, batch normalization, fully connected, and the activation layer ReLU. The CNN input layers feed the input data to the convolutional layer, and the weights are calculated. The activation function is applied using the ReLU layer, and inactive neurons are removed. The classification is performed on features computed using a fully connected layer by SoftMax layer. In our proposed method of stomach disease recognition, two deep CNN models, inception v3 and DenseNet-201 were implanted for deep feature derivation. The next sections comprise a description of the deep-learning models.

3.2.1. Inception v3

Inception v3 is a deep-learning model that has robust performance in classification tasks. The model has a directed acyclic graph (DAG) with 94 convolutional layers, 316 layers, and 350 connections [34]. The architecture of the DAG network is complex due to multiple inputs to different layers at the same time. Several masks are implemented on different layers for the derivation of different features. The diverse architecture of inception v3 allows employing different masks and parameters on different layers as compared to the conventional CNN model, which has predetermined parameters on layers. Inception v3 was trained on ImageNet, a challenging image dataset having over a million images and 1000 classes [35]. The deep-learning model learned information about multiple objects and categories. The input size of inception v3 was $299 \times 299 \times 3$. The first convolutional layer was processed by performing activation to derive a feature matrix having a size of $149 \times 149 \times 3$, and the filter applied had a size of 32. The ReLU function was utilized to perform activation, and the next batch normalization was performed. ReLU, an activation function, is expressed as follows:

$$Te_{j}^{N} = \max\left(lw.lw^{(N-1)}\right) \tag{4}$$

A pooling layer is sandwiched in convolutional layers having a filter size of 2×2, which is illustrated below

$$oz_{1}^{(q)} = oz_{1}^{(q-1)} \tag{5}$$

$$oz_{2}^{(q)} = \frac{oz_{2}^{(q)} - H(q)}{S^{q}} + 1 \tag{6}$$

$$oz_{3}^{(q)} = \frac{oz_{3}^{(q)} - H(q)}{S^{q}} + 1 \tag{7}$$

where $oz_1^{(q)}$, $oz_2^{(q)}$, and $oz_3^{(q)}$ are filters applied on the feature vector, and S^q expresses the stride. Several layers are concatenated before the average pooling layer. The average pooling layer deployed for the derivation of the deep CNN feature map has a size $FV_1 \times 2048$ by performing activation.

3.2.2. DenseNet-201

A deep CNN model is known for the robust classification and recognition tasks. The DenseNet-201 [36] layers, having a directed connection to the next layers, increase the learning rate with minimal information loss. The network has fewer parameters in comparison to the other CNN models. Information maintainability requires less information loss from the first to the last layer. The information and features extracted at the different network layers can easily be predicted. The presence of gradient function decreases the chances of overfitting. The input size of the DenseNet-201 is $224 \times 224 \times 3$.

Suppose an image ξ_z is used as an input to DenseNet-201. The network has M number of layers and transformation filters which are nonlinear $S_m(.)$. The transformation filter $S_m(.)$ is utilized as a concatenated function of convolution, batch normalization pooling, and ReLU. The output of the mth the layer is computed in a classical CNN model as $(m+1)^{\text{th}}$ and mathematical modeling as follows:

$$B_m = S_m(B_m - 1) \tag{8}$$

DenseNet-201 layers have a direct connection to each other, and the mth layers carry information computed from all network layers $b_0, b_1 \ldots \ldots b_{n-1}$ which can be defined as

$$b_n = S_m([b_0, b_1 \ldots \ldots b_{n-1}]) \tag{9}$$

In the above equation, $b_0, b_1 \ldots \ldots b_{n-1}$ are the computed feature map of layer $0, \ldots \ldots n-1$. The resultant feature vector of previous layers is utilized by the average pooling layer and activation is employed to extract the required deep CNN feature map having a size of $FV \times 1920$.

3.2.3. Feature Extraction Using Transfer Learning

Transfer learning has been employed on the deep CNN model for different machine-learning tasks and proved to be robust in recognition and classification tasks. We implanted transfer learning on our pre-trained fine-tuned deep CNN models for deep features extractions. The created stomach disease classification data was used to collect deep features by adopting transfer learning. We utilized 70:30 practice for the data split, training was performed on 70% data, and the remaining 30% was used for testing. The preprocessing step resized the images according to the input size of fine-tuned deep-learning networks. We employed the fine-tuned DenseNet-201, and the convolutional layer of the network was utilized as an input layer. The average pooling layer was used for the implementation of the activation function for features computation. The derived features map had a $FV \times 1920$ size and s expressed as φ^{m1}. The architecture of DenseNet for transfer learning is expressed in Figure 3.

We employed the pre-trained Inception V3 for deep features derivation by utilizing transfer learning. The convolutional layer used for image input and average pooling layer was activated using the activation function. The derived features vector size was $FV_1 \times 2048$ and the feature map of derived features are illustrated using φ^{m2}. The derived features from DensNet-201 and inception V3 were fused for training and testing of the model. The architecture of inception v3 is illustrated in Figure 4. Training of the pre-trained model was performed using a sigmoid function. We utilized different parameters such as 250 epochs, iteration 30, learning rate 0.0001, batch size 64, and data shuffling was performed on all epochs.
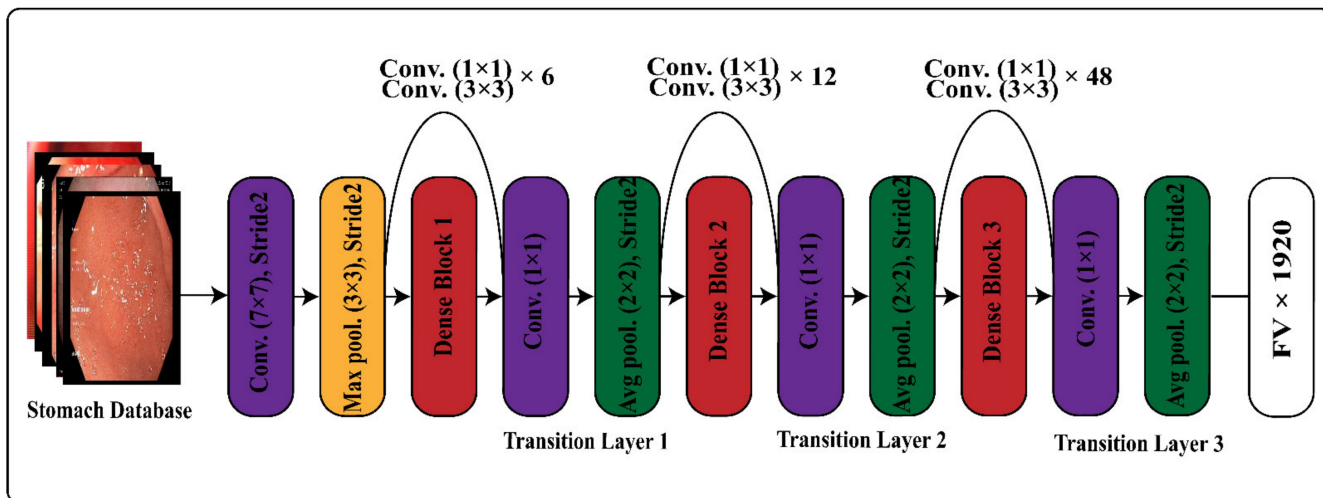
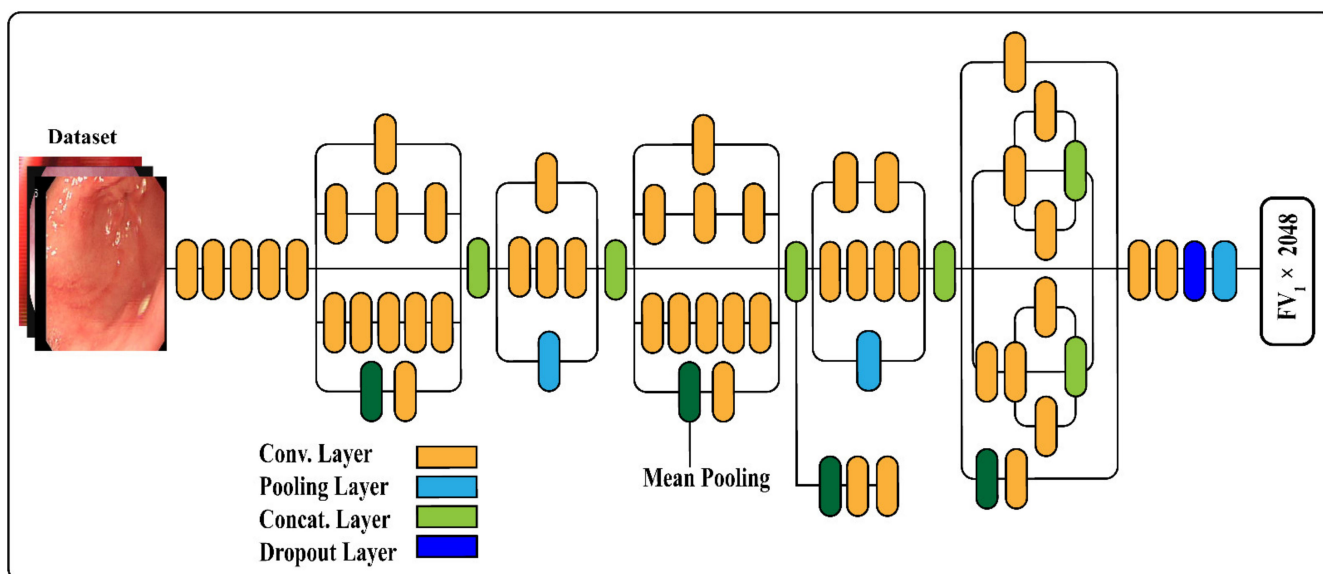**Figure 3.** Transfer-learning architecture of DenseNet-201 for feature extraction.

Figure 4. Transfer-learning formation of Inception V3 for feature derivation.

3.3. Features Fusion

Feature fusion provides rich feature space for pattern recognition applications [37]. Recognition and classification of objects and images require an enriched feature space to recognize the specific image and feature concatenation that provides specific dense feature space. Feature concatenation provides efficient feature space for the classification of objects but impacts the processing time. A novel feature concatenation technique, parallel maximum covariance (PMC), is employed in our method for feature fusion. The integration of two deep CNN feature maps results in a single feature space.

Suppose the deep CNN features extracted from two deep CNN models are φ^{m1} and φ^{m2}. The dimensions of these two feature vectors Fv and Fv1 are $q \times r$ and $q \times s$, the q represents the number of images in a feature space where r and s represents the length of the attribute matrix. The features extracted using the pre-trained inception v3 are $q \times 2048$, and the features derived using the pre-trained deep CNN model DenseNet-201 are $q \times 1920$. The dimensions of feature spaces are equalized by computing the average and the addition of the average as padding. Suppose the d denotes a column vector n in a feature space φ_1,

and pattern field φ_2 has a column vector e. The utilization of time series presents the row vectors as:

$$z_1 = \varphi_1{}^T \varphi^{m1} \tag{10}$$

$$z_2 = \varphi_2{}^T \varphi^{m2} \tag{11}$$

The maximum covariance of φ_1 and φ_2 can be described as

$$\check{f} = Cov[z_1, z_2] \tag{12}$$

$$\check{f} = Cov\left[\varphi_1{}^T \varphi^{m1}, \varphi_1{}^T \varphi^{m1}\right] \tag{13}$$

$$\check{f} = \frac{1}{n-1} Cov\left[\varphi_1{}^T \varphi^{m1}, \varphi_1{}^T \varphi^{m1}\right] \tag{14}$$

$$\check{f} = \varphi_1(F\varphi_1\varphi_2)\ \varphi_2 \tag{15}$$

$$F\varphi_1\varphi_2 = \frac{1}{n-1}\left(\varphi^{(m1)}\varphi^{(m2)T}\right) \tag{16}$$

where $F\varphi_1\varphi_2$ represents covariance between φ_1 and φ_2, and features are expressed as ith and jth for φ_i and φ_j. The maximum covariance of the final concatenated feature space is $F\varphi_1\varphi_2$. The features concatenation process makes the feature map dense and, in addition, creates redundant features. The final feature map obtained after feature fusion has a $FV_2 \times 3968$ dimension.

3.4. Features Optimization

Feature assortment is the robust process of obtaining the most relevant features using features selection algorithms. The prime goal of features optimization is to reduce the presence of irrelevant features that impact classification functioning and computation time. The implanted features optimization method is illustrated in Figure 5. In our propounded technique, we utilized the binary dragonfly metaheuristic optimization method which utilized the KNN fitness function for robust feature selection. A detailed description of the implanted optimization method dragonfly is presented here.
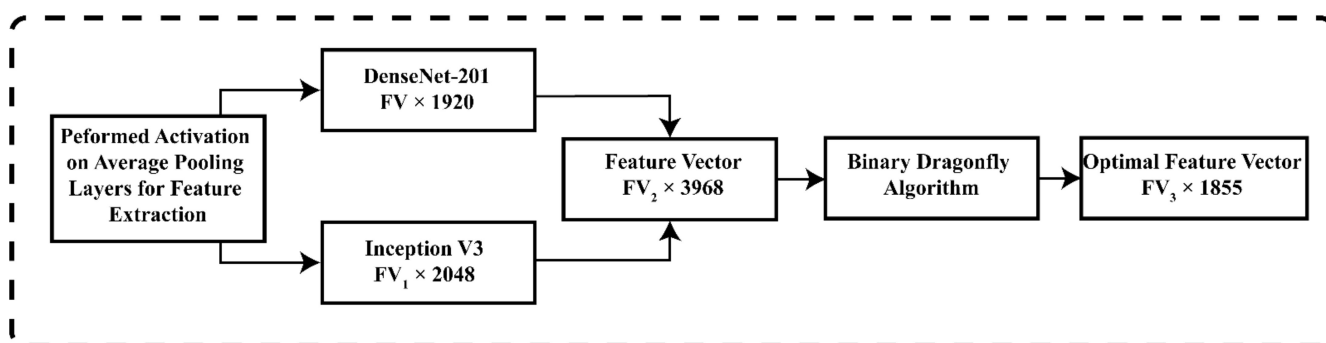

Figure 5. Proposed feature optimization architecture.

Recently, Mirjalili et al. [38] introduced the nature-inspired dragonfly algorithm (DA). This is a population-based metaheuristic technique, motivated by the hunting and migration behavior of dragonflies. Small groups of dragonflies move to find food sources in what is known as the hunting procedure. In the migration process, the larger groups of dragonflies fly in one direction. The swarming behavior of dragonflies is explained by the following five parameters.

Separation: This operator ensures that the search agents in the neighborhood keep away from each other. Mathematically, it can be described as:

$$P_i = -\sum_{j=1}^{N} A - A_j \tag{17}$$

where N represents the neighborhood size, A is the current location of the individual, and A_j denotes the j-th neighbor of the position A.

Alignment: This parameter describes the velocity of an individual according to the other neighboring individuals. The following equation describes this behavior:

$$X_i = \frac{\sum_{j=1}^{N} W_j}{N} \tag{18}$$

W_j is the velocity of the j-th neighbor.

Cohesion: It indicates the individual's movement behavior between the neighborhood to the center of the mass. This can be mathematically explained as:

$$M_i = \frac{\sum_{j=1}^{N} A_j}{N} - A \tag{19}$$

Attraction: This parameter defines the attraction of the flying individual towards a food source. This can be modeled as:

$$H_i = H_{loc} - A \tag{20}$$

where H_{loc} denotes the location of the food source.

Distraction: The movement of an individual away from the enemy is a distraction. It can be given as:

$$K_i = K_{loc} + A \tag{21}$$

where, K_{loc} is the enemy position.

In the DA, the optimization problem is solved using a step vector and a position vector. The following equation defines the step vector:

$$\Delta A_{t+1} = (sS_i + xX_i + mM_i + hH_i + kK_i) + \omega \Delta A_t \tag{22}$$

where s is the separation weight, S_i denotes the i-th individual separation, x denotes the alignment weight, X_i represents the alignment of i-th individual, m is the cohesion weight, M_i represents the cohesion of i-th individual, h is the food factor, H_i denotes the i-th individual's food source, k is the enemy factor, K_i is the location of the enemy of the i-th individual, ω indicates the inertia weight, and t indicates the iteration number which is 100 in our proposed modified optimization model.

In the search space, a step vector is added to the previous position to revise the position of dragonflies. However, the following equation is used in a binary search space:

$$A(t+1) = \begin{cases} A_t, r < T(\Delta a_{t+1}) \\ A_t, r \geq T(\Delta a_{t+1}) \end{cases} \tag{23}$$

where r represents the random number between the range 0 and 1. $T(\Delta a_{t+1})$ is the transfer function that calculates the probability of a location update for all dragonflies, and this is given as:

$$T(\Delta a) = \left| \frac{\Delta a}{\sqrt{\Delta x^2 + 1}} \right| \tag{24}$$

The pseudo-code of the modified Binary Dragonfly is presented below (Algorithm 1):

Algorithm 1: Pseudocode of the Dragonfly Algorithm (DA)

Input: Feature Vector ($N \times \mathbf{3968}$)
Output: Selected Feature vector ($N \times \mathbf{1855}$)
Maximum Iteration = 100
Step 1: Initializing the population $A_i(i = \mathbf{1, 2, 3, \ldots, n})$
Step 2: Initialize $\Delta A_i(i = \mathbf{1, 2, 3, \ldots, n})$
Step 3: while ($t < \textbf{Max Iteration}$) do

-     evaluate each dragonfly
-     Update (H) *and* (K)
-     Update the coefficients (s, x, m, h, k, *and* ω)
-     Calculate (S, X, M, H, **and** K)
-     Update Step Vector (ΔA_{t+1})
-     Update A_{t+1}

Step 4: Return: the best solution

---

4. Results

The proposed technique of stomach disease classification is implanted on the stomach dataset created from challenging datasets. We acquired different disease images such as a polyp, ulcerative-colitis, and esophagitis from challenging datasets such as Kvasir V1 and Kvasir V2 [32] and two classes, bleeding, and healthy obtained in [9]. A few sample images from the created dataset are shown in Figure 6. Cross-validation of the 10-fold and the 70:30 ratio was employed, 70% for training and 30% for testing.

The extensive experiments and simulation of the propounding method were processed on a system having an Intel Core i7 9th generation processor, a memory of 16 gigabytes, and 11 gigabytes of the graphic processing unit. Several machine-learning algorithms were employed for the prediction of stomach diseases, and a robust one was determined based upon accurate disease prediction and processing time. The robustness of the implanted method was assessed using several performance validation metrics such as accuracy, precision, recall, f1-score, false-negative rate (FNR), and computational time.
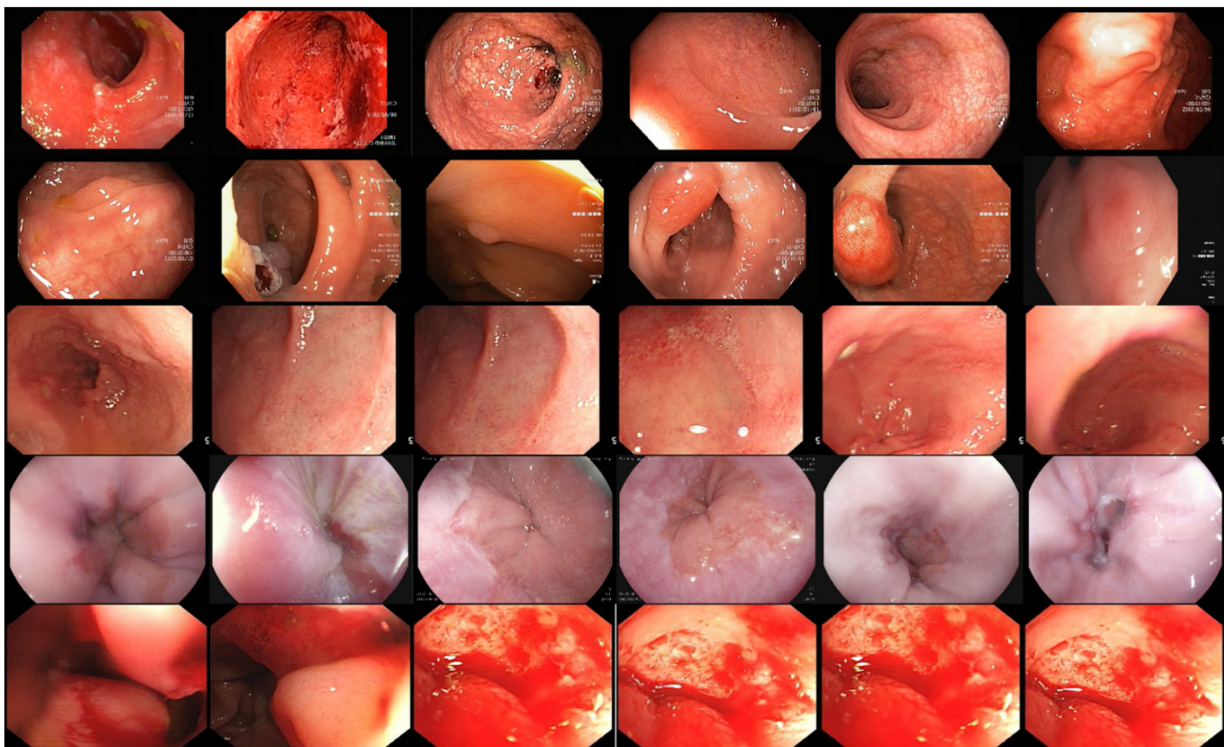
Figure 6. Sample images from stomach disease classification dataset.

4.1. Deep Feature Fusion Results

A brief description of the aspired methodology results has been demonstrated in this section. The numerical outcomes of the implanted deep features fusion technique are shown in Table 1. The deep features are derived from pre-trained deep CNN models and parallel fusion functions. Several machine-learning classifiers were applied to the fused feature map to perform the stomach disease classification. In our experiments, the cubic SVM obtained the highest recognition accuracy of 96.2% with an FNR of 3.8%. The remaining machine-learning algorithms employed for the recognition of stomach disease on the fused feature space were linear support vector machine (L-SVM), quadratic support vector machine (Q-SVM), medium Gaussian support vector machine (MG-SVM), cosine Gaussian support vector machine (CG-SVM), Gaussian-naïve Bayes (GN-Bayes), ensemble subspace discriminant (ESD), cosine K-nearest neighbor (C-KNN), and linear discriminant (LD) and their attained corresponding classification results are 96%, 96.1%, 95.6%, 95.2%, 93.7%, 90.7%, 93.6%, 91.5%, respectively. The concatenation of deep CNN features increased the accurate classification but also increased the processing time.

Table 1. Stomach disease classification results using deep CNN features fusion.

Classifier	Evaluation Measures		
	Accuracy (%)	FNR (%)	Time (s)
L-SVM	96	4	514
Q-SVM	96.1	3.9	730
C-SVM	96.2	3.8	611
MG-SVM	95.6	4.4	1310
CG-SVM	95.2	4.8	1214
GN-Bayes	93.7	6.3	354
ESD	90.7	9.3	125
C-KNN	93.6	6.4	153
LD	91.5	8.5	272

4.2. Deep Feature Optimization Results on (CV = 10)

The main goal of the proposed technique was to accomplish the highest classification accuracy of stomach diseases in minimal computational time. The obtained fused feature space was optimized using a robust feature optimization algorithm called the binary dragonfly algorithm. The optimization technique eliminated the redundant features and picked the most relevant and robust features. The selection of the best features increases the disease classification functioning in less processing time. The proposed features optimization technique outcomes are presented in Table 2. Different machine algorithms are adopted for stomach disease recognition, different performance evaluation measures are utilized for the assessment of the methodology, and a robust one was selected based on processing time and accurate classification rate.

Table 2. Stomach disease classification with proposed feature optimization technique (CV = 10).

Classifier	Evaluation Measures					
	Accuracy (%)	Precision (%)	FNR (%)	Recall (%)	F1_Score (%)	Time(s)
L-SVM	98.6	98.8	1.4	99.1	98.9	65.271
Q-SVM	99.1	99.2	0.8	98.3	98.7	69.563
C-SVM	**99.8**	**99.8**	**0.2**	**99.4**	**99.6**	**33.18**
MG-SVM	96.5	97	3.5	97.5	97.2	252.343
CG-SVM	95.6	95.6	4.4	96.6	96.1	102.235
GN-Bayes	92.6	92.7	7.3	93.7	93.2	89.102
ESD	94.2	94.3	5.8	95.2	94.7	125.36
C-KNN	97.4	98.4	2.6	98.4	98.4	27.151
LD	98.3	98.7	1.7	99.3	99	32.14

C-SVM attained the highest accuracy of 99.8% with 0.2% FNR, 99.8% precision, 99.4% recall, and 99.6% f1-score in a 33.18 s processing time. L-SVM, Q-SVM, MG-SVM, CG-SVM, GN-Bayes, ESD, C-KNN, and LD accomplished an accuracy of 98.6%, 99.1%, 96.5%, 95.6, 92.6%, 94.2%, 97.4%, and 98.3%, respectively. The most robust classifier in terms of processing time was LD with a processing time of 32.14 s, and the worst classifier was MG-SVM with a computational time of 252.343 s as presented in Figure 7.
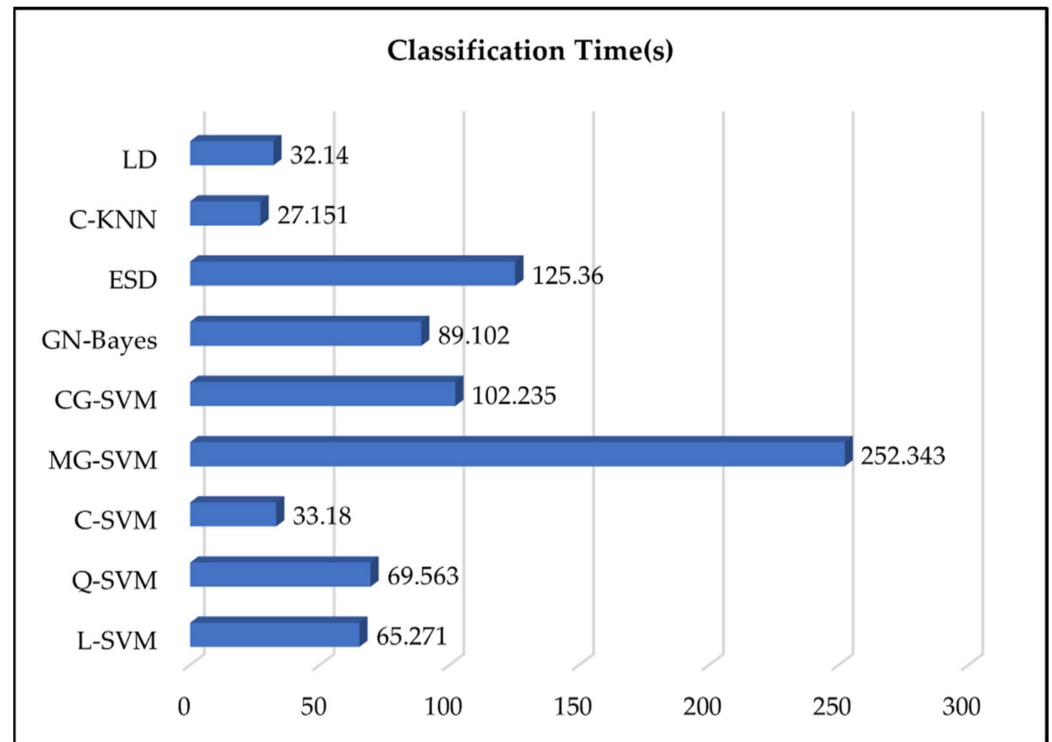

Figure 7. Computational time comparison of utilized machine-learning classifiers.

4.3. Deep Feature Optimization Results on (CV = 15)

This section of results provides the proposed feature fusion and optimization method results on several machine-learning classifiers. Different optimized features vectors were fed to machine-learning algorithms to evaluate the performance of the proposed fusion and optimization approach. The main goal of experiments of the proposed technique at (CV = 15) was to analyze the performance of stomach disease recognition. Results are presented in Table 3 and show the highest accuracy of 99.6% with a 99.3 f1 score.

Table 3. Stomach disease classification with proposed feature optimization technique (CV = 15).

Classifier	Evaluation Measures					
	Accuracy (%)	Precision (%)	FNR (%)	Recall (%)	F1_Score (%)	Time(s)
L-SVM	99.5	98.4	0.5	99.1	98.7	67.271
Q-SVM	99.5	99.1	0.5	98.3	98.7	73.561
C-SVM	**99.6**	99.2	0.4	99.4	99.3	39.18
MG-SVM	99.4	99.3	0.6	99.5	99.4	286.393
CG-SVM	99.2	99	0.8	99.4	99.2	152.135
GN-Bayes	98.3	98.1	1.7	99.3	98.7	92.112
ESD	99.2	99.1	0.8	99.4	99.2	139.36
C-KNN	99.4	98.6	0.6	99.6	99.1	**29.121**
LD	98.9	98.4	1.1	99.3	98.8	35.11

In addition to the highest accuracy, the other performance evaluation measure such as precision 99.2%, recall 99.4% with 0.4% FNR were also computed on the C-SVM machine-

learning algorithm. The worst classification performance was achieved on GN-Bayes with 98.3% accuracy. The best machine algorithm which utilized less computational cost was linear discriminant, and the worst classifier in terms of computation time was MG-SVM. The results show that the C-SVM achieved overall better accuracy, and LD achieved the best computational time. The overall performance of C-SVM is better in stomach disease recognition.

The results highlight the robustness of the proposed deep features extraction and optimization using the binary dragonfly algorithm. The highest classification accuracy of machine-learning algorithms using features optimization also reduced the computational time. The computational time of employed classifiers is presented in Figure 7.

The optimized feature map classification was executed using various machine-learning algorithms, and a robust one was selected based upon the highest accuracy and less processing time. The C-SVM obtained the highest classification accuracy of 99.8% with a FNR of 0.2%. The robustness of the C-SVM classifier was also verified using the confusion matrix expressed in Table 4.

Table 4. Confusion matrix of proposed features optimization for stomach disease classification.

Stomach Disease	Stomach Diseases				
	Healthy	Bleeding	Esophagitis	Polyps	Ulcerative-Colitis
Healthy	100%				
Bleeding		100%			
Esophagitis			100%		
Polyps			<1%	99%	<1%
Ulcerative-Colitis				<1%	99%

The proposed deep CNN-based stomach disease classification methodology compared with other deep CNN models such as AlexNet, VGG19, VGG16, InceptionV3, and GoogleNet utilized for stomach disease classification and recognition accuracy has been presented in Figure 8. We have also trained our fine-tuned models and performed classification of stomach disease and compared with them. The results show that the pre-trained model accuracy of AlexNet, VGG16, VGG19, GoogleNet, and Resnet50 was 94.67%, 93.96%, 95.9%, 97.5, and 92.17, respectively, while the proposed model attained an accuracy of 99.8%.
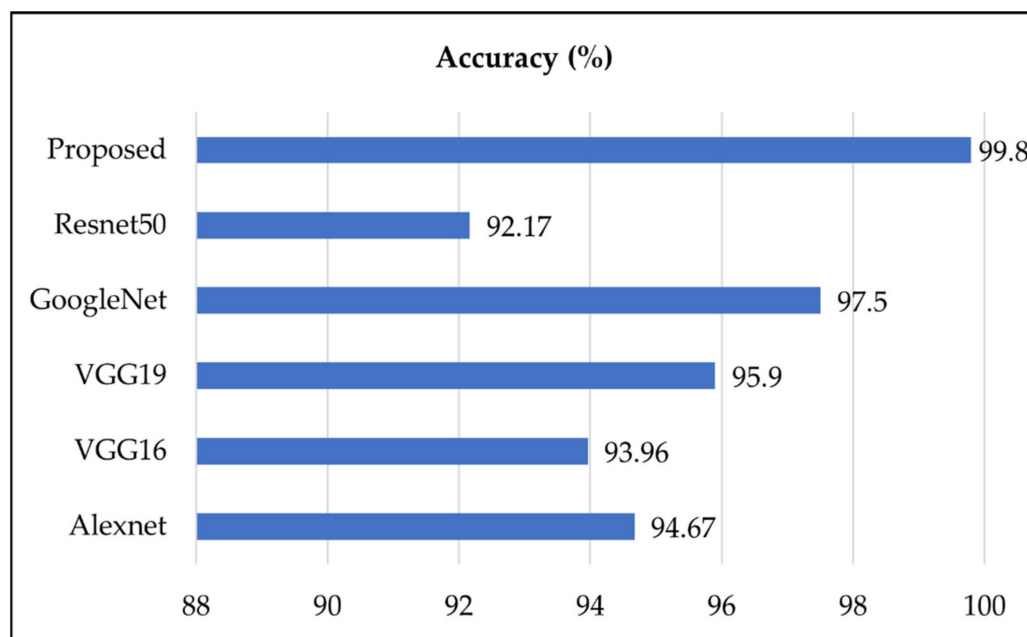
Figure 8. Proposed stomach disease recognition model comparison with CNN models.

5. Discussion

In this work, we performed a fair comparison of our proposed methodology with other pertinent techniques presented for stomach disease recognition. In our proposed features extraction and optimization method, we utilized cross-validation of 10-fold to examine the robustness of our methodology. The maximum accuracy achieved by concatenating deep features spaces was 96.2%, and the deployment of a modified dragonfly algorithm increased the accuracy to 99.8 using C-SVM machine-learning classifiers. The comparison of the proposed feature optimization method with relevant methodologies is presented in Table 5. The techniques that are used for the comparison of the proposed method used the same dataset or same classes of stomach dataset used in this proposed work. Researchers [39] utilized 12,147 stomach disease endoscopic images to extract from Kvasir v2 and endoscopy artifact detection (EAD) [40] for disease recognition, using a deep-learning-based attenuation technique and achieved a classification accuracy of 93.19%. In [41] stomach disease recognition was performed on 6702 images from Kvasir V2, challenging the stomach disease dataset using data augmentation and fine-tuning of CNN models for stomach disease classification with an accuracy of 96.33%. In [42] 2006, capsule endoscopy images acquired from Kiang Wu Hospital used by researchers for the classification of gastric disease using a deep attention model for segmentation to locate the lesion region for accurate recognition of disease recognition and attained 96.76% of stomach disease classification accuracy. In [43] researchers employed CNN and capsule network for stomach disease classification and deformation analysis using a Kvasir V2 dataset having 8000 images with an accuracy of 94.73%. In [44] the authors classified ulcerative colitis from challenging datasets Kvasir, Kvasir V2, and hyper-Kvasir for binary classification using deep learning and achieved an accuracy of 87.50%. Our proposed deep feature extraction and optimization technique for stomach cancer classification was applied on 17,500 images of different stomach disease classes for stomach diseases recognition from Kvasir and Kvasir v2 and two classes of healthy and bleeding obtained from [9] and achieved the highest accuracy of 99.8% on the. cubic SVM classifier. The proposed model also reduced the computational time as compared to modern techniques of stomach infection recognition. We performed experiments on the fused feature vector and compared these results after applying the optimization method. The results clearly show that the proposed binary dragonfly algorithm increases the efficiency of the model as presented in Figure 9.

Table 5. Comparison of the proposed model with existing techniques.

Ref.	Year	Dataset	No. of Images	Accuracy
[39]	2021	Kvasir V2 + EAD2019	12,147	93.19%
[41]	2021	Kvasir V2	6702	96.33%
[42]	2022	Kiang Wu Hospital dataset	2006	96.76%
[43]	2022	Kvasir V2	8000	94.83%
[44]	2022	Kvasir + Kvasir V2	3482	87.50%
Proposed	**2022**	Kvasir + Kvasir V2 + Healthy + Bleeding	**17,500**	**99.8%**

The robustness and consistency of the proposed method were also evaluated using a confidence interval, a statistical method used for the computation of error and continuity exemplified in Figure 10. Normally, in data representation, 5% is used as a statistical significance for the confidence level score of 95% with a standard deviation of 0.08. The presence of error in our proposed method is 99.74 ± 0.118 (±0.12%). The statistical analysis presented in the figure shows that the proposed stomach disease classification method is consistent and accurate at different confidence intervals.
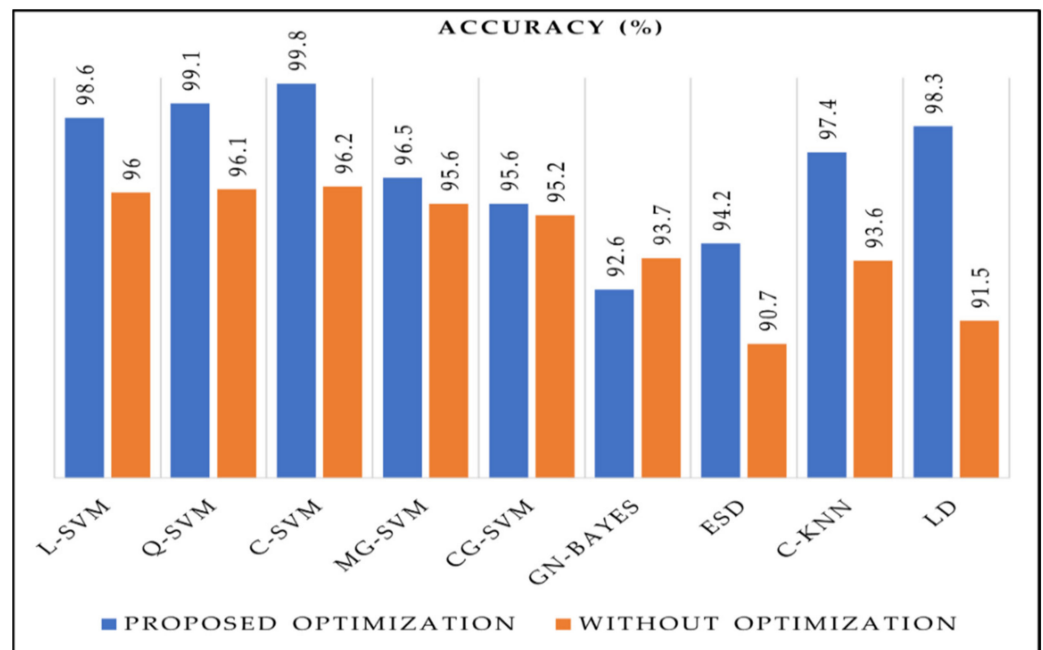
Figure 9. Stomach disease recognition accuracy with and without optimization.

Confidence Level	Margin of Error	Error Bar
68.3%, $\sigma_{\bar{x}}$	99.74 ±0.0358 (±0.04%)	
90%, 1.645$\sigma_{\bar{x}}$	99.74 ±0.0589 (±0.06%)	
95%, 1.960$\sigma_{\bar{x}}$	99.74 ±0.0701 (±0.07%)	
99%, 2.576$\sigma_{\bar{x}}$	99.74 ±0.0922 (±0.09%)	
99.9%, 3.291$\sigma_{\bar{x}}$	99.74 ±0.118 (±0.12%)	
99.99%, 3.891$\sigma_{\bar{x}}$	99.74 ±0.139 (±0.14%)	
99.999%, 4.417$\sigma_{\bar{x}}$	99.74 ±0.158 (±0.16%)	
99.9999%, 4.892$\sigma_{\bar{x}}$	99.74 ±0.175 (±0.18%)	

Figure 10. Statistical confidence interval of proposed stomach classification method.

6. Conclusions

Deep CNN-based feature derivation and optimization methodology have been presented in this paper for stomach disease classification. In the suggested method, the deep CNN features derivation was performed by employing transfer learning using Inception V3 and DenseNet-201, Activation was performed on the feature derivation layers of both models to obtain the specific feature matrix. The derived feature matrix was fused with the help of the parallel maximum covariance method. The maximum accuracy achieved after feature fusion was 96.2% on the C-SVM. The concatenated features matrix was optimized using a modified binary dragonfly algorithm. The features optimizing technique provided better results as compared to the individual deep CNN feature fusion feature map. The maximum recognition accuracy of stomach disease recognition after feature optimization was 99.8% on a C-SVM machine-learning classifier which represents the robustness of the proposed feature optimization technique. The main strength of the proposed method comprised robust feature derivation and selection to enhance the accuracy of stomach diseases recognition. The major disadvantage of the implanted technique is the increase in computational cost due to feature concatenation. Future work will comprise the large

database creation and optimal deep-learning model creation, especially for stomach disease recognition. Furthermore, a deep-learning model will be trained from scratch to perform polyp and ulcer segmentation.

References

1. Sharif, M.; Attique Khan, M.; Rashid, M.; Yasmin, M.; Afza, F.; Tanik, U.J. Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. *J. Exp. Theor. Artif. Intell.* **2021**, *33*, 577–599. [CrossRef]
2. Khan, M.A.; Kadry, S.; Alhaisoni, M.; Nam, Y.; Zhang, Y.; Rajinikanth, V.; Sarfraz, M.S. Computer-aided gastrointestinal diseases analysis from wireless capsule endoscopy: A framework of best features selection. *IEEE Access* **2020**, *8*, 132850–132859. [CrossRef]
3. Lee, J.H.; Kim, Y.J.; Kim, Y.W.; Park, S.; Choi, Y.I.; Kim, Y.J.; Park, D.K.; Kim, K.G.; Chung, J.W. Spotting malignancies from gastric endoscopic images using deep learning. *Surg. Endosc.* **2019**, *33*, 3790–3797. [CrossRef]
4. Ghatwary, N.; Ye, X.; Zolgharni, M.J. Esophageal abnormality detection using densenet based faster r-cnn with gabor features. *IEEE Access* **2019**, *7*, 84374–84385. [CrossRef]
5. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A.J. Cancer statistics, 2021. *CA Cancer J. Clin.* **2021**, *71*, 7–33. [CrossRef]
6. Yabroff, K.R.; Wu, X.-C.; Negoita, S.; Stevens, J.; Coyle, L.; Zhao, J.; Mumphrey, B.J.; Jemal, A.; Ward, K.C. Association of the COVID-19 Pandemic with Patterns of Statewide Cancer Services. *NCI J. Natl. Cancer Inst.* **2021**, djab122. [CrossRef] [PubMed]
7. Masmoudi, Y.; Ramzan, M.; Khan, S.A.; Habib, M. Optimal feature extraction and ulcer classification from WCE image data using deep learning. *Soft Comput.* **2022**, 1–14. [CrossRef]
8. Fan, S.; Xu, L.; Fan, Y.; Wei, K.; Li, L. Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. *Phys. Med. Biol.* **2018**, *63*, 165001. [CrossRef] [PubMed]
9. Khan, M.A.; Rashid, M.; Sharif, M.; Javed, K.; Akram, T.J. Classification of gastrointestinal diseases of stomach from WCE using improved saliency-based method and discriminant features selection. *Multimed. Tools Appl.* **2019**, *78*, 27743–27770. [CrossRef]
10. Charfi, S.; El Ansari, M.J. Computer-aided diagnosis system for colon abnormalities detection in wireless capsule endoscopy images. *Multimed. Tools Appl.* **2018**, *77*, 4047–4064. [CrossRef]
11. Saito, H.; Aoki, T.; Aoyama, K.; Kato, Y.; Tsuboi, A.; Yamada, A.; Fujishiro, M.; Oka, S.; Ishihara, S.; Matsuda, T.J. Automatic detection and classification of protruding lesions in wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointest. Endosc.* **2020**, *92*, 144–151.e1. [CrossRef] [PubMed]
12. Naz, J.; Sharif, M.; Yasmin, M.; Raza, M.; Khan, M.A.J. Detection and Classification of Gastrointestinal Diseases using Machine Learning. *Curr. Med. Imaging* **2021**, *17*, 479–490. [CrossRef] [PubMed]
13. Suman, S.; Hussin, F.A.B.; Malik, A.S.; Pogorelov, K.; Riegler, M.; Ho, S.H.; Hilmi, I.; Goh, K.L. Detection and classification of bleeding region in WCE images using color feature. In Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing, Firenze, Italy, 19–21 June 2017; pp. 1–6.
14. Suman, S.; Hussin, F.A.; Malik, A.S.; Ho, S.H.; Hilmi, I.; Leow, A.H.-R.; Goh, K.-L.J. Feature selection and classification of ulcerated lesions using statistical analysis for WCE images. *Appl. Sci.* **2017**, *7*, 1097. [CrossRef]
15. Li, B.; Meng, M.Q.-H.J. Tumor recognition in wireless capsule endoscopy images using textural features and SVM-based feature selection. *IEEE Trans.* **2012**, *16*, 323–329.
16. Tuba, E.; Tuba, M.; Jovanovic, R. An algorithm for automated segmentation for bleeding detection in endoscopic images. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 4579–4586.
17. Charfi, S.; El Ansari, M. Computer-aided diagnosis system for ulcer detection in wireless capsule endoscopy videos. In Proceedings of the 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Fez, Morocco, 22–24 May 2017; pp. 1–5.

18. Jamil, D. Diagnosis of Gastric Cancer Using Machine Learning Techniques in Healthcare Sector: A Survey. *Informatica* **2022**, *45*, 147–166. [CrossRef]

19. Khan, M.A.; Sarfraz, M.S.; Alhaisoni, M.; Albesher, A.A.; Wang, S.; Ashraf, I.J. StomachNet: Optimal deep learning features fusion for stomach abnormalities classification. *IEEE Access* **2020**, *8*, 197969–197981. [CrossRef]

20. Liaqat, A.; Khan, M.A.; Shah, J.H.; Sharif, M.; Yasmin, M.; Fernandes, S.L. Automated ulcer and bleeding classification from WCE images using multiple features fusion and selection. *J. Mech. Med. Biol.* **2018**, *18*, 1850038. [CrossRef]

21. Zhao, R.; Zhang, R.; Tang, T.; Feng, X.; Li, J.; Liu, Y.; Zhu, R.; Wang, G.; Li, K.; Zhou, W.; et al. TriZ-a rotation-tolerant image feature and its application in endoscope-based disease diagnosis. *Comput. Biol. Med.* **2018**, *99*, 182–190. [CrossRef]

22. Majid, A.; Khan, M.A.; Yasmin, M.; Rehman, A.; Yousafzai, A.; Tariq, U. Classification of stomach infections: A paradigm of convolutional neural network along with classical features fusion and selection. *Microsc. Res. Tech.* **2020**, *83*, 562–576. [CrossRef]

23. Khan, M.A.; Majid, A.; Hussain, N.; Alhaisoni, M.; Zhang, Y.-D.; Kadry, S.; Nam, Y. Multiclass Stomach Diseases Classification Using Deep Learning Features Optimization. *Comput. Mater. Contin.* **2021**, *67*, 3381–3399. [CrossRef]

24. Bora, K.; Bhuyan, M.; Kasugai, K.; Mallik, S.; Zhao, Z. Computational learning of features for automated colonic polyp classification. *Sci. Rep.* **2021**, *11*, 4347. [CrossRef] [PubMed]

25. Ayyaz, M.S.; Lali, M.I.U.; Hussain, M.; Rauf, H.T.; Alouffi, B.; Alyami, H.; Wasti, S. Hybrid Deep Learning Model for Endoscopic Lesion Detection and Classification Using Endoscopy Videos. *Diagnostics* **2022**, *12*, 43. [CrossRef] [PubMed]

26. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

27. Urban, G.; Tripathi, P.; Alkayali, T.; Mittal, M.; Jalali, F.; Karnes, W.; Baldi, P. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* **2018**, *155*, 1069–1078.e8. [CrossRef] [PubMed]

28. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.

29. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef]

30. Khan, M.A.; Sharif, M.; Akram, T.; Yasmin, M.; Nayak, R.S. Stomach deformities recognition using rank-based deep features selection. *J. Med. Syst.* **2019**, *43*, 329. [CrossRef]

31. Billah, M.; Waheed, S.; Rahman, M.M. An automatic gastrointestinal polyp detection system in video endoscopy using fusion of color wavelet and convolutional neural network features. *Int. J. Biomed. Imaging* **2017**, *2017*, 9545920. [CrossRef]

32. Pogorelov, K.; Randel, K.R.; Griwodz, C.; Eskeland, S.L.; de Lange, T.; Johansen, D.; Spampinato, C.; Dang-Nguyen, D.-T.; Lux, M.; Schmidt, P.T. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In Proceedings of the 8th ACM on Multimedia Systems Conference, Taipei Taiwan, 20–23 June 2017; pp. 164–169.

33. Aisu, N.; Miyake, M.; Takeshita, K.; Akiyama, M.; Kawasaki, R.; Kashiwagi, K.; Sakamoto, T.; Oshika, T.; Tsujikawa, A. Regulatory-approved deep learning/machine learning-based medical devices in Japan as of 2020: A systematic review. *PLOS Digit. Health* **2022**, *1*, e0000001. [CrossRef]

34. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.

35. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.

36. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.

37. Khan, M.A.; Alhaisoni, M.; Tariq, U.; Hussain, N.; Majid, A.; Damaševičius, R.; Maskeliūnas, R. COVID-19 case recognition from chest CT images by deep learning, entropy-controlled firefly optimization, and parallel feature fusion. *Sensors* **2021**, *21*, 7286. [CrossRef]

38. Mirjalili, S. Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Comput. Appl.* **2016**, *27*, 1053–1073. [CrossRef]

39. Lonseko, Z.M.; Adjei, P.E.; Du, W.; Luo, C.; Hu, D.; Zhu, L.; Gan, T.; Rao, N. Gastrointestinal Disease Classification in Endoscopic Images Using Attention-Guided Convolutional Neural Networks. *Appl. Sci.* **2021**, *11*, 11136. [CrossRef]

40. Ali, S.; Zhou, F.; Daul, C.; Braden, B.; Bailey, A.; Realdon, S.; East, J.; Wagnieres, G.; Loschenov, V.; Grisan, E. Endoscopy artifact detection (EAD 2019) challenge dataset. *arXiv* **2019**, arXiv:1905.03209.

41. Yogapriya, J.; Chandran, V.; Sumithra, M.; Anitha, P.; Jenopaul, P.; Suresh Gnana Dhas, C. Gastrointestinal tract disease classification from wireless endoscopy images using pretrained deep learning model. *Comput. Math. Methods Med.* **2021**, *2021*, 5940433. [CrossRef] [PubMed]

42. Yu, X.; Tang, S.; Cheang, C.F.; Yu, H.H.; Choi, I.C. Multi-Task Model for Esophageal Lesion Analysis Using Endoscopic Images: Classification with Image Retrieval and Segmentation with Attention. *Sensors* **2022**, *22*, 283. [CrossRef] [PubMed]

43. Wang, W.; Yang, X.; Li, X.; Tang, J. Convolutional-capsule network for gastrointestinal endoscopy image classification. *Int. J. Intell. Syst.* **2022**. [CrossRef]

44. Sutton, R.T.; Zaïane, O.R.; Goebel, R.; Baumgart, D.C. Artificial intelligence enabled automated diagnosis and grading of ulcerative colitis endoscopy images. *Sci. Rep.* **2022**, *12*, 2748. [CrossRef]