Application of Modified Ant Colony Optimization for Computer Aided Bleeding Detection System

Shahed K. Mohammed, Farah Deeba, Francis M. Bui, and Khan A. Wahid.
Department of Electrical and Computer Engineering
University of Saskatchewan
Saskatoon, Canada

***Abstract*— Wireless capsule endoscopy (WCE) plays a significant role in the non-invasive small intestine screening for obscure gastrointestinal bleeding detection. However, the task of reviewing 60,000 frames to detect the bleeding encumbers the clinician, leading to visual fatigue and false diagnosis. In this paper, we propose a color feature based bleeding detection system with feature selection using a modified ant colony optimization (MACO) algorithm. We have utilized the feature selection capability of MACO algorithm to find the optimum feature subset over the color space of RGB and HSV, which provided a classifier that outperforms the classifier formed from RGB and HSV features individually. Comprehensive experimental results reveal that the proposed MACO algorithm can detect the optimal feature subset with performance comparable to exhaustive search in case of individual classifier from RGB and HSV requiring 2% of the computational time compared to exhaustive search. The comparative study of feature selection showed that MACO can provide the most relevant features and improve the performance in terms of accuracy, sensitivity and computational time.**

Keywords—Wireless Capsule Endoscopy, Ant Colony Optimization, Support Vector Machine, Feature Selection.

I. INTRODUCTION

The invention of wireless capsule endoscopy (WCE) system [1], the capsule-sized camera to capture images of the entire gastrointestinal (GI) tract, has enabled the gastroenterologists to monitor the entire small bowel in a patient friendly and non-invasive manner, primarily to detect obscure gastrointestinal bleeding [2]. GI Bleeding in the digestive tract indicates the possibility of a range of diseases such as angiodysplasia, colon cancer, Crohn's disease, esophageal cancer [3], [4]. However, the arduous task of monitoring 60,000 endoscopic frames for one patient often causes visual fatigue, resulting in missing of important diagnostic information [5]. With the advent of computer-aided decision system, automatic bleeding detection can reduce the burden of gastroenterologists and increase the diagnostic yield of WCE.

Keeping the importance of bleeding detection in various pathologies in mind, Given Imaging Ltd. has marketed suspected bleeding indicator (SBI), a software tool for the automatic hemorrhage detection in WCE video frames [6]. SBI's suboptimal performance has motivated several investigations in automated bleeding detection. In these studies, color has been the most utilized feature due to the distinct red hue of bleeding [7]. Different researchers have employed different color spaces for bleeding detection, e.g. RGB [8], [9], CIELab [5], YIQ [10], CMYK [10], HSI [11]. These works have demonstrated that different classifiers have a varying performance depending on the testing dataset indicating that the combination of features from different color space can improve the performance of the classification. To the best of our knowledge, there have been very limited reported works on classifier with features drawn from different color spaces. One of the underlying reasons is the curse of dimensionality introduced by the merging of features extracted from various color planes. For example, in [9], we have presented a classifier working on RGB color plane by extracting 15 features from RGB plane. Following the same procedure, if we want to merge two classifiers using RGB and HSV color-plane, we have to extract 30 features, translating to training and comparing the classifier for $2^{30}-1$ combinations to find the optimum feature subset using the exhaustive search (ES) method. An efficient feature selection method can resolve this issue by selecting the appropriate feature subset from a large feature set, with less computational effort.

To this end, there are several feature subset selection methods in the literature which can be employed to determine the optimal feature subset [12]. Sequential forward search (SFS) and sequential backward search (SBS) are two feature selection methods based on selection or dropping of certain features step by step [12]. SFS and SBS provide a fast approximation of the best feature subset, however often cause selection of suboptimal features set. On the other hand, stochastic feature selection methods based on evolutionary computation algorithms such as Ant Colony Optimization (ACO) [13], Binary Genetic Algorithm (BGA) [14] and Random Subset Feature Selection (RSFS) [15] have the potential to select the optimal feature set with comparable performance to ES.

In this paper, we have proposed a modified ant colony optimization (MACO) algorithm for finding the best feature subset for bleeding detection in the endoscopic image. The algorithm differs from the previous implementation of ACO in two prospects. Firstly, we have proposed a new heuristic information function to resolve the failure of correlation values to reflect the goodness of the color features. The heuristic information based on classifier performance provides a better starting point for the random searching procedure of the ant colony optimization. Secondly, to capture the higher importance of the sensitivity over specificity in bleeding

978-1-5090-0620-5/16/$31.00

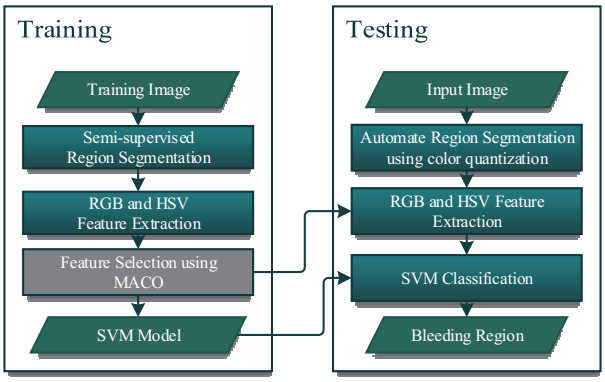

Fig 1. Flowchart of the Proposed Bleeding Detection Algorithm based on MACO

detection, sensitivity has been included in the evaluation function along with accuracy and feature subset length. Thus, MACO can find the optimal feature subset with highest accuracy and sensitivity. To evaluate the performance of MACO, first, its performance was compared with ES method for determining the optimum feature subset from 15 features taken from RGB color space (RGB classifier). Similarly, MACO was also compared with ES in the search for the best feature subset from 15 features taken from HSV color space (HSV classifier). Then we applied the feature selection algorithm to find the best feature subset among 30 features taken from RGB and HSV images or color space to model the RGB-HSV classifier. The results demonstrated that ACO can detect the optimum feature subset while reducing the computational time and increasing the classification accuracy in comparison to other feature selection algorithms. The improvement of the performance of the RGB-HSV classifier over the RGB classifier and HSV classifier also strengthened the implication that by combining the features from different color spaces, the bleeding detection system can be made more robust and accurate.

Organization of the rest of the paper is as follows: Section II describes the proposed bleeding detection algorithm, Section III introduces the modified ant colony optimization algorithm for feature selection, and Section IV demonstrates the experimental setup and result. Finally, in Section V, we draw the conclusion of the paper.

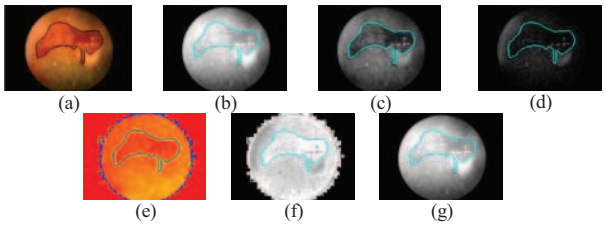

Fig 2. The bleeding in different color map (the line annotates the bleeding region), (a) Original Image (b) R-plane, (c) G-plane, (d) B-plane, (e) Hue (Using Hue Color Map), (f) Saturation, (g) Value

II. Bleeding Detection Using SVM Classifier

A flowchart of the proposed bleeding detection, which is a modified version of the algorithm proposed in [9], is provided in Fig. 1, The main modification is in the process of feature selection as described in detail in Section III (highlighted in gray in Fig. 1). A brief description of these blocks is given in the following subsections.

A. Color Space

The optimum choice of color space depends mostly on the particular application. In this paper, we have chosen RGB and HSV color spaces. As the camera sensor outputs the image in RGB color space, it constitutes a very convenient color space. We have also utilized HSV color space since the distinct red hue in the bleeding region is more distinguishable in HSV color space as it describes colors with intuitive values [7]. Fig. 2 demonstrates the bleeding in different color planes in RGB and HSV.

B. Region Segmentation

In the training stage, a semi-automated region segmentation based on region growing algorithm was utilized to determine the bleeding area. As described in [9], at first, the user provides some bleeding points as the manual seeds to the program. A region growing algorithm grows the regions from these seeds using a distance measure from HSV color space, where the points at a distance less than certain thresholding value are considered as the same region. The distance measured has two components, chrominance d_c measured using (1) and luminance d_v measured using (2) from the HSV value of the two points say $X_1(H_1,S_1,V_1)$ and $X_2(H_2,S_2,V_2)$.

$$d_c = \sqrt{S_1^2 + S_2^2 - 2S_1S_2\cos\theta} \tag{1}$$

where,

$$\theta = \begin{cases} |H_1 - H_2| & if\ |H_1 - H_2| < 180° \\ 360 - |H_1 - H_2| & if\ |H_1 - H_2| > 180° \end{cases}$$

$$d_V = |V_1 - V_2| \tag{2}$$

The distance is then measured using (3):

$$d = \sqrt{(d_c)^2 + (d_v)^2} \tag{3}$$

C. Feature Extraction

In the next step, from each bleeding and non-bleeding regions, 30 features were extracted by calculating the mean, standard deviation (STD), entropy, skew and energy of each color plane in RGB and HSV of those regions. Table-I listed the features considered in this work.

To evaluate the feature selection capability of the modified ACO algorithm, we have first used the F_1-F_{15} features and F_{16}-F_{30} features separately, to train RGB classifier and HSV classifier respectively. Later, we have used all the 30 features

TABLE I. FEATURES EXTRACTED FROM EACH REGIONS

F_1	R-Mean	F_4	R-STD[a]	F_7	R-Entropy	F_{10}	R-Skew	F_{13}	R-Energy
F_2	G-Mean	F_5	G-STD[a]	F_8	G-Entropy	F_{11}	G-Skew	F_{14}	G-Energy
F_3	B-Mean	F_6	B-STD[a]	F_9	B-Entropy	F_{12}	B-Skew	F_{15}	B-Energy
F_{16}	H-Mean	F_{19}	H-STD[a]	F_{22}	H-Entropy	F_{25}	H-Skew	F_{28}	H-Energy
F_{17}	S-Mean	F_{20}	S-STD[a]	F_{23}	S-Entropy	F_{26}	S-Skew	F_{29}	S-Energy
F_{18}	V-Mean	F_{21}	V-STD[a]	F_{24}	V-Entropy	F_{27}	V-Skew	F_{30}	V-Energy

[a] Standard Deviation

for feature selection and optimization of the RGB-HSV classifier with 30 features.

D. SVM Classification

Support vector machine (SVM) has shown relatively lower sensitivity towards sample size and feature selection. The ability of SVM to adapt even using high-dimensional features with a small number of training samples makes SVM ideal for anomaly detection in WCE [4]. With our small training dataset and high dimensional features space, we have chosen SVM for classification. In this model, we have utilized the radial basis function (RBF) SVM, which incorporates translating the feature vector into higher dimensional space where the feature vectors are linearly separable. In RBF, the classification problem translates into determining the decision function $g(x)$ using the (4):

$$g(x) = \sum_{i=1}^{N} \lambda_i y_i k(x, x_i) + w_o \tag{4}$$

Here, λ_i represents the Lagrange multipliers, which are non-zero for only the support vectors. w_o is known as bias. x_i is the training vector and y_i is their corresponding label. $k(x, x_i)$ is the RBF function, which is defined as:

$$k(x, y) = \exp\left(-\frac{\|x-y\|^2}{\sigma^2}\right) \tag{5}$$

Here, σ is a user-defined kernel scale parameter.

E. Feature Selection and Classifier Optimization

As demonstrated in previous works [9], feature selection can reduce the feature dimension. Simultaneously, an optimal feature subset can discard irrelevant features as well as increase the classification performance regarding the computational time and classification performance. Following the previous work [9], ES can give an optimum feature subset with reduced dimension and best accuracy. However, conducting ES over a set of 30 features requires an extensive computational time and resources for optimizing the classifier for approximately 2^{30} combinations. Therefore, we have proposed a modified version of the ant colony optimization (MACO) algorithm to resolve the curse of dimensionality and find the best feature subset. Section III provides a detailed description of the feature selection using MACO.

F. Region Classification

After the model is derived, the testing procedure works in the manner as described in [9]. A color quantization step reduces the color component to 24 and all the pixels with similar intensity values are grouped into one region. Each region is classified as bleeding and non-bleeding region using the SVM model. If the classification model classifies one of the areas in a frame as a bleeding region, the frame is classified as a bleeding frame [9].

III. Modified Ant Colony Optimization

Inspired by ant's foraging behavior, ACO was introduced to solve the well-known Travelling Salesman Problem [13]. Motivated by its meta-heuristic approach in solving the graph minimization problem, several researchers utilized ACO to search the optimum feature subset [16], [17]. To cope with the bleeding detection problem, we have proposed two changes in the ACO algorithm. We have introduced a new evaluation function (described in section III.D) to reflect the higher cost of the low sensitivity. We have also utilized a more informative prior information as the heuristic function to govern the probability between the nodes (described in Section III.B). Fig. 3 demonstrates the flowchart of the proposed feature selection algorithm.

A. Feature Space

First, the feature space is considered as a graph plot, where each node represents a feature. The feature subset is the path between the features. The connection between each node is represented as a probability function P_{ij} (shown in Fig. 4). P_{ij}

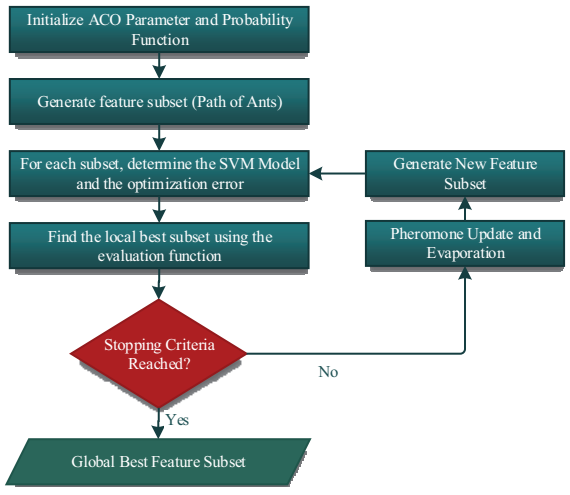

Fig 3. Flowchart of MACO

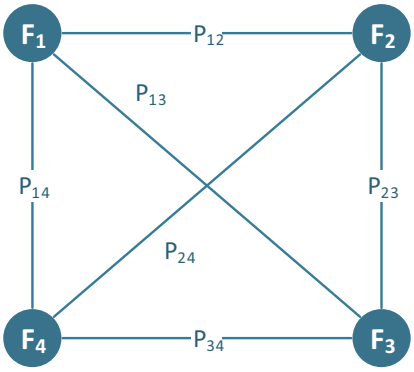

Fig 4. MACO feature space for 4 features where each node representing the probability of selecting both of the features in the feature subset.

represents the probability of selecting the feature i and j simultaneously. So if the P_{ij} is low between two nodes, then, given that F_i is selected for feature subset, the likelihood of selecting F_j for that feature set would be low. So the objective of the ACO is to maximize the P_{ij} between the distinctive features with higher classification performance, and dampen the P_{ij} with features which provide lower classification performance.

B. Proposed Probability Function

To reflect the stochastic nature of the proposed algorithm, the probability function has two components: static heuristic information (HI) η and dynamic pheromone values τ. The probability function is formulated using (6).

$$Pij = \frac{|\tau_{ij}|^{\alpha} \cdot |\eta_{ij}|^{\beta}}{\sum_{k=1}^{N} |\tau_{ik}|^{\alpha} \cdot |\eta_{ik}|^{\beta}} \tag{6}$$

Here α and β govern the trade-off between pheromone trail intensity and HI. From the perspective of an ant, the HI is the a priori knowledge that drives the initial choice of the path by the first team of ants. In the case of feature selection, HI guides the selection of initial feature subsets, which are then evaluated by the evaluation function. The result of the evaluation function dictates the amount of pheromone deposited in each path. By providing a good starting point, the HI can reduce the searching time significantly. On the other hand, a sub-standard prior knowledge can lead to the trapping at the local best subset. In the case of feature selection using ACO, previous research works have utilized the inverse of Pearson correlation coefficient between the nodes, i.e. the feature pairs, as the HI [16]. The underlying assumption of this choice is that the correlation between the features in optimal feature subset must be minimal. However, in our case, where we extracted the features from color space, we have seen that the selection of correlation value as the HI cause the algorithm to select local optimum feature subset. The reason behind the suboptimal performance of the previous HI may be the failure of correlation value in reflecting the goodness of features due to the high correlation between the features extracted from the RGB and HSV color-space.

In previous works on feature selection using ACO, correlation value was selected as the HI. However, in the case of features collected from the different color space, we have seen that correlation does not correspond closely to the goodness of a feature in classification due to the high correlation between each feature extracted from color space. To this end, we have proposed a new heuristic information based on the classification accuracy achieved by each pair of features. Since, we are using SVM classifier here, we have utilized the SVM optimization accuracy resulted from the cross-validation of the training data as the HI, as illustrated in (7)

$$\eta_{ij} = SVM\_Optimization\_Accuracy([F_i, F_j]) \tag{7}$$

Although it is time-consuming to compute the SVM model for each pair of features, as it is only done once and as the feature selection is an offline process, it was a feasible approach to determine a much more informative HI. With this heuristic information as the base, we can form an efficient searching procedure to find the best global feature subset.

C. Ant Generation

Based on the probability function, k feature subsets are selected as the initial path for the ant. To generate the ant, first all the columns in the probability function P_{ij} is accumulated to determine the total probability of selecting $i-th$ feature (U_i) using the following:

$$U_i = \sum_{j=1}^{N} Pij \tag{8}$$

By comparing each row of U_i with a random number R (a value between 0 to 1 generated each time for ant generation), the algorithm forms local feature subset for each iteration. The algorithm selects the rows that have a higher value than R as the features for the particular subset s_t.

$$i \in s_t, \qquad if\ U_i > R \tag{9}$$

D. Evaluation Function

An evaluation function calculates the cost of each path using (10),

$$\gamma(s_t) = E(s_t) + \lambda \times N_l(s_t) \tag{10}$$

Here, $E(s_t)$ function reflects the cost for classification error while $N_l(s_t)$ determines the cost from the subset length. λ controls the trade-off between these two costs. The length function $N_l(s_t)$ is defined as the normalized feature subset length as shown in (11)

$$N_l(s_t) = \frac{length\ of\ s_t}{Total\ Number\ of\ Feature} \quad (11)$$

The cost of the classification error reflects the performance of the classification algorithm defined by accuracy, specificity, and sensitivity. If TP, TN, FP, and FN are respectively the number of true positive (bleeding frame detected correctly), true negative (non-bleeding frame detected correctly), false positive (non-bleeding frame erroneously detected as bleeding) and false negative (bleeding frame erroneously detected as non-bleeding) events, the accuracy, sensitivity and specificity are defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (12)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (13)$$

$$Specificity = \frac{TN}{TN + FP} \quad (14)$$

Ideally, the priority of the system should be to achieve the highest accuracy. In our experiment with the exhaustive search of the optimum feature subset for RGB-classifier and HSV-classifier, we have found several feature subsets with comparable accuracy. To reflect the priority of sensitivity over specificity in bleeding detection, we have chosen the feature set with higher sensitivity in such case. To prioritize higher sensitivity for searching the optimal feature subset in MACO, we have proposed a new evaluation function as shown in (15)

$$E(s_t) = (1 - Accuracy(s_t)) + \omega.(1 - Sensitivity(s_t)) \quad (15)$$

Here ω is the constant dictating the trade-off between the accuracy and sensitivity. Here accuracy and sensitivity are measured using the five-fold cross validation of the SVM model using feature subset s_t.

E. Updating of the Pheromone and Probability Function

Depending on the evaluation function, pheromones are either deposited to or evaporated from each path. The amount of pheromone deposited, $\Delta\tau_{ij}$, depends on the cost of each path as given in (16)

$$\Delta\tau_{ij} = \begin{cases} \dfrac{Q}{\gamma(s_t)} & if\ link(i,j) \in s_t \\ 0 & otherwise \end{cases} \quad (16)$$

Here Q is a fixed quantity of pheromone. So the amount of pheromone deposited in each path is inversely proportional to the loss of that way. Then the pheromone of each path is updated following the (17)

$$\tau_{ij}(t+1) = (1-\rho).\tau_{ij}(t) + \rho.\Delta\tau_{ij}(t) \quad (17)$$

Here ρ, which takes a value in the range of 0 to 1, controls the convergence of the algorithm and avoid the risk of stranding in local minima.

F. Stopping Criterion

There are two conditions for stopping in the proposed algorithm. A maximum number of iteration T was set. If the number of iteration reached T, the iteration would stop, and the local best subset with the smallest cost would be given as the global best subset. Another condition was set to check for convergence before the number of iteration reached T. In this case, if all the local best subsets converged to one subset, and remained the same for a fixed number of iteration, then we assumed that the algorithm has converged. The converged set is given as the global best subset. Fig. 5 shows the complete pseudo-code.

IV. RESULTS

For the training of the classifier and feature selection, we have utilized the features from a dataset of 300 images, with 150 bleeding images and 150 non-bleeding images. For testing, we have utilized six different videos sequence. Table II illustrates the statistics of each video sequences taken from the DVD resource available in [18].

Algorithm Feature Selection Using MACO

```
Input: Training Database
Initialize MACO parameter α, β, ρ, Q, λ, K, T, ω, τ
Determine Heuristic Information:
for each feature pair (i,j) where i!=j
        ηij= ηji=Accuracy ({Fi, Fj})
end
Repeat
    Calculate P
    for ant=1 To K
        Ui=Sum of each column (Pij)
        Generate a random number (R)
        st (ant)= Find the index of Ui >R
    end
    for each subset
        if Loss(st)< Loss(S)
            Local Best S st
        else
            Local Best S S
        end
    end
    Update pheromone for Local Best S
    Evaporate pheromone of all links
    if (iteration is divisible by 10)
        Check Convergence
    end
Until (iteration> T) Or (Convergence==1)
return last Local Best Subset S as Global Feature Subset
end
```

Fig. 5. The proposed MACO Algorithm for Feature Selection

TABLE II. STATISTICS OF THE TESTING DATABASE

Video	Total No. of Frames	Bleeding Frames	Non-Bleeding Frames
1	99	0	99
2	99	0	99
3	99	0	99
4	99	99	0
5	99	99	0
6	100	96	4
Total	595	294	301

To assess the performance of ACO in determining the optimum feature subset, we have first compared the performance of ACO with exhaustive search method regarding feature length reduction, classification performance (Sensitivity, Specificity, and Accuracy) and computational time. Although the calculation of heuristic information can be done offline, we have included the time to compute the heuristic information in the MACO. Later we have utilized the ACO algorithm for determining the optimum feature subset for RGB-HSV classifier. Table III listed the parameter that we used in this model for MACO. This parameters were selected by optimizing the performance of the MACO for a smaller subset of features (with only the first 6 features) on a smaller dataset containing 25 bleeding and 25 non-bleeding images.

A. Performance on RGB and HSV individual classifier

We compare the performance of our proposed feature selection method, the modified ant colony optimization algorithm, with the ES, SFS, SBS, BGA [14], RSFS [15] and the ant colony optimization (ACO) algorithm implemented in [16]. The results for RGB classifier and HSV classifier are listed in Table IV. The best result is shown in bold and the worst result is shown in italic. By observing the result, we conclude that feature selection using MACO shows much better feature selection capability than ACO, SFS, SBS, BGA and RSFS. For RGB classifier, MACO provided the exact feature set as the exhaustive search (ES) method with the highest accuracy and smallest feature set. Although ACO showed excellent sensitivity for RGB classifier, however, the feature set length and the accuracy were substandard. The superiority of MACO over ACO indicates that the proposed heuristic information in this paper can provide better performance for features taken from highly correlated features such as color features.

TABLE III. PARAMETERS FOR MACO

α	0.5	Q	0.01	λ	0.15	T	1000
β	0.75	ρ	0.5	K	10	ω	0.2

For HSV classifier, the MACO provides a different feature set than ES. These sets differ in only one feature while both of them contain the mean of hue and saturation. The accuracy and sensitivity obtained are comparable to each other. While doing the ES, we have seen different subsets can yield similar results. This is may be the reason behind the different feature subset generated by ES and MACO. However, MACO finds the feature set in 1.85% time of the exhaustive search. BGA also gives a good result. However, it takes a longer time. As evident from the Table IV, the most relevant two features in RGB are the mean of red and mean of

TABLE IV. COMPARISON OF FEATURE SELECTION METHOD FOR RGB CLASSIFIER AND HSV CLASSIFIER

	Feature Selection	Feature Set Selected	Number of Features	Sensitivity	Specificity	Accuracy	Percentage of computational time on Exhaustive Search[b] (%)
RGB	ES	$\{F_1,F_2\}$	**2**	0.9798	0.9800	**0.9799**	*100*
	MACO	$\{F_1,F_2\}$	**2**	0.9798	0.9800	**0.9799**	1.88
	ACO	$\{F_1,F_2,F_3,F_4,F_{10},F_{11},F_{12},F_{13}\}$	8	**0.9866**	*0.6411*	0.8084	**1.86**
	SFS	$\{F_1,F_2,F_6,F_{14}\}$	4	0.6938	**1**	0.8387	1.98
	SBS	$\{F_1,F_2,F_3,F_6,F_7,F_{10},F_{15}\}$	7	*0.0306*	**1**	*0.5110*	3.05
	BGA	$\{F_1, F_2, F_3, F_5\}$	4	*0.8586*	0.8581	*0.8566*	15.41
	RSFS	$\{F_1, F_2, F_3, F_5\}$	4	*0.8586*	0.8581	*0.8566*	21.36
HSV	ES	$\{F_{16},F_{17},F_{18}\}$	**3**	**0.9866**	0.9800	**0.9833**	*100*
	MACO	$\{F_{16},F_{17},F_{19}\}$	**3**	0.9797	0.9700	0.9748	1.85
	ACO	$\{F_{16},F_{17},F_{18},F_{19},F_{20},F_{21},F_{22},F_{25},F_{26},F_{27}\}$	9	*0*	**1**	*0.5058*	**1.52**
	SFS	$\{F_{16},F_{17},F_{19},F_{25}\}$	4	0.9764	0.9634	0.9715	1.98
	SBS	$\{F_{16},F_{17},F_{22},F_{23},F_{30}\}$	5	*0*	**1**	*0.5058*	3.39
	BGA	$\{F_{16},F_{17},F_{18}\}$	**3**	0.9797	0.9700	0.9748	15.41
	RSFS	$\{F_{16},F_{17},F_{19},F_{20},F_{22},F_{26}\}$	6	0.5102	0.9834	0.7496	9.61

[b] Computational time for ES: 12 hours (approximately)

TABLE V. COMPARISON OF FEATURE SELECTION METHOD FOR FEATURES TAKEN FROM RGB-HSV CLASSIFIER

Feature Selection	Feature Set Selected	Number of Features	Sensitivity	Specificity	Accuracy	Computational Time (min)
MACO	$\{F_2,F_{16},F_{17}\}$	**3**	**0.9966**	0.9801	**0.9882**	147
ACO	$\{F_{17},F_{20},F_{22},F_{25},F_{26},F_{28}\}$	6	0.5059	**1**	*0.5058*	47
SFS	$\{F_{16},F_{17},F_{19},F_{25}\}$	4	0.9864	0.9900	**0.9882**	133
SBS	$\{F_1,F_2,F_{11},F_{16},F_{17},F_{18},F_{22},F_{23},F_{29},F_{30}\}$	*10*	*0.0034*	**1**	0.5075	*362*
BGA	$\{F_1,F_2,F_3,F_5\}$	*4*	*0.9864*	0.7475	0.8655	111
RSFS	$\{F_2,F_3,F_4,F_{16},F_{17},F_{18},F_{19},F_{22},F_{28}\}$	9	*0.0404*	1	0.6162	**44**

green. While in HSV, mean of the hue, saturation and value are the prominent features. These result closely matches with the features used in [9].

B. Performance on Merged Classifier

The result of feature selection over a feature space of 30 features taken from RGB and HSV color space is shown in Table V. To give a perspective of the result, theoretically, if the exhaustive search were conducted using the current resource, it would take approximately 1.5 years to complete the feature selection procedure. Between the feature selection methods, SBS takes the maximum time, which may be caused by the reason that the best feature set has a much smaller number of features than the entire feature set. So by dropping one in each step, the SBS required a larger number of steps to obtain the optimal feature set. By observing the result, we can conclude that the MACO provides the best performance regarding sensitivity, accuracy and computational time. The feature set selected by the MACO constitutes mean of hue, mean of saturation and mean of the green channel. These features match closely with features used in [9] and [11], where the value of hue and saturation were used for classifying the bleeding region. The distinctive nature of green channel values can be seen from Fig. 2. The lack of higher order statistical features such as standard deviation suggests that the texture of bleeding and the non-bleeding region does not contain discriminative characteristics. As can be seen from Table IV, SFS also provides a comparable accuracy, however with a larger feature subset and higher computational time for feature selection. If we compare the performance of individual classifiers, we can see that by utilizing features from two color space RGB-HSV classifier can provide better performance with regard to accuracy and sensitivity.

V. CONCLUSION

We have introduced a modified ant colony optimization (MACO) algorithm to devise a bleeding detection algorithm based on color features extracted from RGB and HSV color space. By utilizing a new heuristic information function and evaluation function, the proposed feature selection algorithm MACO has illustrated the initial promise of optimal feature selection capability. By comparing the performance of the Exhaustive search (ES) method, we showed that while reducing the computational time by 1.5% MACO can select the optimal feature set with comparable accuracy and sensitivity to Exhaustive Search method. We have also utilized MACO to provide an optimal feature set for RGB-HSV classifier, which has shown better performance than the RGB classifier and HSV classifier designed by ES. In the future, our work will extend to incorporate features from other color spaces to devise robust anomaly detection algorithm for WCE over a broad range of applications.

REFERENCES

[1] G. Iddan, G. Meron, A. Glukhovsky, and P. Swain, "Wireless capsule endoscopy," *Nature*, vol. 405, no. 6785, pp. 417–418, 2000.

[2] M. Pennazio, R. Santucci, E. Rondonotti, C. Abbiati, G. Beccari, F. P. Rossini, and R. De Franchis, "Outcome of patients with obscure gastrointestinal bleeding after capsule endoscopy: Report of 100 consecutive cases," *Gastroenterology*, vol. 126, no. 3, pp. 643–653, 2004.

[3] W. Tan, Y. J. Gao, X. B. Li, J. Dai, S. W. Fu, Y. Zhang, H. B. Xue, and Y. J. Zhao, "Long-term outcome in patients with obscure gastrointestinal bleeding after negative capsule endoscopy.," *J. Dig. Dis.*, vol. 16, pp. 125–134, 2015.

[4] M. Liedlgruber and A. Uhl, "Computer-aided decision support systems for endoscopy in the gastrointestinal tract: A review," *IEEE Rev. Biomed. Eng.*, vol. 4, pp. 73–88, 2011.

[5] Y. Fu, W. Zhang, M. Mandal, and M. Q. H. Meng, "Computer-aided bleeding detection in WCE video," *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 2, pp. 636–642, 2014.

[6] J. M. Buscaglia, S. A. Giday, S. V. Kantsevoy, J. O. Clarke, P. Magno, E. Yong, and G. E. Mullin, "Performance characteristics of the suspected blood indicator feature in capsule endoscopy according to indication for study.," *Clin. Gastroenterol. Hepatol.*, vol. 6, no. 3, pp. 298–301, 2008.

[7] D. K. Iakovidis and A. Koulaouzidis, "Software for enhanced video capsule endoscopy: challenges for essential progress," *Nat. Rev. Gastroenterol. Hepatol.*, vol. 12, no. 3, pp. 172–186, 2015.

[8] T. Ghosh, S. A. Fattah, and K. A. Wahid, "Automatic bleeding detection in wireless capsule endoscopy based on RGB pixel intensity ratio," in *International Conference on Electrical Engineering and Information & Communication Technology*, 2014, pp. 2–5.

[9] S. Sainju, F. M. Bui, and K. A. Wahid, "Automated bleeding detection in capsule endoscopy videos using statistical features and region growing.," *J. Med. Syst.*, vol. 38, no. 4, p. 25, 2014.

[10] T. Ghosh, S. K. Bashar, S. A. Fattah, C. Shahnaz, and K. A. Wahid, "A feature extraction scheme from region of interest of wireless capsule endoscopy images for automatic bleeding detection," *2014 IEEE Int. Symp. Signal Process. Inf. Technol.*, pp. 000256–000260, 2014.

[11] S. Zhou, X. Song, M. A. Siddique, J. Xu, and P. Zhou, "Bleeding detection in wireless capsule endoscopy images based on binary feature vector," in *Fifth International Conference Intelligent Control and Information Processing (ICICIP)*, 2014, pp. 29–33.

[12] Y. Saeys, I. Inza, and P. Larranaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[13] M. Dorigo and T. Stützle, "The ant colony optimization metaheuristic : algorithms , applications , and advances," in *Handbook of metaheuristics*, Springer US, 2003, pp. 250–285.

[14] B. Oluleye, A. Leisa, and D. Dean, "A Genetic Algorithm-Based Feature Selection," *Br. J. Math. Comput. Sci.*, vol. 5, no. 4, pp. 899–905, 2014.

[15] J. Pohjalainen, O. Rasanen, and S. Kadioglu, "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits," *Comput. Speech Lang.*, vol. 29, no. 1, pp. 145–171, 2015.

[16] M. A. J. Ghasab, S. Khamis, F. Mohammad, and H. J. Fariman, "Feature decision-making ant colony optimization system for an automated recognition of plant species," *Expert Syst. Appl.*, vol. 42, no. 5, pp. 2361–2370, 2015.

[17] A. Rashno, S. Sadri, and H. SadeghianNejad, "An efficient content-based image retrieval with ant colony optimization feature selection schema based on wavelet and color features," in *2015 International Symposium on Artificial Intelligence and Signal Processing*, 2015, pp. 59–62.

[18] D. O. Faigel and D. R. Cave, *Capsule endoscopy*. Saunders Elsevier, 2008.