Received July 30, 2019, accepted August 15, 2019, date of publication August 21, 2019, date of current version September 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2936621

A Novel Feature Selection Method for Video-Based Human Activity Recognition Systems

MUHAMMAD HAMEED SIDDIQI[1], MADALLAH ALRUWAILI[1], AND AMJAD ALI[2]
[1]College of Computer and Information Sciences, Jouf University, Sakaka, Saudi Arabia
[2]Department of Computer Science, COMSATS University Islamabad at Lahore, Lahore 54000, Pakistan

Corresponding author: Muhammad Hameed Siddiqi (mhsiddiqi@ju.edu.sa)

This work was supported by the Jouf University, Sakaka, Al Jouf, Saudi Arabia.

ABSTRACT This paper introduces a new feature selection approach for human activity recognition systems to accurately recognize the human activities. We proposed normalized mutual information-based feature selection (NMIFS) method that will select good features extracted from numerous existing feature extraction techniques. The proposed method is an extension of the max-relevance and min-redundancy method. The ability of this method is to combine the strengths of different extraction techniques. However, the selection process might be influenced because of the inequality among the feature's classification power and the feature's redundancy. To escape this influenced selection, we normalize both terms by the proposed feature independent upper bound of the mutual information function. Moreover, we exploit the curvelet transform for feature extraction, and linear discriminant analysis for reduction of feature space. Moreover, we use the hidden Markov model for activity recognition based on our proposed method of feature selection. Finally, by using the benchmark datasets such as KTH and Weizmann datasets, we experimentally compare the proposed scheme with state-of-the-art. Simulation results show that the proposed scheme is not only more accurate for some datasets, but outperforms competing method by weighted average accuracy 98%.

INDEX TERMS Human activity recognition, curvelet transform, mutual information, minimal redundancy, maximal relevance, video surveillance.

I. INTRODUCTION

Contextual computing is the way forward to the success of ubiquitous and pervasive computing. The context of a consumer or customer assists a ubiquitous application in an autonomous setting to trigger for example: making an emergency call when an elderly person being monitored falls down, turning the lights on or off when a person enter or leaves the premises. To correctly synthesize the data coming from various monitoring sources/sensors, activity recognition determines the situation under observation by providing tailored services to ubiquitous applications. Such as monitoring activities of an elderly person and recognition of their activities to understand the mental and physical conditions is an example of healthcare applications [1]. Human activity recognition (HAR) has gained a lot of interest because of its importance in the fields of pervasive computing, human computer interaction [38], computer vision and healthcare.

The associate editor coordinating the review of this article and approving it for publication was Yongqiang Zhao.

A classical HAR system formally comprised on various modules, such as feature selection, feature extraction, human body segmentation, and activity recognition. Among them, feature selection is the important one as it is used for dimension reduction and highly impacts the recognition rate of the HAR system. Feeding a large set of dimensions to an activity recognition system does not only mean heavy computation but also compliments the well known problem *the curse of dimensionality*. Therefore, an efficient feature selection technique reduces the complexity of the dimensionality problem and expedite the recognition process. Feature selection has been studied in literature in detail and helps in understanding the data [2]. However, most of them have their limitations and are bounded by certain applications. We present some of them.

An adaptive learning model based on genetic on genetic algorithm was proposed in [3] for facial expression recognition. However, the genetic algorithm has a well-known drawback i.e., if the population has many subjects then the genetic algorithm provides less assurance in determining the global optimum [4]. Moreover, genetic algorithms are not well-suited for real-time applications and their performance degrade as the update rates increase. Similarly, in [5], the authors apply correlation based feature selection method in the facial expression recognition and show that their has improved significantly. However, this approach only provides a heuristic merit for a subset of features and ignores the independent nature of features [6], and it also is computationally expensive [7]. A wrapper based method for feature selection proposed by Fong, *et.al.*, in [8] produces good results but lacks the global optimum results and mainly generate local optimums [9].

A widely used feature selection method used by many feature extraction systems is the ranking method. Although ranking method performs the general goal of feature selection [11] but it lacks to provide the optimality because of the possibility of getting a redundant subset from the selected subset of features [9]. A recent work by Gharsalli, *et.al.*, [12] have proposed the use of random forest for feature selection. Though their work shows efficiency in their specific application; however, random forests are biased in the favor of multi-level variables/attributes i.e., variables with multiple level of granularity [13]. Likewise, in [14] the authors have proposed swarm search based method in order to select the prominent features from human silhouettes. This method has the drawback of increasing delays and performance degradation in real-time scenarios. Therefore, is only suitable for offline validation [15].

Regarding the feature selection in HAR, several approaches have been developed. However, mutual information based feature selection is the most widely adopted method. Some limitations about this method are outlined in [16]–[19]. For instance, $X_1, X_2, \ldots, X_N$ is a dataset of N different features, where $\left(S_{j-1} = \left\{s_1, s_2, \ldots, s_{j-1}\right\}\right)$, represent a set of $j-1$ selected index. However, the problem is to select a feature X_{s_j}, in which the redundancy $\left(RD\left(X_{s_j}\right) = \sum_{s \in s_{j-1}} I_m\left(X_s; X_{s_j}\right)\right)$ and the relevance $\left(RE\left(X_{s_i}\right) = I\left(C; X_{s_i}\right)\right)$ are minimized and maximized, respectively at the same time. However, in such a hybrid problem finding a single or a common solution is difficult and complex. Therefore, we are interest to address such a hybrid problem where we can solve both issues jointly. Thus, in the above given sub-problems, we want to maximize $RL(X_{s_i}) - \beta \times RD(X_{s_i})$ by considering a scalar factor, say β and the feature X_{s_i}. We summarize the existing solution as follows:

- In the experiments of mutual information feature selection (MIFS) and mutual information feature selection-unsupervised (MIFSU) the value of β was manually selected.
- In the max-relevancy min-redundancy scheme the value of β was set to $1/(S_{I_{m1}})$.
- In the normalized mutual information feature selection (NMIFS) scheme the value of β was set to $1/(S_{j-1}) \times 1/(min(G(X_s), G(X_{s_j})))$.

In this paper, we investigate a robust feature selection scheme by exploiting the information measurements that helps to extracts the hidden information or estimates the potential of features. However, due to impracticality of searching algorithms a greedy forwarding scheme is being adopted, in which quality of feature characteristic is used to combined each individual feature with the set of features [16], [18], [19]. We adopted curvelet transformation method in combination with the proposed scheme to extract and select the most significant features. Moreover, we used Latent dirichlet allocation (LDA) to decrease the dimension of the feature space and hidden Markov model (HMM) to label the activities. The main contributions of the paper are as follows:

- We proposed a mutual information-based feature selection method that will select good features extracted from numerous existing feature extraction techniques. The ability of this method is to combine the strengths of different extraction techniques. However, the selection process might be influenced because of the inequality among the feature's classification power and the feature's redundancy. To escape this influenced selection, we normalize both terms by the proposed feature independent upper bound of the mutual information function.
- We carefully calculated the value of β by considering the maximization and minimization problem, which means the proposed method tracks filtering approach to consider the benefit of low computational cost. Hence, it is appropriate for comprehensive domains. Similarly, the proposed method overwhelmed the drawback of the previously proposed techniques. Therefore, it provides much better recognition rate against the existing one
- We use combination of methods, such as curvelet transform, LDA, and HMM to prove that the proposed method works fine.
- A comprehensive set of experiments have been performed under different combinations of methods using publicly available datasets namely KTH action dataset [20] and Weizmann action dataset [21].

The rest of the paper is organized as follows: Section II presents the details of the proposed MNF-HAR scheme. Section III discusses our experimental testbed and setup, and Section IV, provides the analysis of the results. Finally, Section V, concludes the discussion and put forward future directions.

II. METHODOLOGY

In this section, we present our approach to apply normalized feature selection on extracted features and then present the application of HMM.

A. FEATURE EXTRACTION

In our proposed scheme, we exploit the curvelet method of wrapping because of computationally less expensive and it is an application of curvelet transform [22] which has a great potential to extract the distinct features from each

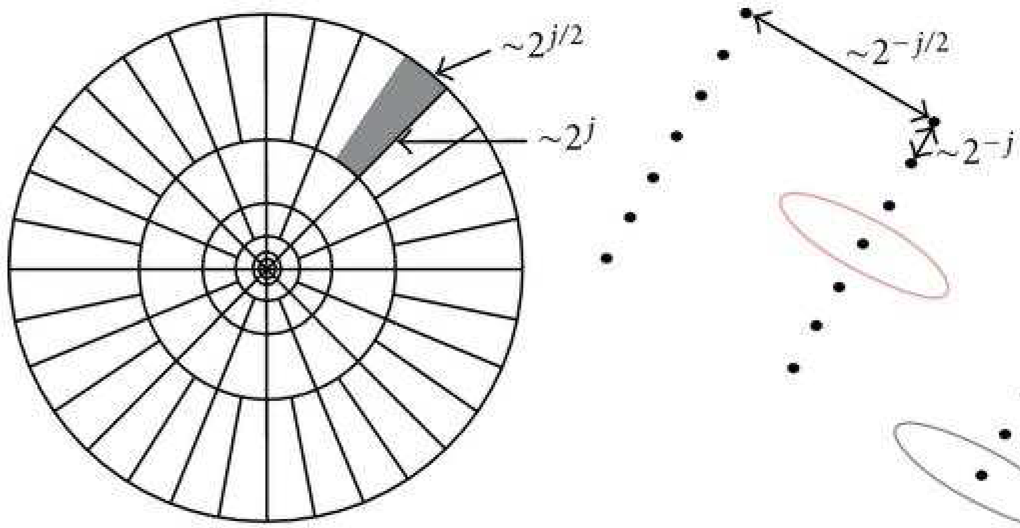

FIGURE 1. Demonstration of scale and angle in frequency and spatial domains [10].

activity frame being processed by using curve, line, and edge information. Moreover, this technique is widely adopted in the cases where the image reconstruction bear the severe ill-posed problem. Furthermore, wrapping scheme employed a minimal number of coefficients that include the matrices of scale and angle which are given as follows:

$$C\,(i, l, k) = \langle f, \varphi_{i,l,k} \rangle \tag{1}$$

where l and i represents the angle and scale, respectively. k indicates the inside products project f onto the $\varphi_{j,l,k}$, and the parameter position. Figure 1 demonstrates the scale and angle. To apply the wrapping method, we employed the following steps.

- First, to get the Fourier samples, we apply Fast Fourier transform as: $\hat{f}\,[n_1, n_2]\,, -\frac{n}{2} \le n_1, n_2 \le \frac{n}{2}$ (see Figure 1).
- Second, we re-smaple (interpolate) each pair of angle l, and scale j to obtain $\hat{f}\,[n_1, n_2 - n_1 \tan\theta_l]$ as: $\hat{f}\,[n_1, n_2]$
- Third, the results of the discrete localizing window function are multiplied with a interpolation function $\hat{f}$ as $\tilde{U}_j\,[n_1, n_2]$.

$$\hat{f}_{j,1}\,[n_1, n_2] = \hat{f}\,[n_1, n_2 - n_1 \tan\theta_l]\,\tilde{U}_j\,[n_1, n_2]$$

- Last, on each $\hat{f}_{j,1}$, we apply the inverse Fast Fourier transform to obtain the associated curvelets $C\,(j, l, k)$.

Here, the range of n_1 and n_2 is between $0 \le n_1 \le L_{1,j}$ and $0 \le n_2 \le L_{2,j}$, whereas the range of θ is between $-\frac{\pi}{4}$ and $+\frac{\pi}{4}$.

B. NORMALIZED MUTUAL INFORMATION-BASED FEATURE SELECTION (NMIFS)

The NMIFS technique is taken from the *max-relevance and min-redundancy* scheme. The NMIFS technique has the limitation to develop normalized mutual information to remove the dominance of relevance and redundancy. We extend this technique by developing an approach to incorporate the functionality to develop normalized mutual information used in this method. We do so by considering the upper bound of the mutual information of random variables, so for that we define a bound. We presume that a pair X, Y of discrete random variables is given as input with their marginal and joint distributions. Then, the joint mutual information are computed by:

$$I_m\,(X; Y) \le \min\,(G\,(X)\,, G\,(Y)) \tag{2}$$

where I_m denotes the combined mutual information of X and Y. Whereas, G represents the defined entropy function as follows:

$$G\,(X) \le \log_2 \left(\sum_{x \in \Omega_X} p\,(x)\,\frac{1}{p\,(x)} \right) \tag{3}$$

$$G\,(X) \le \log_2\,(|\Omega_X|) \tag{4}$$

We get the following from (2) and (4)

$$I_m\,(X; Y) \le \min\left(\log_2\,(|\Omega_X|)\,, \log_2\,(|\Omega_Y|)\right) \tag{5}$$

TABLE 1. Key notations.

N^s	Total data samples
K	Selected number of features
a	Selected measurement index
X^{ij}	Feature values
N	Quantization levels numbers
M	Number of total feature
ξ	Quantization error
$X(1 \ldots M)$	Training data
C^j	Class labels for data samples
S^k	Index of selected feature
$Y(1 \ldots M)$	Quantized data
CB	Code book

The proposed scheme achieves the expected quantization error, ξ by quantizing each feature over the same number of levels (N). Algorithm 1 represents the quantization process.

Algorithm 1: Quantization Algorithm

Input : $M, X(1 \ldots M)$, *and* ξ
Output: N, and $Y(1 \ldots M)$

begin
  Set $N = 2$
  while *1* **do**
    $Max^{Error} = -1e + 16$
    for $i = 1$ *to* M **do**
      $low = min(X(i))$
      $up = max(X(i))$
      $S^{tep} = (up - low)/N$
      $Par = [low : S^{tep} : up]$ $CB =$ $[low - S^{tep}, low : S^{tep} : up][Y(i), Q^{Error}] =$ $Q^{uantiz}(X(i)), P^{artition}, CB$
      if $Q^{Error} > Max^{Error}$ **then**
        $Max^{Error} = Q^{Error}$
      end
    end
    if $Max^{Error} < \xi$ **then**
      Break
    end
    $N = N + 1$
  end
end

It can be observed that the quantization levels number increases gradually till the quantization error reaches to smaller than the predefined threshold, (ξ) which is set to 0.05 in our evaluation. A value less than $\xi = 0.05$ tends to increase the computational cost and did not effect/improve the accuracy. Therefore, threshold value is set to $\xi = 0.05$, the minimum value where the algorithm keeps improving without incurring additional computational costs. For every feature, X such that $|\Omega_X| = N$. So,

$$I_m(X; Y) \leq \log_2(N) \tag{6}$$

where $\log_2(N)$ is a feature-independent upper bound and it does not based on X and Y. Moreover, $\log_2(N)$ is used in order to reduce the problem of unsatisfactory regularizing weights [17]. Therefore, we calculate the normalized feature-to-feature mutual information, $NI_m(X; Y)$ as given below:

$$NI_m(X; Y) = \frac{I_m(X; Y)}{\log_2(N)} \tag{7}$$

In (7), the range of $NI(X; Y)$ lies within [0, 1]. So, to obtain a balance among the redundancy and the relevance, the class-feature mutual information (CFMI) is divided by $\log_2|\Omega_C|$. Balance is described as:

$$NI_m(C; X) = \frac{I_m(C; X)}{\log_2|\Omega_C|} \tag{8}$$

by combining (6) and (7), the feature's potential is measured as follows:

$$f^1(X_i) = NI_m(C; X_i) - \frac{1}{|S_{i-1}|} \sum_{X_s \in S_{i-1}} NI_m(X_s; X_i) \tag{9}$$

where, $S = X_1, X_2, \ldots, X_i$ is a set of features.

In this paper, we also combined the normalized CFMI with the $NI_m(X; Y)$ as in [17] to assess the effects of the inequity among the relevance and redundancy. Thus, the goodness of the feature is calculated as below:

$$f^1(X_i) = NI_m(C; X_i) - \frac{1}{|S_{i-1}|} \sum_{X_s \in S_{i-1}} \frac{I_m(X_s; X_i)}{\min(H(X_s), G(X_i))} \tag{10}$$

The step-by-step selection process based on greedy forwarding search scheme is illustrated in algorithm 2. Next, we discuss the dimension reduction.

Algorithm 2: Feature Selection Based Mutual Information Coupled With Greedy Forwarding Searching Method

Input : N^s, K, X^{ij}, C^j, *and* a
Output: S^k

begin
  $S = \phi$
  for $n = 1$ *to* N **do**
    $\sigma = $ *Standard deviation of* X^n
    $\mu^n = $ *Mean value of* X^n
    $X^n = X^n - \mu^n$
    $X^n = X^n / \sigma^n$
  end
  $\overline{X} = Q^{uantiz}(X)$
  for $i = 1$ *to* K **do**
    for $j = 1$ *to* N **do**
      $Computef^a(\overline{X}_j)$
    end
    $s = argmax_{i \notin S(f^a(\overline{X}_i))}$
    $S = S \cup s$
  end
end

C. DIMENSION REDUCTION

Dimension reduction has been extensively studied in the past. Linear Discriminant Analysis(LDA) [23], Generalized Discriminant Analysis(GDA) [24], and Kernel Discriminant Analysis(KDA) [25] are some commonly used dimension reduction methods. However, LDA is widely adopted in human activity recognition cases.

1) LINEAR DISCRIMINANT ANALYSIS

The LDA is used to map the input set over a classification space with the help of optimal linear discriminant function. This space is used to decide the different classification of samples. Moreover, LDA can better handles the cases where indie-class occurrence are not equal. Let the comparison within class be denoted by S_W and between different classes by S_B. S_W and S_B are computed as follows:

$$S_B = \sum_{i=1}^{c} V_i \left(\overline{m}_i - \overline{\overline{m}}\right) \left(\overline{m}_i - \overline{\overline{m}}\right)^T \tag{11}$$

$$S_W = \sum_{i=1}^{c} \sum_{m_k \in C_i} \left(m_k - \overline{m}_i\right) \left(m_k - \overline{m}_i\right)^T \tag{12}$$

Here, V_i is the number of vectors numbers in the class $C_i \forall i < c$, and c is the total number of classes. In this paper, we refer to each activity as a class, and the total number of classes is then given by the total number of activities i.e. $c = numberofdistinctactivities$. Moreover, m_k denotes the vector of a specific class C_i, $\overline{m}$ is the mean of the class, and $\overline{\overline{m}}$ denotes the mean of all the vectors. Whereas D_{opt} represents the optimal discrimination projection matrix which is determined by enlarging the ratio of determinant and within-class scatter matrices.

More details on this topic can be found in [26]. In the following subsection, we discuss our proposed classification method which is based on Hidden Markov Model (HMM).

D. PROPOSED CLASSIFICATION METHOD USING HMM

Among the existing classifiers, the HMM is widely adopted for sequential data classification such as activity recognition [27]. The HMM offers a statistical modeling factor, λ over a set of observation-points which is also called frames in the context of activity recognition. Generally, the HMM has a T frames, where $T = O_1 \ldots O_T$, and a state sequence is denoted by S, such that $S = S_1 \ldots S_N$. Whereas Q represents the set of states, such that $Q = q_1, q_2 \ldots q_N$. N provides the total states number in the model over a time t. Over time t, a transition takes place from S_i state to S_j state with transition probability, a_{ij}. Where Π_j is the initial model probability for any S_j state. Thus, observation, $b_j(O_t)$ is produced with the help of multivariate Gaussian distribution with mean μ_j. Moreover, V_j denotes the covariance matrix and it is correlated with the state. The set of parameters, $\lambda = A, B, \Pi$ represents the HMM, where A denotes the state transition probability which is defined as $A = a_{ij}$, $a_{ij} = Prob(q_{t+1} = S_j | q_t = S_i)$, $1 \leq i, j \leq N$), and B represents the observation probability which is written as $B = b_j(O_t)$, $b_j = Prob(O_t | q_t = S_j) 1 \leq j \leq N$), where Π indicates the initial state probability which is represented $\Pi = \Pi_j$, $\Pi_j = Prob(q_1 = S_1)$ [28]. Moreover, in training phase, the multiplication of a_{ij} with output probability gives the joint likelihood, $P(O|\lambda)$ of state sequence corresponding to its observation-point O, which is evaluated as follows:

$$P(O|\lambda) = \sum_{Q} P(O, Q|\lambda) \tag{13}$$

Eq. (13) can be maximized by obtaining the value of λ, which is based on $\beta_t(j) = P(O_(t+1) \ldots O_T | q_t = j, \lambda)$, $\alpha_t(j) = P(O_1 \ldots O_t, q_t = j|\lambda)$ [29]. These two variables are initiated by the following processes.

$$\alpha_1(j) = \pi_j b_j(O_1), \quad 1 \leq j \leq N \tag{14}$$

$$\beta_T(j) = 1, \quad 1 \leq j \leq N \tag{15}$$

Hence, the trained value of parameter λ can be used to determine the mean of likelihood estimation for the set of sequence observations which provides the labels of activities and its calculated as follows:

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{16}$$

$$\lambda = (O, Q, \pi) \tag{17}$$

Eq. (17) is used to model the parameter valaue λ in HMM. Where O represent the set of total observation-points, such as $O_1, Q_2, \ldots, O_T$).

III. EXPERIMENTAL SETUP

This section describes the experimental setup and evaluation criteria of the proposed technique. We first present the datasets we use for experiments and then present the set of experiments to validate and evaluate our approach. In our experiments and evaluation, we use values for the parameters O, Q, and π as 64, 4, and 4, respectively to model HMM. We choose these values after testing on various combinations of different values for these parameters [30].

A. DATASETS

The performance of the proposed MNF-HAR scheme is validated with the aid of two publicly available standard action datasets; 1) KTH action dataset, and 2) Weizmann action dataset.

- **KTH Action Dataset (KTH-AD)**: In the KTH-AD, total of 25 subjects performed six different activities such as hand-clapping, hand-waving, boxing, running, jogging, and walking under four different indoor and outdoor scenarios. There are a total of 2391 observations taken over homogeneous backgrounds with a static camera. The frame size is 160×120 pixels.
- **Weizmann Action Dataset (Weizmann-AD)**: In Weizmann-AD, total of 9 subjects performed 10 types of different activities (i.e., two-hand-waving, one-hand-waving, side-movement, place-jumping, jumping forward, skipping, running, walking, and bending). In Weizmann-AD, total 90 video clips are used with average of 15 frames in each video clip. The frame size is 144×180 pixels.

B. EXPERIMENTS OVERVIEW

We performed the following experiments for a thorough testing and validation of the system.

- In first experiment, we assess the significance of the proposed MNF-HER system against each datasets individually. For each dataset, we exploit $10-$fold cross-validation method for each dataset, in which one out of

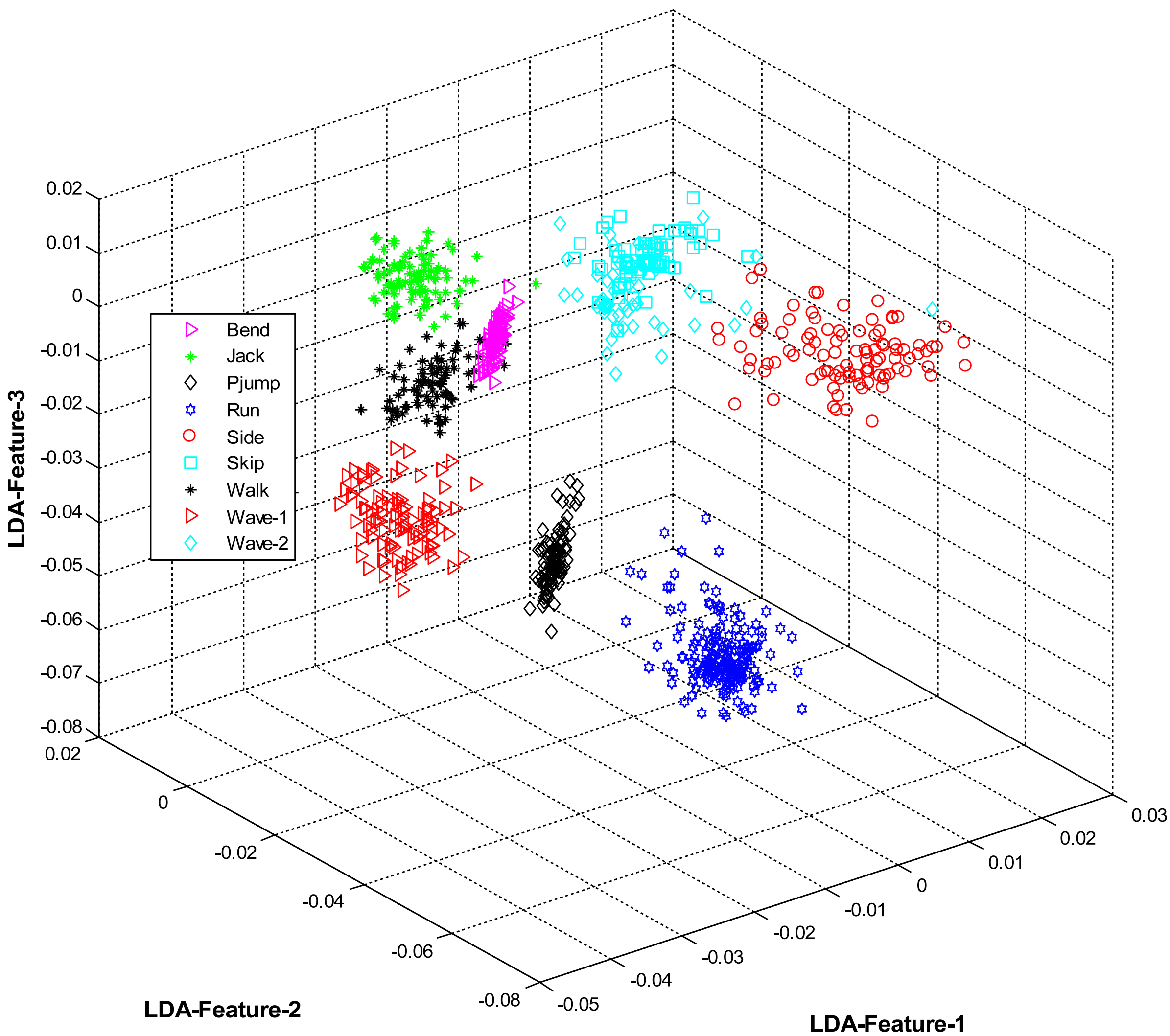

FIGURE 2. 3D plot for the proposed MNF-HAR system using Weizmann action dataset.

ten subjects is employed for validation and remaining nine are used for training. 10 times, this process is repeated, and data from each individual subject is only used once once for validation.

- In second experiment, the performance of the proposed feature selection method is evaluated. We perform the experiments over the both datasets under the absence of the proposed feature selection method. This experiment basically shows the importance of the proposed method.
- Finally, we compare the proposed MNF-HAR and the weighted average recognition rate under the state-of-the-art techniques.

IV. EXPERIMENTAL EVALUATION

A. MNF-HAR PERFORMANCE

We test our proposed system on the two datasets described above. Figure 2 and Table 2 depict the results of MNF-HAR on the first dataset, and Figure 3, Table 3 depicts the results over the second dataset.

From Tables 2 and 3 it is clear that the proposed MNF-HAR scheme obtain better recognition rate on each dataset, which means that the proposed scheme is more robust in achieving the higher recognition rate on multiple datasets. The recognition rate on Weizmann action dataset is 98.22% and 99% on KTH action dataset. This accuracy is a significant improvement over the existing approaches.

B. MNF-HAR FEATURE SELECTION PERFORMANCE

This experiment presents the importance of the proposed mutual information feature selection method in the MNF-HAR system. In this experiment, we test and validate the performance of the MNF-HAR system under the absence of the proposed feature selection method on both datasets. In this experiment, the extracted features (i.e., the output of curvelet transform) were fed directly to HMM without invoking the feature selection process and Table 4 and 5 provides the overall results on two datasets.

TABLE 2. Confusion matrix for proposed MNF-HAR using Weizmann action dataset. It can be seen that the proposed HAR system showed significant performance in classification rate (Unit: %).

Activities	Wave2	Wave1	Walk	Skip	Side	Run	Pjump	Jack	Bend
Wave2	**96**	0	1	1	0	1	0	1	0
Wave1	0	**98**	0	0	2	0	0	0	0
Walk	0	0	**99**	0	0	0	0	0	1
Skip	0	1	0	**99**	0	0	0	0	0
Side	0	0	0	0	**100**	0	0	0	0
Run	0	0	1	0	0	**97**	2	0	0
Pjump	0	0	0	1	0	0	**99**	0	0
Jack	0	0	1	0	2	0	1	**96**	0
Bend	0	0	0	0	0	0	0	0	**100**
Average					**98.22**				

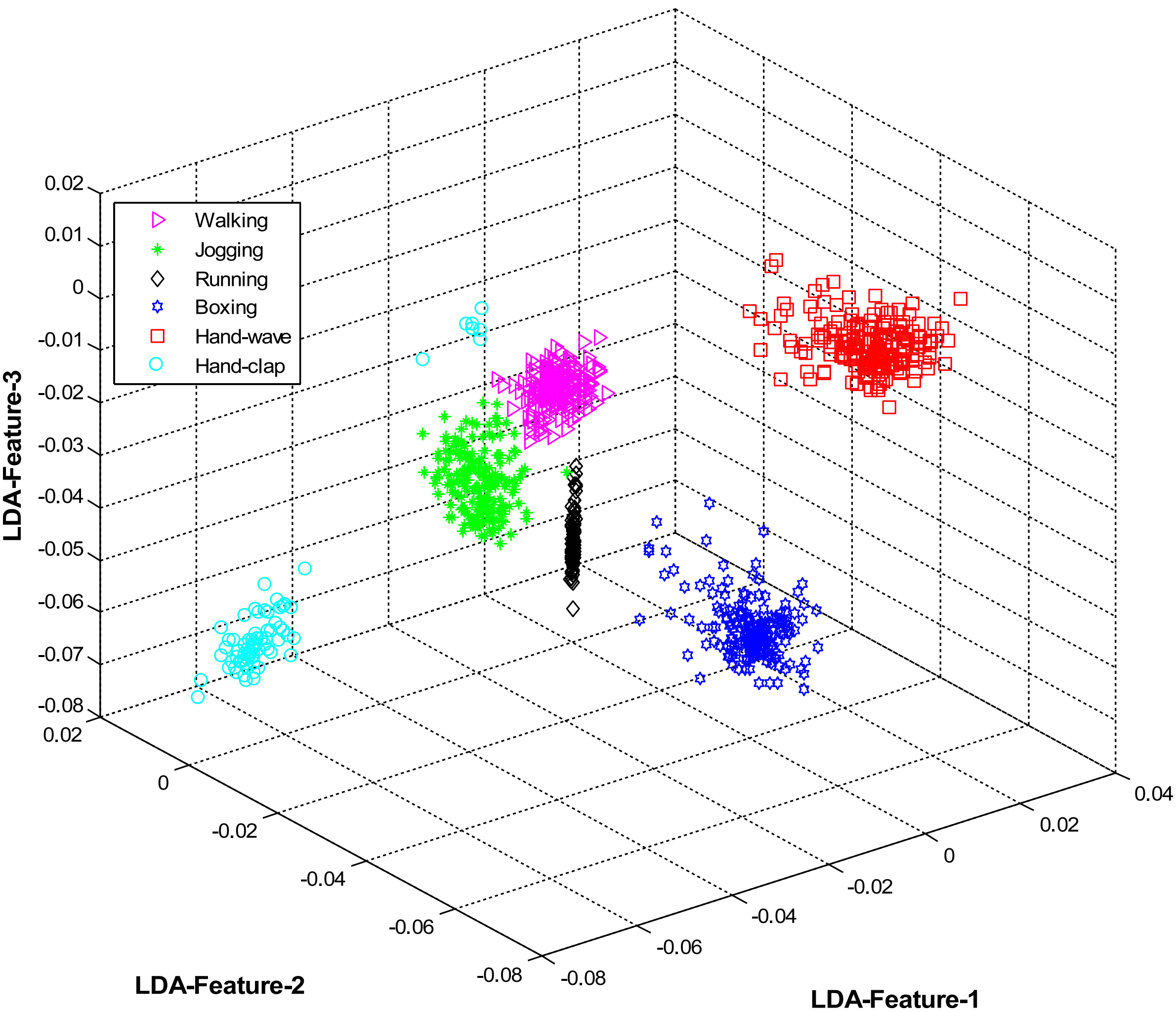

FIGURE 3. Proposed MNF-HAR scheme using KTH action dataset.

The proposed mutual information feature selection method performed major role and has a significant contribution in the high recognition of the MNF-HAR as shown in Tables 4 and 5. It is clear that in the absence of the proposed method the recognition rate significantly decreased that validate the high similarity problem among the feature of various activities. Thus, the simulation results show that the proposed scheme is more robust in terms

TABLE 3. Confusion matrix for proposed MNF-HAR using KTH action dataset. It can be seen that the proposed HAR system shows significant performance in classification rate (Unit: %).

Activities	Hand-clap	Hand-wave	Boxing	Running	Jogging	Walking
Hand-clap	**100**	0	0	0	0	0
Hand-wave	0	**99**	1	0	0	0
Boxing	0	0	**98**	1	1	0
Running	0	1	0	**98**	1	0
Jogging	0	0	0	0	**100**	0
Walking	0	0	0	1	0	**99**
Average			**99.00**			

TABLE 4. Confusion matrix for proposed MNF-HAR using Weizmann action dataset under the absence of the proposed mutual information feature selection method (Unit: %).

Activities	Wave2	Wave1	Walk	Skip	Side	Run	Pjump	Jack	Bend
Wave2	**95**	0	1	1	0	1	0	1	0
Wave1	0	**91**	0	0	2	0	0	0	0
Walk	0	0	**90**	0	0	0	0	0	1
Skip	0	1	0	**92**	0	0	0	0	0
Side	0	0	0	0	**94**	0	0	0	0
Run	0	0	1	0	0	**89**	2	0	0
Pjump	0	0	0	1	0	0	**90**	0	0
Jack	0	0	1	0	2	0	1	**93**	0
Bend	0	0	0	0	0	0	0	0	**92**
Average					**98.22**				

TABLE 5. Confusion matrix for proposed MNF-HAR using KTH action dataset under the absence of the proposed mutual information feature selection method (Unit: %).

Activities	Hand-clap	Hand-wave	Boxing	Running	Jogging	Walking
Hand-clap	**91**	0	0	0	0	0
Hand-wave	0	**89**	1	0	0	0
Boxing	0	0	**92**	1	1	0
Running	0	1	0	**97**	1	0
Jogging	0	0	0	0	**90**	0
Walking	0	0	0	1	0	**87**
Average			**99.00**			

of classification accuracy and it selects the better feature set.

C. COMPARISON TO THE STATE-OF-THE-ART METHODS

In this subsection, we compare our proposed MNF-HAR system with other state-of-the-art methods over both datasets. The implementations for other schemes are borrowed or used publicly available codes. Moreover, for a fair comparison we reported their published results. Furthermore, all implementations are modified to perform the simulations for same settings as given in their respective works. Tables 6 and 7 summaries the average classification rates of the proposed MNF-HAR using Weizmann and KTH action datasets respectively.

It is clear from Tables 6 and 7 that the proposed MNF-HAR scheme provide better performance and achieved significant recognition rate than the other state-of-the-art schemes against two datasets. Hence, our proposed system has a great potential to accurately recognize human activities by using video data.

TABLE 6. Comparison results of the proposed MNF-HAR with state-of-the-art methods against Weizmann action dataset (Unit: %).

State-of-the-art Works	Average Classification Rates	Standard Deviation
[31]	84.3	±3.2
[32]	78.3	±2.6
[33]	92.5	±3.2
[34]	83.3	±2.9
[35]	75.0	±1.9
Proposed MNF-HAR	**98.2**	±1.8

TABLE 7. Comparison results of the proposed MNF-HAR with state-of-the-art methods against KTH action dataset (Unit: %).

State-of-the-art Works	Average Classification Rates	Standard Deviation
[31]	92.1	±1.2
[36]	90.5	±2.3
[37]	90.4	±4.6
[33]	94.1	±3.2
[38]	93.7	±2.7
Proposed MNF-HAR	**99.0**	±1.0

V. CONCLUSION

Currently the video-based Human Activity Recognition (HAR) is one of the hottest candidates that can be employed in many applications of image processing and computer vision. The accuracy of a general HAR systems is reliant on the selection of the most relevant and informative features. Most of the existing works forced to develop methods but a very few works can be found on feature selection. Therefore, in this paper, we thoroughly reviewed and investigated the recent feature selection algorithms and developed a new feature selection method that derived from normalized mutual information feature selection (NMIFS). The proposed feature selection method has two improvements; 1) feature-independent normalizing weights, and 2) the normalization of the mutual information. We used curvelet transform for the feature extraction and LDA has been utilized for dimension reduction. Finally, activities are classified by using HMM. The performance of the proposed scheme is validated on two different datasets. It is clear from our simulation results that the proposed scheme achieves 98% accuracy. In our used datasets, mostly RGB cameras are used which may raise privacy concern. Hence, this concern can be avoided in future by using the depth camera's. Hence, the accuracy and robustness of the proposed system need to be investigated under new settings.

REFERENCES

[1] M. H. Siddiqi, A. M. Khan, and S.-W. Lee, "Active contours level set based still human body segmentation from depth images for video-based activity recognition," *KSII Trans. Internet Inf. Syst.*, vol. 7, no. 11, pp. 2839–2852, 2013.

[2] R. Kamimura, "Structural enhanced information and its application to improved visualization of self-organizing maps," *Appl. Intell.*, vol. 34, no. 1, pp. 102–115, 2011.

[3] A. Jamshidnezhad and M. J. Nordin, "An adaptive learning model based genetic for facial expression recognition," *Int. J. Phys. Sci.*, vol. 7, no. 4, pp. 619–623, 2012.

[4] (1995). *Advantages and Disadvantages of Genetic Algorithms*. Accessed: Feb. 2, 2016. [Online]. Available: http://www.ro.feri.uni-mb.si/predmeti/int_reg/Predavanja/Eng/3.Genetic%20algorithm/_18.html

[5] M. F. Valstar, T. Almaev, J. M. Girard, G. McKeown, M. Mehu, L. Yin, M. Pantic, and J. F. Cohn, "FERA 2015—Second facial expression recognition and analysis challenge," in *Proc. 11th IEEE Int. Conf. Workshops Autom. Face Gesture Recognit. (FG)*, vol. 6, May 2015, pp. 1–8.

[6] Y. Wang, I. V. Tetko, M. A. Hall, E. Frank, A. Facius, K. F. X. Mayer, and H. W. Mewes, "Gene selection from microarray data for cancer classification—A machine learning approach," *Comput. Biol. Chem.*, vol. 29, no. 1, pp. 37–46, 2005.

[7] Z. M. Hira and D. F. Gillies, "A review of feature selection and feature extraction methods applied on microarray data," *Adv. Bioinf.*, vol. 2015, Mar. 2015, Art. no. 198363.

[8] S. Fong, S. Deb, X.-S. Yang, and J. Li, "Feature selection in life science classification: Metaheuristic swarm search," *IT Prof.*, vol. 16, no. 4, pp. 24–29, Jul./Aug. 2014.

[9] Z. Ji, G. Meng, D. Huang, X. Yue, and B. Wang, "NMFBFS: A NMF-based feature selection method in identifying pivotal clinical symptoms of hepatocellular carcinoma," *Comput. Math. Methods Med.*, vol. 2015, Jul. 2015, Art. no. 846942.

[10] E. Candès, L. Demanet, D. Donoho, and X. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, Sep. 2006.

[11] F. Bellal, H. Elghazel, and A. Aussem, "A semi-supervised feature ranking method with ensemble learning," *Pattern Recognit. Lett.*, vol. 33, no. 10, pp. 1426–1433, 2012.

[12] S. Gharsalli, B. Emile, H. Laurent, X. Desquesnes, and D. Vivet, "Random forest-based feature selection for emotion recognition," in *Proc. Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Nov. 2015, pp. 268–272.

[13] H. Deng, G. Runger, and E. Tuv, "Bias of importance measures for multi-valued attributes and solutions," in *Artificial Neural Networks and Machine Learning—ICANN*. 2011, pp. 293–300.

[14] S. Fong, J. Liang, I. Fister, Jr., I. Fister, and S. Mohammed, "Gesture recognition from data streams of human motion sensor using accelerated PSO swarm search feature selection algorithm," *J. Sensors*, vol. 2015, Mar. 2015, Art. no. 205707.

[15] K. Y. Lee and J.-B. Park, "Application of particle swarm optimization to economic dispatch problem: Advantages and disadvantages," in *Proc. IEEE PES Power Syst. Conf. Expo. (PSCE)*, Oct./Nov. 2006, pp. 188–192.

[16] R. Battiti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Trans. Neural Netw.*, vol. 5, no. 4, pp. 537–550, Jul. 1994.
[17] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *IEEE Trans. Neural Netw.*, vol. 20, no. 2, pp. 189–201, Feb. 2009.
[18] N. Kwak and C.-H. Choi, "Input feature selection for classification problems," *IEEE Trans. Neural Netw.*, vol. 13, no. 1, pp. 143–159, Jan. 2002.
[19] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.
[20] C. Schuldt, I. Laptev, and B. Caputo, "Recognizing human actions: A local SVM approach," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, Cambridge, U.K., vol. 3, Aug. 2004, pp. 32–36.
[21] M. Blank, L. Gorelick, E. Shechtman, M. Irani, and R. Basri, "Actions as space-time shapes," in *Proc. 10th IEEE Int. Conf. Comput. Vis.*, Beijing, China, vol. 2, Oct. 2005, pp. 1395–1402.
[22] M. Tang and F. Chen, "Facial expression recognition and its application based on curvelet transform and PSO-SVM," *Optik*, vol. 124, no. 22, pp. 5401–5406, 2013.
[23] S. Mika, "Kernel Fisher discriminants," Ph.D. dissertation, Universitätsbibliothek, 2002.
[24] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Comput.*, vol. 12, no. 10, pp. 2385–2404, 2000.
[25] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K. Müllers, "Fisher discriminant analysis with kernels," in *Proc. IEEE Signal Process. Soc. Workshop.*, Aug. 1999, pp. 41–48.
[26] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
[27] M. H. Siddiqi, R. Ali, M. S. Rana, E.-K. Hong, E. S. Kim, and S. Lee, "Video-based human activity recognition using multilevel wavelet decomposition and stepwise linear discriminant analysis," *Sensors*, vol. 14, no. 4, pp. 6370–6392, 2014.
[28] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
[29] L. E. Baum, "An equality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes," *Inequalities*, vol. 3, pp. 1–8, 1972.
[30] F. S. Samaria, "Face recognition using hidden Markov models," Ph.D. dissertation, Univ. Cambridge, Cambridge, U.K., 1994.
[31] A. Roy, B. Banerjee, and V. Murino, "A novel dictionary learning based multiple instance learning approach to action recognition from videos," in *Proc. 6th Int. Conf. Pattern Recognit. Appl. Methods (ICPRAM)*, 2017, pp. 519–526.
[32] C. Gu, C. Sun, D. A. Ross, C. Vondrick, C. Pantofaru, Y. Li, S. Vijayanarasimhan, G. Toderici, S. Ricco, R. Sukthankar, C. Schmid, and J. Malik, "Ava: A video dataset of spatio-temporally localized atomic visual actions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6047–6056.
[33] S. Dharmalingam and A. Palanisamy, "Vector space based augmented structural kinematic feature descriptor for human activity recognition in videos," *ETRI J.*, vol. 40, no. 4, pp. 499–510, Aug. 2018.
[34] A. P. Sirat, "Activity recognition using particle network," *Int. J. Appl. Innov. Eng. Manage.*, vol. 7, no. 8, pp. 49–54, Aug. 2018.
[35] B. Lakshmi and S. Parthasarathy, "Complex human activity recognition using Gaussian polar coordinate and sequence likelihood support classifier," *Int. J. Creative Res. Thoughts*, vol. 6, no. 2, pp. 767–790, 2018.
[36] K.-P. Chou, M. Prasad, D. Wu, N. Sharma, D.-L. Li, Y.-F. Lin, M. Blumenstein, W.-C. Lin, and C.-T. Lin, "Robust feature-based automated multi-view human action recognition system," *IEEE Access*, vol. 6, pp. 15283–15296, 2018.
[37] V.-D. Hoang, "Multiple classifier-based spatiotemporal features for living activity prediction," *J. Inf. Telecommun.*, vol. 1, no. 1, pp. 100–112, 2017.
[38] M. Elmezain and A. Al-Hamadi, "Vision-based human activity recognition using LDCRFs," *Int. Arab J. Inf. Technol.*, vol. 15, no. 3, pp. 1683–3198, 2018.

MUHAMMAD HAMEED SIDDIQI received the bachelor's degree (Hons.) in computer science from Islamia College Peshawar (Chartered University), Peshawar, Pakistan, in 2007, and the master's and Ph.D. degrees from the Ubiquitous Computing (UC) Lab, Department of Computer Engineering, Kyung Hee University, Seoul, South Korea, in 2012 and 2016, respectively. He was a Graduate Assistant with Universiti Teknologi PETRONAS, Malaysia, from 2008 to 2009. He was a Postdoctoral Research Scientist with the Department of Computer Science and Engineering, Sungkyunkwan University, Suwon, South Korea, from March 2016 to August 2016. He is currently an Assistant Professor with the College of Computer and Information Sciences, Jouf University, Sakaka, Saudi Arabia. He has published more than 50 research articles in highly reputable international journals and conferences. His research interest is image processing, pattern recognition, machine intelligence, activity recognition, and facial expression recognition. He is also Reviewer for different journals and conferences.

MADALLAH ALRUWAILI received the bachelor's degree (Hons.) from Jouf University, Saudi Arabia, in 2005, the M.S. degree from the University of Science, Malaysia, in 2009, and the Ph.D. degree from Southern Illinois University, Carbondale, IL, USA, in 2015. His Ph.D. dissertation entitled "Enhancement and Restoration of Dust Images." He is currently an Assistant Professor of computer engineering and networks with Jouf University. He is also the Dean of the College of Computer and Information Sciences. His research interests include image processing, image quality analysis, pattern recognition, computer vision, and biomedical imaging.

AMJAD ALI received the B.S. and M.S. degrees in computer science from the COMSATS Institute of Information Technology, Pakistan, in 2006 and 2008, respectively, and the Ph.D. degree from the Electronics and Radio Engineering Department, Kyung Hee University, South Korea, in 2015. From 2018 to 2019, he was a Postdoctoral Research Scientist with the UWB Wireless Communications Research Center (formerly the Key National IT Research Center), Department of Information and Communication Engineering, Inha University, South Korea, and also with the Mobile Network and Communications Lab, School of Electrical Engineering, Korea University, Seoul, South Korea. Since 2015, he has been an Assistant Professor with the Department of Computer Science, COMSATS University Islamabad at Lahore, Pakistan. He has published many peer-reviewed international journals and conference papers. His research interests include the Internet of Things, cognitive radio networks (interference modeling, dynamic spectrum access, power and admission control, spectrum management, spectrum trading, MAC protocols, performance modeling, and optimization), 5G cellular networks (spectrum/resource management, coexistence, distributed wireless access, scheduling, power control, network selection, and mobility/handover management), ad hoc social networks, multimedia cloud computing, smart grid, and vehicular networks. He received the Excellent Research Contribution Award from the UWB Wireless Communications Research Center, for his excellent research performances. He is serving as a Reviewer for the *IEEE Communications Magazine*, the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, the IEEE TRANSACTIONS ON CIRCUITS AND Systems for Video Technology, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE INTERNET OF THINGS JOURNAL, IEEE ACCESS, *IET Communications*, and *Computer Networks*. He is also serving as an Editorial Board Member for IEEE ACCESS.