CHAPTER-4 AN EFFECTIVE APPROACH FOR SEGMENTATION OF MULTIPLE ABNORMALITIES FROM CAPSULE ENDOSCOPY IMAGES

In this chapter, we propose a Gaussian mixture model (GMM) and improve Expectation-Maximization (EM) based segmentation approach. Unlike other approaches observed in the literature, the proposed approach is capable of segmenting any normal or abnormal CE image instead of just one particular abnormality. This chapter further demonstrates the efficacy of the segmentation method in the post-processing part. Here, we present a backward elimination based feature selection algorithm, a fusion of feature method to reduce information loss and a multi-class classifier system for CE images.

4.1 Introduction

Segmentation helps us focus only on the region of interest (RoI) instead of the entire image. Since the CE videos are lengthy, producing approximately 55000 to 60000 frames, segmentation plays a crucial role. Largely these solutions depend on the labeled data. This chapter presents a generic and automatic segmentation technique. The term generic signifies that this same method is capable of segmenting any CE image. It can take any normal or abnormal CE image and present the segmented image with RoI. Following are the core contributions of this chapter:

- A generic and automatic segmentation technique.
- Feature selection and fusion of features to improve system accuracy.
- Comparative analysis of the proposed system on different classifiers.

4.2 Methods and Models

Figure 4.1 shows a schematic diagram of the entire proposed system. As seen in Figure 4.2, the CAD system sequences through pre-processing, segmentation, feature extraction including selection & fusion and finally leading to classification.

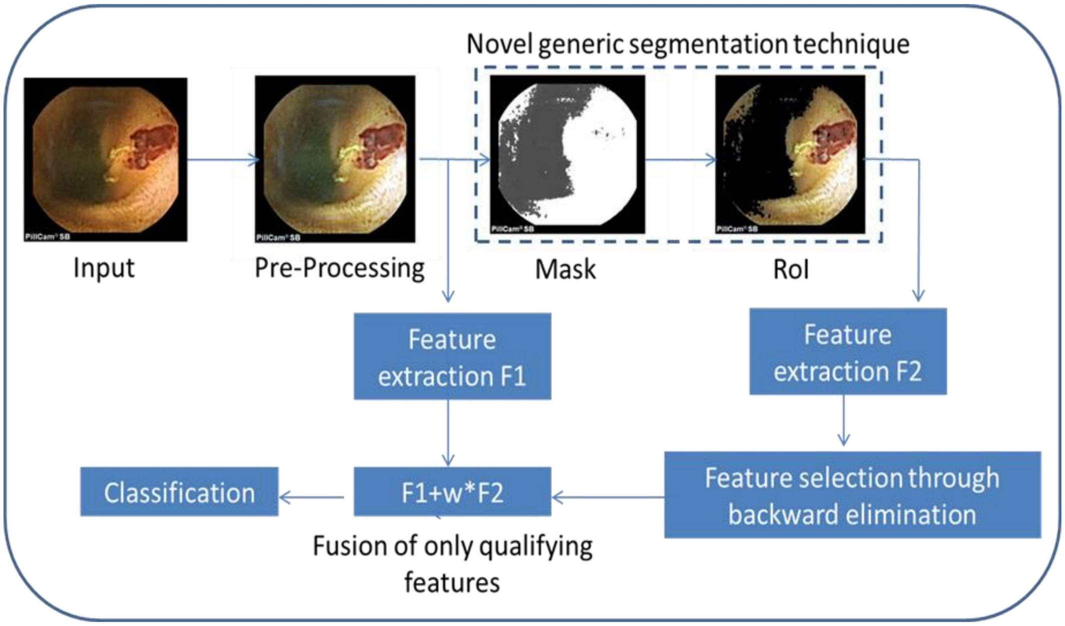


Figure 4.1: Schematic diagram of the entire proposed system

4.2.1 Pre-Processing using Contrast Stretching

Here, contrast stretching is used to improve the contrast of the input CE images. This technique stretches the range of intensity values to the desired range of values [104]. It saturates top and bottom 1% of pixel intensity and maps the intensity of a given image to the desired image using (4.1).

$$s = (r - c)\left(\frac{b - a}{d - c}\right) + a \tag{4.1}$$

where, r is original pixel value, s is the output value of the pixel, a is new min value, b is new max value, c is old min value and d is the old max value of intensity. This is a linear mapping procedure to enhance the contrast.

4.2.2 Segmentation using Improved GMM

One can categorize regions of a CE image as outer region, non-informative region, informative normal region, and informative abnormal region. Our aim here is to segment out the informative region automatically. In particular, the outer region and non-informative inner region are to be omitted. Automatic segmentation of CE images faces the following challenges:

- Lack of huge amount of labeled ground truth data.
- Low contrast images.
- Complex nature of the GI tract.
- Multiple local minima.
- Non-guarantee of the closed contour.

This study presents a segmentation technique comprising of the Gaussian mixture model (GMM) and improved expectation–maximization (EM). The algorithm design criterion kept to overcome challenges as mentioned above and ensure generic and automatic segmentation of CE images are as follows:

- The initial value of the parameters of the objective function is kept high.
- Reduce parameters of the objective function at a slow reduction rate to ensure heuristically obtained range.
- Annihilate and penalize if the range is missed.

In real life, many applications can be modelled by Gaussian distribution univariate or multivariate. It is very intuitive to assume that the clusters come from different Gaussian Distributions. Or in other words, it is tried to model the dataset as a mixture of several Gaussian Distributions. This is the core idea of this model. In one dimension the probability density function of a Gaussian distribution is given by equation (4.2)

$$G(X|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{4.2}$$

Where μ and σ^2 are respectively mean and variance of the distribution. For Multivariate (say n-variate) Gaussian distribution, the probability density function is given by equation (4.3)

$$G(X|\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)|\Sigma|}} e^{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)} \tag{4.3}$$

Where μ is the n-dimensional vector denoting the mean of the distribution and Σ is the n x n covariance matrix.

GMM is a natural model to utilize in problems such as CE which contains multiple distinct sub-classes. Existing methods to obtain the parameters of Gaussian distribution are maximum likelihood estimation, posterior probability, and EM. This study presents improved EM such that it does not stuck in local minima and ensures a proper estimation of Gaussian parameters leading to desired segmentation.

The EM algorithm is an iterative way to find maximum-likelihood estimates for model parameters. EM chooses some random values for the missing data points and estimates a new set of data. These new values are then recursively used to estimate a better first data, by filling up missing points, until the values get fixed. There are the two basic steps of the EM algorithm, namely E Step or Expectation Step or Estimation Step and M Step or Maximization Step. In the “Expectation” step, we will calculate the probability that each data point belongs to each cluster (using our current estimated mean vectors and covariance matrices). In the "Maximization" step, we'll re-calculate the cluster means and covariances based on the probabilities calculated in the expectation step.

A CE image essentially has four sub-regions thus; we have kept the number of Gaussian components as four. Further, EM is an optimization strategy for an objective function that is interpreted as a likelihood. We intend to minimize the defined objective function iteratively. For all elements of the image, a membership probability is computed concerning all Gaussian components by (4.4).

$$p_j(X|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n|\Sigma_j|}} e^{-\frac{1}{2}(X-\mu_j)^T\Sigma_j^{-1}(X-\mu_j)} \tag{4.4}$$

Where $p_j(X)$ is the PDF of multivariate Gaussian for cluster j; the probability of this Gaussian producing the input X, j is the cluster number, X is the input vector, n is the input vector length, $\sum_j$ is the nxn covariance matrix for cluster j, $|\sum_j|$ is the determinant of the covariance matrix and $\sum_j^{-1}$ is the inverse of the covariance matrix. The objective function to be minimized is defined by equation (4.5).

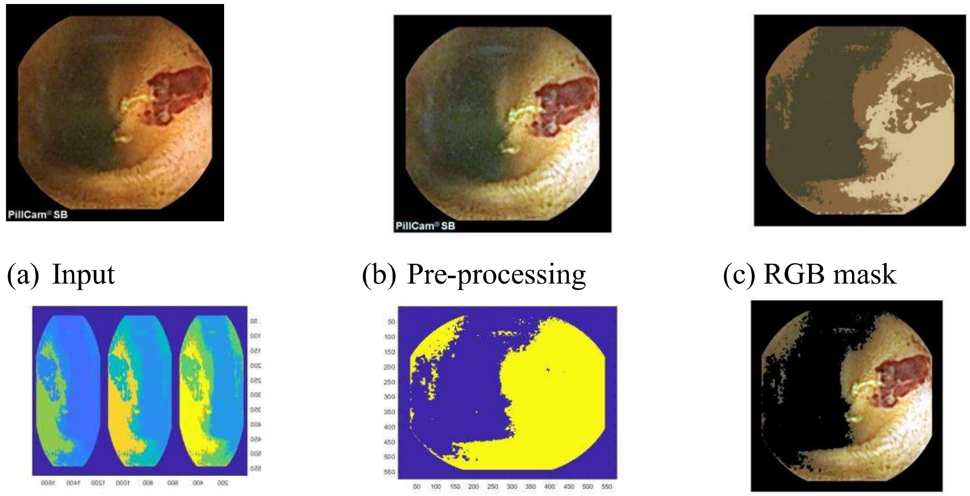
$$Energy = \varepsilon + \sum(u - u_0)^2 + \sum(v - v_0)^2 + \sum(w - w_0)^2 \tag{4.5}$$

where ε is the positive value, u is the estimated mean of Gaussian and v is the estimated standard deviation of the Gaussian and w is the weight parameter.

The ε is slowly reduced while ensuring that it does not become negative. The reduction rate of 0.0125 is chosen so that the expected range is likely to meet. However, ε is not reduced further if it gets less than zero. Algorithm also takes care of a special case where the Energy gets non-computable. In such scenario the re-initialization of all values is done thereby, the process restarts. In case a divergence is found then one step rollback is performed and a penalty of 0.5 is imposed. The proposed generic segmentation algorithm 4.1 is as follows and Figure 4.2 presents stepwise visual results of the same.

Algorithm 4.1: Generic segmentation algorithm for CE images

Input: CE color Image I, no. of Gaussian components j =4
Output: Assignment matrix, mask
Initialize mean u = random[1,n] from image vector, standard deviation v = standard deviation(image vector),weight w=1/j, assignment matrix p=0, u0=0 ,v0=0, w0=1.5 and ε =0.075 n= no. of rows * no. of columns of I
$Energy = \varepsilon + \sum(u - u_0)^2 + \sum(v - v_0)^2 + \sum(w - w_0)^2$
while (Energy between 0.0025 and 0.009)
for j= 1 to 4
calculate p as per equation (4.4)
end for
save current u,v,w as u0,v0,w0 respectively
update u,v,w
if(ε - 0.0125 > 0) // to insure that ε does not become negative
ε = ε -0.0125
end if
$New\_Energy = \varepsilon + \sum(u - u_0)^2 + \sum(v - v_0)^2 + \sum(w - w_0)^2$
if(New_Energy > Energy) // roll back u,v,w values and penalise
u=u0,v=v0,w=w0
$Energy = 0.5 + \varepsilon + \sum(u - u_0)^2 + \sum(v - v_0)^2 + \sum(w - w_0)^2$ //penalty of 0.5
else
update Energy = New_Energy
end if
if (Energy = NaN) // special case where Energy gets non-computable; restart
u0=0,v0=0,w0 =1.5, ε = 0.075
$Energy = \varepsilon + \sum(u - u_0)^2 + \sum(v - v_0)^2 + \sum(w - w_0)^2$
end if
end while
Obtain RGB mask
Extract R channel from RGB mask
Get binary mask from R channel components greater than threshold 0.4
Obtain a segmented image by masking the input image with an obtained mask



(a) Input (b) Pre-processing (c) RGB mask

(d) Mask colour map (e) R-Channel with threshold (f) Output

Figure 4.2: Stepwise visual results of algorithm 4.1 from (a) to (f)

As seen in Figure 4.2, the input image is pre-processed to improve the contrast after which the proposed segmentation technique is applied. As an output of the segmentation technique, a membership probability of each pixel is obtained. The interim output is referred to as an RGB mask. Figure 4.2(d) shows three different channels red, green, and blue. Close observation of R channel components shows that a certain threshold value can lead to a precise mask for segmenting out the RoI. A threshold of 0.4 is applied to obtain the R channel mask. Finally, this R channel mask is converted to a binary mask. RoI is obtained by masking the input image with the final binary mask, as seen in Figure 4.2(f).

4.2.3 Results and Analysis of Proposed Segmentation Method

4.2.3.1 Dataset

A total of 700 images from CE videos [9] is extracted out of which 100 images are of angioectasia, 200 images are of ulcers, 200 images are of bleeding, and 200 images are normal. The dimension of each image is 576x576 pixels. All the images were manually diagnosed and annotated by physicians providing the ground truth.

4.2.3.2 Performance Metrics

The performance of the proposed segmentation technique is evaluated quantitative metrics namely Global Consistency Error (GCE), Probabilistic Rand Index (PRI) and Variation of Information (VoI) as given by [105].

GCE for segments s_i and g_j contain a pixel, say p_k, such that $s \epsilon S$, $g \epsilon G$ where S denotes the set of segments that are generated by the segmentation algorithm being evaluated and G denotes the set of reference segments. To begin with, a measure of local refinement error is estimated using equation (4.6) and then it is used to compute local and global consistency errors, where n denotes the set difference operation and R(x,y) represents the set of pixels corresponding to region x that includes pixel y. Using equation (4.6) the GCE is computed using equation (4.7) where n denotes the total number of pixels of the image.

$$E(s_i, g_j, p_k) = \frac{\left|\frac{R(s_i, p_k)}{R(\,g_j, p_k)}\right|}{|R(s_i, p_k)|} \tag{4.6}$$

$$GCE(S, G) = \frac{1}{n}\min\left(\sum_i E(S, G, p_i), \sum_i E(G, S, p_i)\right) \tag{4.7}$$

PRI is the non-parametric measure of goodness of segmentation algorithms. PRI between test (S) and reference (G) is estimated by summing the number of pixel pairs with the same label and number of pixel pairs having different labels in both S and G and then dividing it by the total number of pixel pairs. Given a set of reference segmentations G_k, the PRI is estimated using equation (4.8) such that c_{ij} is an event that describes a pixel pair (i,j) has same or different label in the test image S_{test} and p_{ij} is probability that a pixel pair (i,j) has same or different label in test image S_{test}

$$PRI(S_{test}, G_k) = \frac{1}{n/2} \sum_{\forall i,j\&i<j} [\, c_{ij}p_{ij} + (1 - c_{ij}) + (1 - p_{ij})\,] \tag{4.8}$$

VoI measures the sum of information loss and information gain between the two clustering, and thus it roughly measures the extent to which one clustering can explain the other. The VoI metric is nonnegative, with lower values indicating greater similarity.

$$VoI(A, B) = H(A) = H(B) - 2I(A, B) \tag{4.9}$$

where, $H(A)$ is the entropy of A, $H(B)$ is the entropy of B and $I(A, B)$ is the mutual information between A and B.

GCE measures the extent to which a segment is a refinement of others. VoI is the nonnegative measure that indicated the sum of information loss and gain. PRI gives a measure of similarity between two segments. For better performance, PRI should be higher while GCE and VoI should be lower.

4.2.3.3 Results

The results of proposed approach is compared with four other methods namely Otsu [106], Fuzzy c mean [107], Color Based [108] and Saliency Based [109]. Figure 4.3 presents the visual results of different CE image segmentation methods. Table 4.1 presents a comparison between proposed segmentation technique on quantitative measures and Figure 4.4 and Figure 4.5 presents a graphical representation of the results.

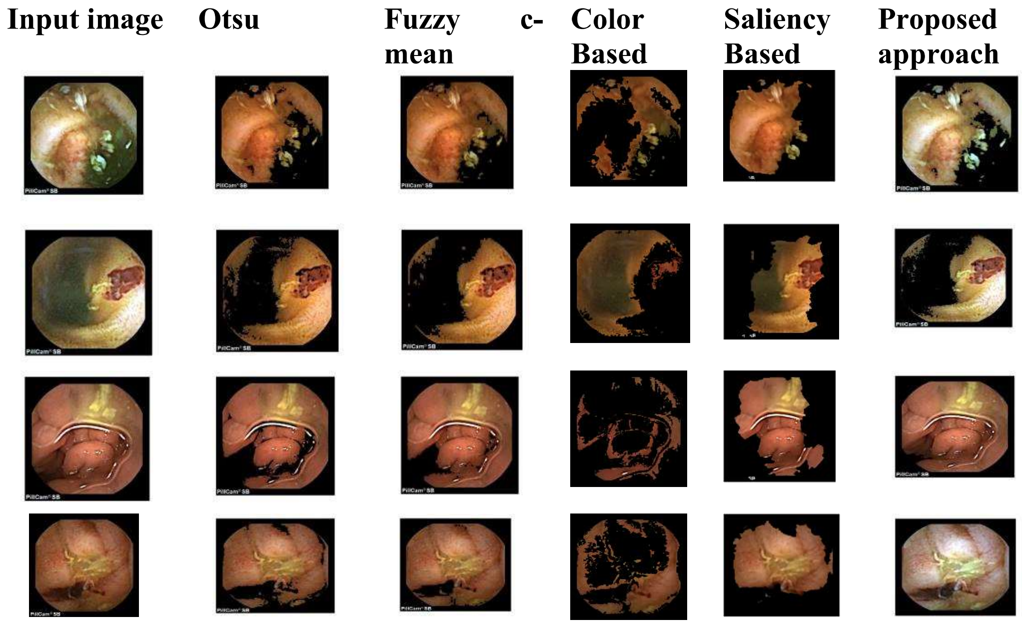


Figure 4.3: Visual results of different CE image segmentation methods

Table 4.1 Quantitative performance evaluation of segmentation techniques

	Otsu [106]	Fuzzy c mean [107]	Color Based [108]	Saliency Based [109]	Proposed
GCE	0.331225	0.386025	0.238925	0.538325	0.018575
PRI	0.837775	0.8845	0.5728	0.745125	0.893075
VoI	3.621375	3.497025	3.852425	6.876375	1.0708

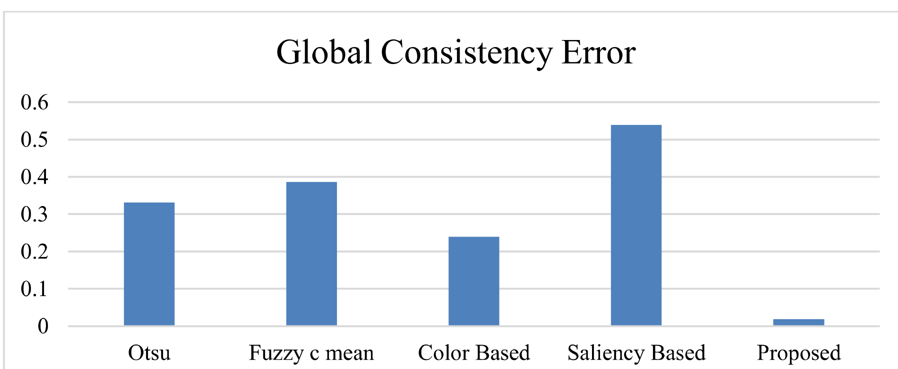


Figure 4.4: Graphical representation of GCI results

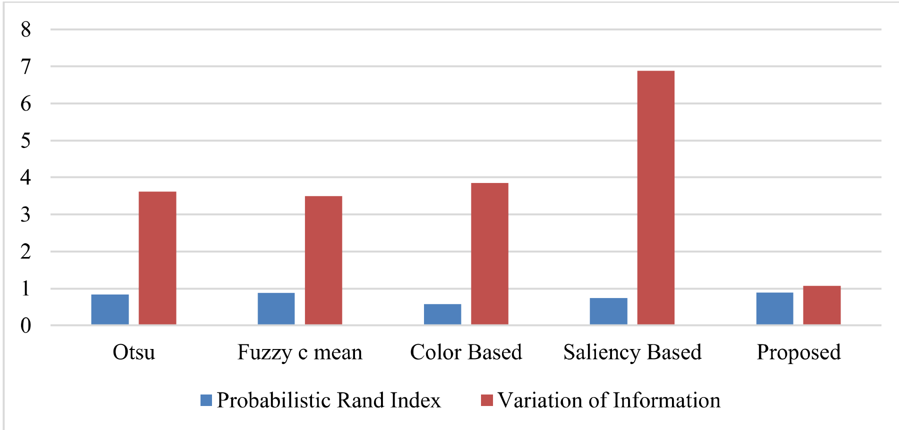


Figure 4.5: Graphical representation of PRI and VoI results

As seen through the visual and quantitative results the proposed segmentation technique outperforms the other four techniques. Further, this chapter throws some light upon the post-processing part after the segmentation.

4.3 Evaluation of Proposed Segmentation Approach in Classification of Abnormalities

The objective here is to evaluate the efficacy of the proposed segmentation approach for the automatic abnormality detection system in CE. Essentially we take the study further to the post-processing part where a feature selection, feature fusion and multi-class classification is presented.

4.3.1 Feature Extraction, Selection, and Fusion

One CE image produces 321 features by combining all the above set of features. The preliminary experiment on the classification of abnormalities using all 321 features yields an accuracy of less than 85%. This experiment further leads to two important observations: (1) some of the features are spoiling the learning accuracy which needs to be eliminated, and (2) some of the vital information is lost during the segmentation processes leading to the poor performance of the system. To overcome these shortcomings, we further propose a backward elimination feature selection algorithm and a fusion of feature technique. The algorithm 4.2 for feature selection is as follows:

Algorithm 4.2: Feature selection using backward elimination

---

1. Input: Entire feature vector F; Group of features G such that $\{G1 \cup G2 \cup G3 \cup \ldots Gn\} = F$ and $Gi \cap Gj = \phi$ where i,j = 1 to n and $i \neq j$
2. Output: Reduced feature set Fr

3. Initialize master feature set F containing all 321 features
4. Initialize reduced feature set Fr = F
5. Calculate classification accuracy Ai for every Gi
6. Calculate accuracy AF for superset F
7. i=1
8. while (all feature groups are not evaluated)
9. temp_index = min Ai // identify individual worst performers
10. temp = Gtemp_index // identify corresponding Gi for worst performer Ai
11. tempF = F – temp
12. compute system accuracy AtempF for feature set tempF
13. if (AtempF <= AF)
14. Fr = F – tempF // omit the non-performer group of feature
15. end if
16. i++
17. end while

By executing algorithm 4.2, the redundant or non-performer features are omitted from the superset resulting in the final feature set Fr containing 192 features. The following are the feature groups surviving the elimination algorithm: Local binary patterns (LBP), color moments from H component of HSI color space as well as G and B components from RGB color space, color coherence vector (CCV), mean squared energy and mean amplitude. Color moments for H, G, and B components are computed as it provides a measurement for the color similarity between images [110]. Statistical features namely mean standard deviation, kurtosis and, skewness for each component are computed. CCV shows coherent pixels versus incoherent pixels for each color [111]. CCV consists of 27 different colors, and thus, it will create a feature of 54 values indicating coherent and noncoherent pixels for each color. Gabor features have been successful in many image processing and computer vision applications. It extracts local pieces of information which are then combined to recognize an object or region of interest [112]. Here we have utilized two features from Gabor response: mean-squared energy & mean amplitude. The final feature vector is then obtained by combining all features and selecting features with variance greater than zero. Variance greater than zero indicates that there is at least some discrimination between features of different observations. This final feature set reduces to 168 features.

To compensate for the information loss during the segmentation process, we propose a fusion of feature technique. Here, we fuse the surviving features from a non-segmented image with the segmented image in such a way that dimensions of the feature vector does not increase. In this way, the total number of features remains the same, while some useful information is restored. The formula for the fusion of feature is (4.6).

$$F_F = F_1 + w * F_2 \tag{4.6}$$

where, F_F is the final feature set, F_1 is the feature set from the un-segmented image, F_2 is the feature set from RoI and w is the weight. Here, the value of w is set to 1.

The sections present the details on the dataset and results of the proposed algorithm is compared on six criteria with three classifiers.

4.3.2 Results

The system is implemented on the Dell Optiplex 9010 machine with an Intel Core i7 processor and 6GB RAM using MATLAB R2017a. Table 4.2 summarizes the performance of the proposed system, and Figure 4.6 shows a graphical representation of the same. To validate the obtained results and avoid overfitting, we have used a ten-fold cross-validation technique.

Table 4.2: Comparative analysis of the proposed system on different classifiers

Classifier	Accuracy	Sensitivity	Specificity	Precision	F-measure	MCC
KNN	84.43	80.75	94.69	83.93	81.58	76.99
Ensemble	89.45	88.75	96.44	89.05	88.80	85.32
SVM	92.86	91.75	97.57	92.74	92.16	89.83

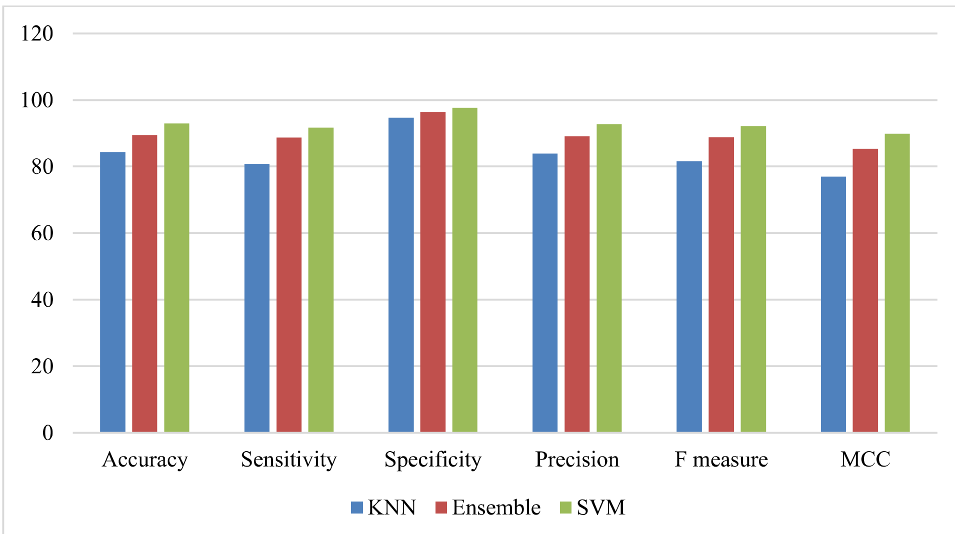


Figure 4.6: A comparative performance analysis of CAD system in graphical form

As seen in Figure 4.6, the SVM classifier outperforms the other two classifiers with an accuracy of 92.86%. The system as a whole is capable of retaining minute details which are segmented out by other techniques.

4.4 Conclusions

The proposed segmentation approach is capable of precisely differentiating between the red regions and dark regions. This ability is the biggest advantage of the proposed segmentation technique as compared to other methods. This fact is reflected in the sample visual results of various techniques. This study also highlights the following observation and properties for the proposed segmentation technique. The EM sequence increases the likelihood and converges to some solution; however; it is difficult to verify that the solution is not trapped at local maxima or a stationary point. To ensure a solution avoiding the above condition, we have incorporated the following aspects:

- The stationary point is taken care of as a special case where the algorithm reaches a non-computational value. In such a scenario, the process is re-started.
- Trapping into a local solution is avoided by ensuring a slow monotonous reduction in energy function, annihilate and penalize in case of a violation.
- For all well-separated Gaussian distributions proposed solution converges to maximum likelihood.

Various methods available for segmentation of CE images target only a specific abnormality such a polyp or bleeding but, the proposed method is capable of segmenting the multiple abnormalities with one generalized approach.

Further, the efficacy was also examined for the classification of multiple abnormalities for CE images using feature extraction and classification techniques. Based on the results and analysis it was observed that the proposed segmentation approach effectively segments all types of abnormalities and it is found to be a suitable method for abnormality detection from CE images.