UTM UNIVERSITI TEKNOLOGI MALAYSIA
INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

# A Hybrid Multiwavelet Transform with Grey Wolf Optimization Used for an Efficient Classification of Documents

Ahmed Hussein Salman
Informatics Institute for Postgraduate Studies
Iraqi Commission for Computers and Informatics
Baghdad, Iraq

Waleed Ameen Mahmoud Al-Jawher
Department of Electronics and Communication Engineering
Uruk University
Baghdad, Iraq
Email: phd202110677@iips.icci.edu.iq

Submitted: 30/11/2022. Revised edition: 31/3/2023. Accepted: 31/3/2023. Published online: 13/9/2023
DOI: https://doi.org/10.11113/ijic.v13n1-2.418

***Abstract*—In machine learning, feature selection is crucial to increase performance and shorten the model's learning time. It seeks to discover the pertinent predictors from high-dimensional feature space. However, a tremendous increase in the feature dimension space poses a severe obstacle to feature selection techniques. Study process to address this difficulty, the authors suggest a hybrid feature selection method consisting of the Multiwavelet transform and Gray Wolf optimization. The proposed approach minimizes the overall downsides while cherry picking the benefits of both directions. This notable wavelet transform development employs both wavelet and vector scaling functions. Additionally, multiwavelets have orthogonality, symmetry, compact support, and significant vanishing moments. One of the most advanced areas of study of artificial intelligence is optimization algorithms. Grey Wolf Optimization (GWO) here produced artificial techniques that yielded good performance results and were more responsive to current needs. Keywords — About four key words or phrases in order of importance, separated by commas, used to compile the subject index for the last issue for the year.**

**Keywords—Multiwavelet Transform, Greg Wolf, classifications, vector scaling, optimisation algorithms**

## I. INTRODUCTION

In machine learning, feature selection is crucial to increase performance and shorten the model's learning time. It seeks to discover the pertinent predictors from high-dimensional feature space. However, a tremendous increase in the feature dimension space poses a severe obstacle to feature selection techniques [1,2]. Study process to address this difficulty, we suggest a hybrid feature selection method consisting of the Multiwavelet transform and Gray Wolf optimization. The proposed approach minimizes the overall downsides while cherry picking the benefits of both directions. This notable wavelet transforms development employs both wavelet and vector scaling functions [3]. Additionally, multiwavelets have orthogonality, symmetry, compact support, and significant vanishing moments. One of the most advanced areas of study of artificial intelligence is optimization algorithms. Grey Wolf Optimization (GWO) here produced artificial techniques that yielded good performance results and were more responsive to current needs [4]. The process of representing documents is usually as vectors using various forms, including TFIDF unigram representation, bigram, word bag, etc.

The document representation in text corpora is frequently relatively high dimensional and corresponds to the vocabulary size. The "curse of dimensionality" affects machine-learning models, decreasing performance [5]. In particular, SMS, tweets, and other short text corpora suffer from a sparse high-dimensional feature space because of their vast vocabulary and short document length [6]. We contrast the Reuters and Twitter data corpora to grasp better how these elements impact the size of the feature space. After standard preparation procedures, the Reuters-21578 corpus comprises 14506 distinct vocabulary entries and a total word count of around 2.5 million (the feature space's dimensionality). However, the extent of the feature space in the Twitter 1 corpus we used in our study is 7423 words, and there are roughly 15,000 words in all. Additionally, a document in the Reuters-21578 news categorization corpus has an average length of 200 words compared to 10 to 12 for an English tweet [7]. Therefore, the dimensionality is very high, even for small corpora with brief texts [8]. Additionally, on average, detailed text data has much fewer words per document, resulting in a sparser feature space representation of documents. Due to this high dimensionality issue, feature selection is crucial in text classification procedures. [9-12].

The Improvised Adaptive Discriminant MultiWavelet Packet Transform is the technique we offer in this research for lowering the dimensionality of a brief text. (IADMWPT) and define the most distinct parameters to represent all documents in the set using the MultiWavelet transform as input to the Gray Wolf optimization algorithm. For feature selection and classification, researchers now turn to optimization approaches first. Data mining is now a significant area of study. Gaining knowledge is the main objective of the data mining process because it allows for the creation of any conclusion or output. The complexity of the data always affects computing costs in the field of data mining. Each data collection includes several sample sets. These examples offer details, or features, regarding a particular scenario. In addition to multi-dimension, duplicated and useless features are significant restrictions. Large feature sets have become a barrier for conventional machine learning techniques. Thus, the many goals of dimension reduction and redundancy elimination have been achieved by utilizing feature selection approaches. This study suggests a multiwavelet-based grey wolf optimization strategy to categorize the papers. The proposed method uses DMWT to extract features linked to documents. The time and frequency domains are simultaneously accessible for the information derived from transient signals. The optimal output parameters for identifying the documents are the ones provided by grey wolf optimization using the DMWT output signals. Hand-labeling groupings of records for certain variables is a common task in social paper (e.g., manual coding). It was necessary to hire a team of research assistants and teach them how to read and analyze content manually. It was infeasible, expensive, error-prone, and demanded intensive data quality checks if you had a vast corpus of text that needed to be classified. As a replacement, we may now type text using machine-learning models into particular groups of classes.

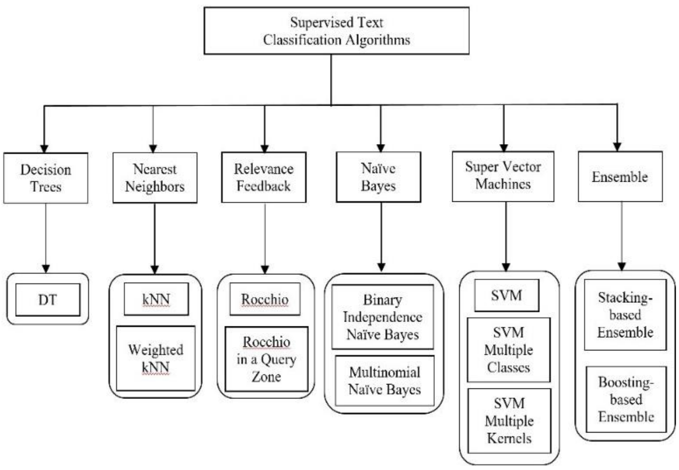


Fig. 1. Block diagram for supervised text Classification Algorithms

## II. Related Work

Previously, the dimensionality of the text dataset was frequently reduced through feature selection. Yiming Yang and colleagues compared some of these methods, including document frequency, information gain (IG), mutual 321 information (MI), 2-test (CHI), and term strength (Yang and Pedersen, 1997). (TS). They found that the most effective methods for aggressive dimensionality reduction are IG and CHI. [10] approach determines the optimum feature subset by boosting feature relevance to the target class and lowering redundancy between selected features. The wavelet transform converts a signal into a representation of temporal frequency. Can use the time-frequency format to describe movements with time-varying frequency content. Beyond the purview of this study, a thorough explanation of wavelet transform theory is not possible. Hammad *et al.* proposed compressing data using wavelet transformation and recommended Adaptive Discriminative Wavelet Transformation (ADWPT). Wavelet transform has previously been used for applications involving natural language processing, which also address why wavelets are such a powerful tool for data mining, cover the fundamentals and characteristics of wavelets in a survey of wavelet applications.

Aggarwal2002 classifies strings using wavelets. Geraldo Xexeo and colleagues (2008) proposed the multi-wave. They stated that the wavelet technique uses the new representation to create a hierarchical data segmentation, which can recognize trends at different resolutions. For feature selection issues, [13] presented the sigmoid function. The addition of simulated annealing improves exploitability. After each iteration, our hybrid technique improved the best answer. The whale optimization algorithm in 2018 [14] uses crossover and mutation operators. By using the whale optimization algorithm in combination with chaotic search, the issue of local optima stagnation and sluggish convergence is addressed [15]. In 2019, a BGOA (binary grasshopper approach) with two improvisation mechanisms is planned. The first mechanism uses a V-shaped function for the transformation function, whereas the second method uses a mutation operator to improve classification accuracy [16].

## III. Multiwavelet Transform

One type of transformation that can be applied to compress images is the wavelet transform, representing a more modern alternative [6]. Wavelets and multi-waves are pretty similar but differ significantly [7]. In the context of a polycyclic study, wavelets can be characterized using the scaling function $\Phi$ (t) and the wave function $\Psi$ (t) in the actual number of scales (and the wavelet function) possible [8]. Multiwave has this as its central tenet. It is a valuable tool for signal processing tasks such as image compression and waveforms [9]. Scalar wavelets, which are waves produced by a single scale function, were the only ones known until recently. However, it is possible to envision a scenario in which multiple measurement functions are [10]. As a result, it has been codified, which has several benefits over scalar waves [13]. Important known properties in signal processing include short support, orthogonality, symmetry, and vanishing moments. Not all of these qualities can coexist in a scalar wave. [14]. In contrast, a multi-wave system can offer a high degree of approximation, flawless reconstruction while preserving length (orthogonality), and excellent performance at boundaries (through linear phase symmetry) (vanishing moments). As a result, multiwave performs better than a scalar wave in applications involving image processing [14].

The inputs and outputs of each channel in the filter bank are evaluated with a vector value, which is one of the main differences between multiwave and scalar waves. It is necessary to convert the input with a scalar value into a signal with an appropriate vector value. Preprocessing is a term for this conversion [15]. Here is a summary of several possible justifications for choosing a multi-wave [14]:

Reduced limitations on the filter qualities can be achieved by using the additional degrees of freedom provided by multiwavelets. For instance, it is widely known that a scalar wavelet cannot have an impulse response that is symmetric, longer than 2 and orthogonal.

In contradiction to symmetric signal extension, which requires for symmetric filters, orthogonality makes the transformation easier to design and implement.

- For scalar wavelets, the filter length and the number of vanishing moments directly correlate with the filter length. This shows that longer filter lengths must be utilized at the expense of the wavelet's support interval to achieve higher-order approximation (in the time domain). A high order of approximation is sought for higher coding benefit, even if shorter wavelet support usually is preferred to generate a more localized approximation of the input function.
- Unlike the limitations of scalar wavelets, multiwavelets can simultaneously have the best of all these characteristics. For instance, the GHM multiwavelet [16,17] is orthogonal and has second-order approximation, symmetric scaling, and wavelet functions. It also has short support for its scaling tasks (and thus symmetric filters). These beneficial characteristics cannot be combined with scalar wavelets. For instance, the four-tap Daubechies filter approximates to the second order, enables scaling functions at level 3, and is orthogonal. However, since this scalar filter is orthogonal, it lacks the critical aspect of symmetry.

The coefficients of the GHM multiskilling and multiwavelet functions are 2x2 matrices in the multiwavelet configuration. Additionally, they must multiply vectors during the transformation stage (instead of scalars). It's orthogonal. Thus, the multifilter bank requires two rows of input data. Therefore, repeating a signal is the most apparent approach to extracting two input rows from its [18]. This process, known as "Repeated Row," introduces a twice oversampling of the input [19,20].

For a given scalar input signal Xk of length N, the input stream is replayed with the same stream multiplied by a constant (N is assumed to be a power of 2, making it even in length). In this way, the original as is used to produce the input length-2 vector.

Whereas Gk for the GHM system is comprised of four wavelet matrices (G0, G1, G2, and G3) [17], Hk for the GHM system is made up of four scaling matrices (H0, H1, H2, and H3). Scalar wavelet transform matrices for discrete MultiWavelets computation can be seen in Fig. 2:

$$W = \begin{bmatrix} H_0 & H_1 & H_2 & H_3 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & H_0 & H_1 & H_2 & H_3 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ H_2 & H_3 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & H_0 & H_1 \\ G_0 & G_1 & G_2 & G_3 & \vdots & \vdots & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & G_0 & G_1 & G_2 & G_3 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & G_0 & G_1 & G_2 & G_3 \\ G_2 & G_3 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & G_0 & G_1 \end{bmatrix}$$

Fig. 2. MultiWavelets transform matrices

Where the low-pass and high-pass filter impulse responses, respectively, are Hi and Gi.

### A. DMWT for 1-D Signal Computation

Employing a redundant sampling technique from preprocessing, the process of performing a Discrete Multi-Wave Transformation (DMWT) matrix multiplication in dimension relative to an input, where N should be a power of two square matrices (the iterative row) [19]. After preprocessing, the transformation matrix's dimensions match the input signal's dimensions. The procedures listed below must be taken in order to calculate a discrete one-dimensional one-level multi-wave transformation: Should check Input dimensions, and the input vector should have a length of N, where N must be a power of two.

- By combining the matrices for the GHM, low and high pass filters produce a transformation matrix W. By entering the values for the GHM matrix filter coefficients, a 2Nx2N transformation matrix is created.
- For the GHM system to prepare the input signal, 1. The input stream must be repeated while increasing it by a fixed amount $\alpha = \frac{1}{\sqrt{2}}$.
- The 2Nx2N transformation matrix, created by the 2Nx1 preprocessing input vector, can be multiplied to modify the input vector.

## IV. Grey Wolf Optimizati

With the optimal wolves, mathematical equations may be used to formally reflect the hunting social team member's attitude to arrive at the best option [22]. The w wolves participate in the hunt procedure by following the other famous wolf members [20-26]. The following is a list of the main steps involved in hunting:

- Get close to the prey.
- Interrogate the prey while surrounding it until the stopping condition is met.
- Strike at the prey.

### *A. Encirclement of Prey*

Grey wolf is circling and chasing the prey while hunting. It is modeled in mathematics as shown in Equations (1) and (2):

D=|C .X_p (t)-X(t)| (1)
X(t+1)=X_p (t)-A .D (2)

In this contract, iteration t is the designated present iteration, while iteration t + 1 is the designated future iteration. Gray wolves' feature vector is X(t), and pray's feature vector is X_P (t). And are elements that they are identified as:

D=2 .A.r_1-a (3)
C=2. r_2 (4)

During an iteration, r1 and r2 are the randomized vectors with periods of 0 and 1 that decrease linearly from two to zero.

### *B. The Hunting Method*

When prey positions are known, a grey wolf will circle the area in preparation for hunting. Members of the grey wolf pack direct the hunting mechanism. The locations of the prayers are not indicated anywhere in the search area; instead, it is assumed that the wolf members α, β, δ, and have sufficient knowledge of the prayer places. As a result, the top candidates are retained while the rest are rejected, meaning that W gray wolf pack members are still updating their locations with the best options.

X(t+1)= 〖(X〗 _(1+) X_(2+) X_3)/3 (5)

Exploration and Exploitation while hunting and attacking the target. The process' random element helps prevent becoming stuck in the local minimum.

## V. GWO Mathematical Example

Find the maximum value of f(x)=1+2x-x^2 for -5<=x1, x2<=5 number of population =5, number of iteration=2.

TABLE I. Input of GWO

| wolves | x | F(x) | Select max f(x) | position |
|---|---|---|---|---|
| 1 | 0.2 | 1.36 | 1.91 | $x_\alpha$=1.3 |
| 2 | 0.3 | 1.51 | 1.51 | $x_\beta$=0.3 |
| 3 | 0.3 | 1.91 | 1.36 | $x_\&$=0.2 |
| 4 | 2.5 | -0.25 | | |
| 5 | 5 | -14 | | |

TABLE II. GWO 1 Iteration

| initial | Best wolf position αβ& | R1 Rand(0-1) | R2 Rand(0-1) | A A=2ar1-a | C C=2r2 | D $D = \|C . X_p(t) - X(t)\|$ | X X=x-AD | X= (x1+x2+x3)/3 | F(x) |
|---|---|---|---|---|---|---|---|---|---|
| a=1.33 x=0.2 | $x_\alpha$=1.3 | 0.532 | 0.95 | A1=0.08538 | C1=1.9 | D1=2.27 | X1=1.06 | $x_{new}$=0.54 | 1.79 |
| | $x_\beta$=0.3 | 0.4373 | 0.2477 | A2=0.1665 | C2=0.049 | D2=0.054 | X2=0.309 | | |
| | $x_\&$=0.2 | 0.5477 | 0.669 | A3=0.1264 | C3=1.338 | D3=0.068 | X3=0.193 | | |
| A=1.33 X=0.3 | $x_\alpha$=1.3 | 0.861 | 0.6091 | A1=0.697 | C1=1.218 | D1=1.284 | X1=0.059 | $x_{new}$=0.124 | 1.46<1.51 Greedy=1.51 |
| | $x_\beta$=0.3 | 0.7490 | 0.3807 | A2=0.662 | C2=0.761 | D2=0.072 | X2=0.252 | | |
| | $x_\&$=0.2 | 0.8172 | 0.3397 | A3=0.844 | C3=0.679 | D3=0.164 | X3=0.124 | $x_{new}$=1.3 | |
| a=1.33 x=1.3 | $x_\alpha$=1.3 | 0.9726 | 0.755 | A1=1.260 | C1=0.755 | D1=0.32 | X1=0.897 | $x_{new}$=0.27 | 1.23<1.91 Greedy=1.91 |
| | $x_\beta$=0.3 | 0.3382 | 0.6053 | A2=-0.430 | C2=1.21 | D2=0.94 | X2=0.704 | | |
| | $x_\&$=0.2 | 0.8984 | 0.9280 | A3=1.060 | C3=1.856 | D3=0.93 | X3=0.790 | $x_{new}$=1.3 | |
| a=1.33 x=2.5 | $x_\alpha$=1.3 | 0.2568 | 0.4254 | A1=-0.647 | C1=0.5807 | D1=1.39 | X1=2.199 | $x_{new}$=0.34 | 1.56 |
| | $x_\beta$=0.3 | 0.7799 | 0.2855 | A2=-0.740 | C2=0.571 | D2=2.33 | X2=-1.42 | | |
| | $x_\&$=0.2 | 0.4925 | 0.7014 | A3=-0.0199 | C3=1.4 | D3=2.22 | X3=0.244 | | |
| a=1.33 x=5.0 | $x_\alpha$=1.3 | 0.9058 | 0.8147 | A1=1.709 | C1=1.6294 | D1=3.89 | X1=-2.89 | $x_{new}$=-2.59 | -10.89 |
| | $x_\beta$=0.3 | 0.9134 | 0.1270 | A2=1.0996 | C2=0.254 | D2=4.92 | X2=-5.11 | | |
| | $x_\&$=0.2 | 0.5341 | 0.7837 | A3=-0.091 | C3=1.467 | D3=4.960 | X3=-0.23 | | |

## VI. Proposed Method

The most effective and widely used method for expressing features in image processing is MultiWavelet Transform. Give an approach based on the Adaptive Discriminant MultiWavelet Packet Transform for categorizing meningioma subtypes (ADMWPT). By maximizing the discriminatory power of different properties, ADMWPT develops a representation based on MultiWavelet.

The suggested short text feature selection method ADMWPT is presented in this section. To extract relevant discriminative characteristics from the sub-bands at varying depths, ADMWPT employs a Multiwavelet packet transformation of the data. This document's numerical vector representation is the same as its signal representation. Must make the following computations to obtain an ADMWPT picture of the record:

Create a comprehensive multiwavelet packet transform from the vector representation of the content.
Calculate each coefficient's discrimination power in the multiwavelet packet transform representation.
Pick the discriminative coefficients that best represent all of the corpus's documents.

Once the 1-D multiwavelet transform has been computed at the selected level, produce2^I the multiwavelet packet transform (MWPT) produces two unique sets of coefficients (nodes in the MWPT tree). These coefficients display the relative weights of the various frequencies currently present in the signal at any given time. We select the most discriminative coefficients to represent all the documents in the corpus by evaluating the discriminative strength of each coefficient. These coefficients show the relative weights of the various frequencies that are, at any given time, present in the signal. We select the most discriminative coefficients to represent all the documents in the corpus by evaluating the discriminative strength of each coefficient. The classification task consists of C classes and texts. The 1-D MultiWavelet Packet Transform in the document generates l levels with f sub-bands and m coefficients on each. The symbols 〖X〗_(m.f.l) And l denote the coefficients of the MultiWavelet Packet Transform. As for Algorithm 2.

The process of calculating 1-D multi-wave packet transformation (MWPT) at the given level, produce2^IMulti-wave packet transformation (MWPT) produces two unique parameters (the nodes in the MWPT tree). These coefficients show the relative weights of the different frequencies currently present in the signal at any given time. We choose the most discriminative coefficients to represent all documents in the set by evaluating the discriminative power of each parameter. These coefficients show the relative weights of the different frequencies present in the signal at any time. We choose the most discriminative coefficients to represent all documents in the set by evaluating the discriminative power of each parameter. The classification task consists of C classes and scripts. 1-D multiband beam conversion in the document generates l levels with f subbands and m coefficients. The symbols 〖X〗_(m.f.l) And l denote the multi-wavelength beam conversion coefficients. For algorithm 2.

**ADMWPT Algorithm**

- **for all** classes C **do**
- Calculate MultiWavelet Packet Transform for all the documents $D_K$ in class ci
- **for all** Documents $D_K$ **do**
- Calculate probability density estimates $S_{m,f,l}$
- **end for**
- **for all** Levels of WPT l and their sub bands f **do**
- **for all** MultiWavelet Packet Transform Coefficients m in sub and f **do**
- Calculate average probability density $A_{m,f,l}$
- **end for**
- **end for**
- **end for**
- **for all** Class Pairs $c_a$, $c_b$ **do**
- Calculate discriminative power $D_{m,f,l}$
- **end for**

Select top $m^l$ coefficients for representing documents in corpus.

Fig. 3. Algorithm 2 .ADMWPT Algorithm for best discriminative feature selection

## VII. Results

We used many short-text datasets to verify the proposed approach's effectiveness compared to cutting-edge techniques. Twitter 1: Nakov *et al*. SemEval .'s 2013 task B dataset include the Twitter 1 dataset for two-class sentiment categorization. We obtained 624 instances from our experiments' positive and negative classes.

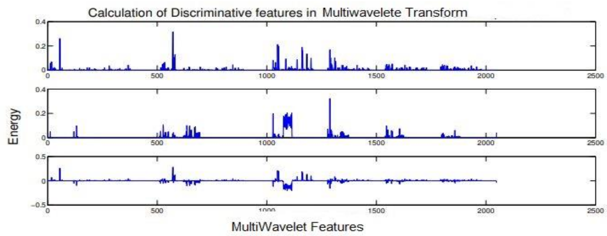


Fig. 4. Shows amplitude on the y-axis and the multiwavelet packet transform coefficients on the x-axis. In all directions

| Dataset | Baseline | MI-avg | $\chi^2$ | PCA | ADWPT | MWT-GWO |
|---|---|---|---|---|---|---|
| Classification accuracy - SVM | | | | | | |
| Twitter 1 | 47.04 | 47.57 | 45.62 | 59.28 | 46 | **63.44** |
| SMS Spam 1 - HAM Accuracy | 99.82 | 99.71 | 99.77 | 99.87 | 99.63 | **99.94** |
| SMS Spam 1 - SPAM Accuracy | 83.10 | 83.32 | 82.96 | 83.81 | 83.57 | **83.92** |
| SMS Spam 2 - HAM Accuracy | 55.2 | 56.31 | 55.62 | 81.12 | 54.1 | **87.7** |
| SMS Spam 2 - SPAM Accuracy | 46.6 | 46.7 | 46.49 | 92.39 | 47.3 | **99.42** |
| Total dimensions in best classification accuracy result - SVM | | | | | | |
| Twitter 1 | 7423 | 2065 | 515 | 540 | 7423 | 23 |
| SMS Spam 1 | 9394 | 3540 | 550 | 750 | 9394 | 815 |
| SMS Spam 2 | 10681 | 2985 | 490 | 855 | 1068 | 250 |
| Classification accuracy - Logistic Regression | | | | | | |
| Twitter 1 | 75.8 | 74.97 | 75.21 | 76.28 | 68.2 | **76.72** |
| SMS Spam 1 - HAM Accuracy | 97.91 | 94.67 | 95.28 | 98.71 | 98.03 | **99.61** |
| SMS Spam 1 - SPAM Accuracy | **95.48** | 85.34 | 86.37 | 91.37 | 82.2 | 87.54 |
| SMS Spam 2 - HAM Accuracy | 96.02 | 89.54 | 92.76 | 71.21 | 95.09 | **98.5** |
| SMS Spam 2 - SPAM Accuracy | 91.2 | 88.37 | 91.38 | 89.15 | 92.2 | **94.51** |
| Total dimensions in best classification accuracy result - Logistic Regression | | | | | | |
| Twitter 1 | 7423 | 5600 | 3250 | 3575 | 7423 | 2749 |
| SMS Spam 1 | 9394 | 7545 | 1755 | 2350 | 9394 | 1680 |
| SMS Spam 2 | 10681 | 6550 | 3000 | 3050 | 10681 | 9981 |

Fig. 5. Classification Results

## VIII. Conclusion

In this article, the authors present a method for efficiently classifying documents. The approach can be applied in several situations where high dimensionality and sparsity present difficulties. For brief text data, experiments demonstrate the effectiveness of the method-based dimensionality reduction. Numerous social media data analysis applications can benefit from this method. the authors plan to investigate theoretical upper limits for the optimal number of dimensions for multiwavelet representation in the future.

## ACKNOWLEDGMENTS

The authors would like to thank the Iraqi Commission for Computer and Informatics, as well as the Informatics Institute for Post Grad, for their help and encouragement in performing this work.

## REFERENCES

[1] Hao, X., Wang, J. T., & Ng, P. A. (1993, October). Nested segmentation: An approach for layout analysis in document classification. *Proceedings of 2nd International Conference on Document Analysis and Recognition (ICDAR'93)* (pp. 319-322). IEEE.

[2] Li, X., & Ng, P. A. (1999, September). A document classification and extraction system with learning ability. *Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR'99 (Cat. No. PR00318)* (pp. 197-200). IEEE.

[3] Aldhuhaibat, M. J., Amana, M. S., Aboud, H., & Salim, A. A. (2022). Radiation attenuation capacity improvement of various oxides via high density polyethylene composite reinforcement. *Ceramics International*, *48*(17), 25011-25019.

[4] Waheed, S. R., Suaib, N. M., Rahim, M. S. M., Adnan, M. M., & Salim, A. A. (2021, April). Deep learning algorithms-based object detection and localization revisited. *Journal of Physics: Conference Series* (Vol. 1892, No. 1, p. 012001). IOP Publishing.

[5] Wang, C. Y., Liu, Q., & Ng, P. A. (1997, December). Intelligent browser for TEXPROS. *Proceedings Intelligent Information Systems. IIS'97* (pp. 388-398). IEEE.

[6] Al-Helali, A. H. M., Mahmmoud, W. A., & Ali, H. A. (2012). A Fast personal palmprint authentication based on 3d-multi wavelet Transformation. *International Journal of Scientific Knowledge*, *2*(8).

[7] Al-Jawher, W. A. (2007). New fast method for computing multiwavelet coefficients from 1D up to 3D. *Proc. 1st Int. Conference on Digital Comm. & Comp. App., Jordan* (pp. 412-422).

[8] Al-Jouhar, W. A., & Abbas, T. M. (2006). Feature combination and mapping using multiwavelet transform. *Iraq Academic Scientific Journals*, *3*(19), 13-34.

[9] Mahmoud, W. A., Abdulwahab, M. S., & Al-Taai, H. N. (2005). The determination of 3D multiwavelet transform.

[10] Mahmoud, W. A. (2011). A smart single matrix realization of fast Walidlet transform. *International Journal of Research and Reviews in Computer Science*, *2*(1), 144.

[11] Salim, A. A., Ghoshal, S. K., Bakhtiar, H., Krishnan, G., & Sapingi, H. H. J. (2020, April). Pulse laser ablated growth of Au-Ag nanocolloids: Basic insight on physiochemical attributes. In *Journal of Physics: Conference Series* (Vol. 1484, No. 1, p. 012011). IOP Publishing.

[12] Salim, A. A., Bidin, N., & Islam, S. (2017). Low power CO2 laser modified iron/nickel alloyed pure aluminum surface: Evaluation of structural and mechanical properties. *Surface and Coatings Te chnology*, *315*, 24-31.

[13] Al-Jadir, I., & Mahmoud, W. A. (2021). A grey wolf optimizer feature selection method and its effect on the performance of document classification problem. *Journal Port Science Research*, *4*(2).

[14] Adnan, M. M., Rahim, M. S. M., Al-Jawaheri, K., Ali, M. H., Waheed, S. R., & Radie, A. H. (2020, September). A survey and analysis on image annotation. *2020 3rd International Conference on Engineering Technology and its Applications (IICETA)* (pp. 203-208). IEEE.

[15] Mahmoud, W. A., Hadi, A. S., & Jawad, T. M. (2012). Development of a 2-D wavelet transform based on Kronecker product. *Al-Nahrain Journal of Science*, *15*(4), 208-213.

[16] Al-Helali, A. H., Ali, H. A., Al-Dulaimi, B., Alzubaydi, D., & Mahmmoud, W. A. (2009). Slantlet transform for multispectral image fusion. *Journal of Computer Science*, *5*(4), 263.

[17] Al-Helali, A. H. M., Mahmmoud, W. A., & Ali, H. A. (2012). A Fast personal palmprint authentication Based on 3d-multi Wavelet Transformation. *International Journal of Scientific Knowledge*, *2*(8).

[18] Waheed, S. R., Rahim, M. S. M., Suaib, N. M., & Salim, A. A. (2023). CNN deep learning-based image to vector depiction. *Multimedia Tools and Applications*, 1-20.

[19] Kattoush, A. H., Al-Jawher, W. A. M., & Al-Thahab, O. Q. (2013). A radon-multiwavelet based OFDM system design and simulation under different channel conditions. *Wireless Personal Communications*, *71*, 857-871.

[20] Islam, S., Bidin, N., Osman, S. S., Krishnan, G., Salim, A. A., Riaz, S., ... & Sanagi, M. M. (2017). Synthesis and characterization of Ni NPs-doped silica–titania nanocomposites: structural, optical and photocatalytic properties. *Applied Physics A*, *123*, 1-9.

[21] Mahmoud, W. A., Abdulwahab, M. S., & Al-Taai, H. N. (2005). The determination of 3D multiwavelet transform.

[22] Salim, A. A., Bakhtiar, H., & Ghoshal, S. K. (2021). Improved fluorescence quantum yield of nanosecond pulse laser ablation wavelength controlled cinnamon nanostructures grown in ethylene glycol medium. *Optik*, *244*, 167575.

[23] Waheed, S. R., Adnan, M. M., Suaib, N. M., & Rahim, M. S. M. (2020, April). Fuzzy logic controller for classroom air conditioner. *Journal of Physics: Conference Series* (Vol. 1484, No. 1, p. 012018). IOP Publishing.

[24] A. A Salim, Ghoshal, S. K., & Bakhtiar, H. (2022). Prominent absorption and luminescence characteristics of novel silver-cinnamon core-shell nanoparticles prepared in ethanol using PLAL method. *Radiation Physics and Chemistry*, *190*, 109794.

[25] Eid, Heba F. 2018. Binary whale optimisation: an effective swarm algorithm for feature selection. *International Journal of Metaheuristics*. *7*(1), 67-79.

[26] Hathot, S. F., Jubier, N. J., Hassani, R. H., & Salim, A. A. (2021). Physical and elastic properties of $TeO_2$-$Gd_2O_3$ glasses: Role of zinc oxide contents variation. *Optik*, *247*, 167941.