IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 28, 2020, accepted January 6, 2021, date of publication January 13, 2021, date of current version January 22, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051281

Multi-Domain Feature Fusion for Emotion Classification Using DEAP Dataset

MUHAMMAD KHATEEB[1,2], SYED MUHAMMAD ANWAR[3], AND MAJDI ALNOWAMI[4], (Member, IEEE)

[1]Department of Software Engineering, University of Engineering and Technology, Taxila 47050, Pakistan
[2]Faculty of Computing, Riphah International University, Islamabad 46000, Pakistan
[3]Department of Software Engineering, University of Engineering and Technology, Taxila 47050, Pakistan
[4]Department of Nuclear Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: Syed Muhammad Anwar (s.anwar@uettaxila.edu.pk)

ABSTRACT Emotion recognition in real-time using electroencephalography (EEG) signals play a key role in human-computer interaction and affective computing. The existing emotion recognition models, that use stimuli such as music and pictures in controlled lab settings and limited number of emotion classes, have low ecological validity. Moreover, for effective emotion recognition identifying significant EEG features and electrodes is important. In our proposed model, we use the DEAP dataset consisting of physiological signals collected from 32 participants as they watched 40 movie (each of 60 seconds) clips. The main objective of this study is to explore multi-domain (time, wavelet, and frequency) features and hence, identify the set of stable features which contribute towards emotion classification catering to a larger number of emotion classes. Our proposed model is able to identify nine classes of emotions including happy, pleased, relaxed, excited, neutral, calm, distressed, miserable, and depressed with an average accuracy of 65.92%. Towards this end, we use support vector machine as a classifier along with 10-fold and leave-one-out cross-validation techniques. We achieve a significant emotion classification accuracy which could be vital towards developing solutions for affective computing and deal with a larger number of emotional states.

INDEX TERMS Affective computing, electroencephalography, emotions classifications, features extraction, machine learning.

I. INTRODUCTION

Affective computing is a specialized field of artificial intelligence (AI) and is used for processing, interpreting, and identifying emotional states. Emotions play a vital role in our daily life activities including decision making, communication, and personal development. Although emotions are natural to us as human beings, a significant attention has been given for detecting emotions during human-robot interaction to enable affective computing [1]. In a seminal work, it was shown that computers that can recognise and respond to human emotions are critical for the progress of human computer interaction [2]. Moreover, it was shown that analysis of affective physiological signals can lead towards machine intelligence. In particular, it was observed that using physiological signals can be more beneficial for machine intelligence than using vocal or visual data [3]. For smart robots, using neural codes could be a way forward to process complex information from real world [4], where physiological signals could be significant for emotion recognition. For AI systems, the field of affective computing is leading the efforts towards developing efficient techniques for emotion recognition [5]. To this end, different methods have been proposed that use speech, facial expressions, and skin conductance for an automated emotion recognition. For emotion classification using speech, different models have been adopted such as hidden Markov model and artificial neural networks [6]. In such cases, the performance is effected by utterance as well as window size selection. Emotion recognition and analysis using facial expressions has been used to identify various classes of emotions [7]. An automated facial expression recognition model was designed based on features extracted using static, dynamic, and geometry based facial characteristics [8].

There are two taxonomy models used to study and discern various emotional states: 1) discrete model and

The associate editor coordinating the review of this manuscript and approving it for publication was Bohui Wang.

2) dimensional model [9]. The discrete model incorporates limited emotional states represented as six basic emotions including joy, sadness, surprise, fear, anger, and disgust. In the dimensional model, emotional states are divided in a two dimensional space represented as valence and arousal. The valence scale ranges from sad to joyful, and the arousal scale ranges from bored to excited. The valence arousal (VA) space is systematically mapped to discrete emotions, where various emotional states can be identified [10].

Recently, physiological signals such as electroencephalography (EEG) have been used for classification of emotions [11]. EEG is easy to use, cost effective, mobile (with the availability of off-the-shelf wearable headsets), and requires lesser physical constraints [12]. Since EEG records the underlying human brain activity, it is considered reliable for emotion recognition systems. In comparison, those techniques which are based on facial expression and speech are subject to human subjectivity and manipulation. Hence, there is recent evidence of efforts towards identifying human emotions in real-time using EEG signals [10], [13]. In [14], signals were obtained from five EEG channels (FP1, AF3, P3, FC2, and O1) and the proposed model provided high performance in emotion classification. In another study, selective frequency sub-bands (alpha, beta, and gamma) were shown to classify emotions more accurately [15]. A relationship was shown to exist between the emotional states and the asymmetric ratios between EEG electrodes for detecting emotional states [16], [17]. In systems using EEG signals, a noise free data acquisition could be a challenge. In particular, when using EEG electrodes placed within a cap, subjects are known to suffer from discomfort which could translate in to motion artifacts. Moreover, with a high density electrode placement, it is not always necessary that data from each of those electrodes contribute significantly towards emotion recognition [11]. A comprehensive review on feature and electrode selection algorithms was presented [18]. The focus was on developing an emotion recognition system using wearable mobile EEG sensors. While the best results were reported for differential entropy-based features, it was shown that feature and electrode selection was a significant step for developing such systems. Hence, there is a need to develop methods that can identify relevant features and electrodes which would play a major role towards emotion classification.

OUR CONTRIBUTIONS

There exists a need to further improve emotion classification models using EEG data and identification of stable features which are suitable for this purpose. Towards this overreaching goal, we propose an emotion recognition technique which is designed based on the DEAP dataset [19]. DEAP is a large collection of physiological signals recorded with the aim of developing emotion recognition systems. We used data from selected channels (those that show significant contribution for emotion recognition) and extracted features from three (time, frequency, and wavelet) different domains. For classification, we used individual as well as hybrid features and in the process identified stable features and frequency bands. In our experiments, we obtained nine discrete states of emotions mapped from the VA space. We used support vector machine (SVM) for classification using 10-fold and leave-one-out cross-validation (LOOCV) techniques. For the classification of nine emotional states, our proposed method achieved an average accuracy of 65.72% (10-fold) and 65.92% (LOOCV), which to the best of our knowledge is the highest among state-of-the-art methods. Our proposed method has the following contributions,

1) We used hybrid features (selected after carefully designed experiments) from time, frequency, and wavelet domains which improved the accuracy for emotion classification.
2) Using grid search, we selected four EEG electrodes (FP1, FP2, F3, and C4), out of 32, to best suite the classification problem at hand.
3) Our proposed model classified nine different emotions (happy, pleased, relaxed, excited, neutral, calm, distressed, miserable, and depressed) with the highest accuracy among state-of-the-art methods.

II. RELATED WORK

Affective computing is an emerging field where a large number of studies have been conducted for emotion classification [10], [11], [20]–[22]. An accurate classification of emotions is a significant factor in developing effective brain-computer interface systems. Emotions have a key role in determining the stress level and personality development of an individual as suggested by modern neuro-science and brain analysis studies [23]. Specific emotions could be evoked in response to viewing a certain video content or listening to a certain genre of music [24]. Recent studies have used the human nervous system to identify emotional states through physiological signals observed using EEG [25]. In [26], participant dependent and independent emotional states were identified using stable patterns of EEG, which were represented using the neural activity related to critical brain areas and frequency bands. A total of six different features were extracted and minimum redundancy maximum relevance parameter was used for feature selection. The method achieved an accuracy of 69.67% on DEAP dataset for four valance/arousal states.

A model used hybrid features from time, frequency, and wavelet domains with audio music as an external stimuli and achieved an accuracy of 94% using multi-layer perceptron for seven discrete emotions [27]. A standardized film clip database was used as a stimuli, involving 9 participants, for classifying three positive emotions (joy, amusement, tenderness) and four negative emotions (anger, disgust, fear, sadness) with an accuracy of 86.43% [6]. In [16], three features were computed using EEG signals from DEAP dataset and the model achieved an accuracy of 73.5% for two valance and arousal states. In [28], two classifiers based on neural networks were used and achieved an accuracy of 71.00%

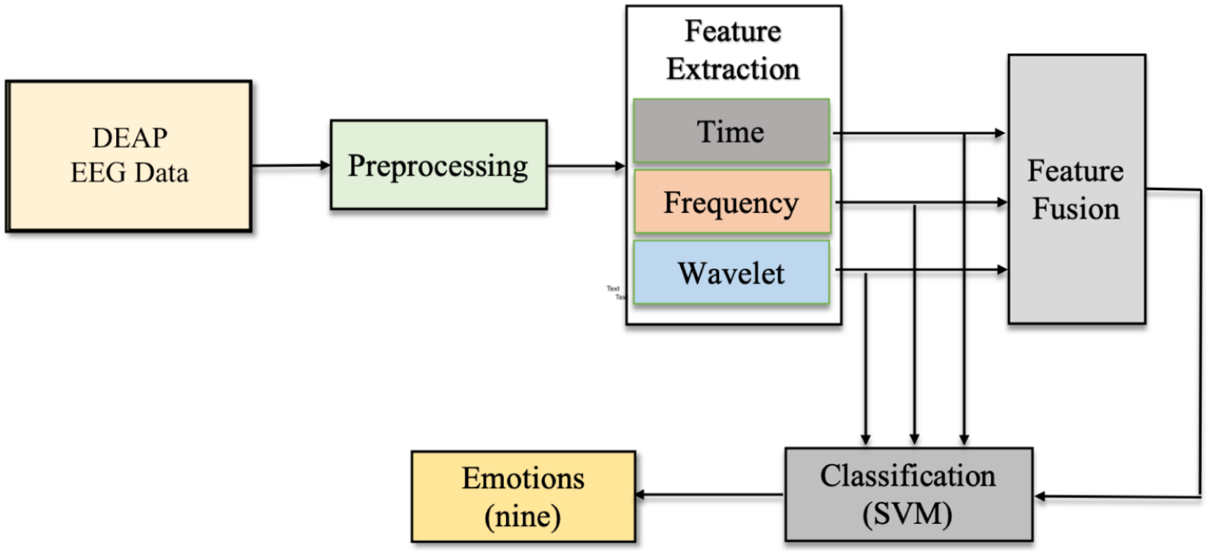


FIGURE 1. **The proposed emotion recognition model based on features extracted from EEG signals.**

for two-class emotion classification using DEAP dataset. In [29], a novel approach based on using combined classifiers i.e., SVM and hidden Markov model (HMM), along with features extracted from time and discrete wavelet domain improved the emotion classification accuracy.

To summarize, there are multiple methods proposed for EEG based emotion classification, but have certain limitations including the number of emotion states considered. For methods that deal with a higher number of classes for a greater number of participants, a decrease in performance is observed. Although there are certain methods with limited number of participants that have reported higher accuracies (>80%), the generalization of such methods needs more careful consideration. Towards this end, we have used the DEAP dataset which could be considered as a baseline for emotion recognition and contains a large collection of data from 32 precipitants. We report a significant performance in emotion classification when a large number of emotional states are considered.

III. PROPOSED METHODOLOGY

Our proposed methodology includes a machine learning approach for emotion classification using DEAP dataset. Towards this end, we used multi-domain features both individually and in a hybrid setting. We incorporated preprocessing in our model for cleaning the EEG data. Feature selection was used to come up with a carefully selected set of features which are best suited to the classification task. The details of the sub-blocks (Figure 1) are presented in the following subsections.

A. MATERIALS

In our experiments, we used the DEAP dataset [19], which consists of EEG based physiological signals. A total of 32 participants (50% of male and 50% female) with an average age of 26.9 years (with age ranging between 19-37) participated in the study. All participants signed a consent prior to experiments and filled out the questionnaires. A set of instructions were provided to all participants for recording the self-assessment reports. The EEG data from all participants was collected while watching videos. A total of 40 videos, each represented with a specific ID and covering different genres, were presented. All participants watched videos in a sequence where each video lasted for 60 seconds. The data obtained from EEG electrodes have two signal arrays: 1) The data array ($40 \times 40 \times 8064$) represents that a user watched 40 videos and 40 channels of EEG were used to collect a total of 8064 data samples. 2) The second array contains four target labels (valence, arousal, dominance, and liking) for each video.

The level of valence, arousal, dominance, and liking was assessed using the self-assessment manikin (SAM) scale [30]. In our experiments, we used the valence and arousal ratings and ignored the dominance and liking values. For both valence and arousal, the rating scale varies from 1-9. We divided the scale into three levels of valence and arousal. For both valence and arousal, the ratings in the range 1-3 were mapped as "LOW", 4-6 were mapped as "MEDIUM", and 7-9 were mapped as "HIGH". Hence a total of nine emotions (Figure 2) were used in the model including distressed, excited, happy, miserable, neutral, pleased, depressed, clam, and relaxed.

B. EEG ELECTRODE SELECTION

The EEG electrodes were placed according to 10-20 system, where a symbol is assigned to the lobe on each area of scalp and a number is assigned to identify the hemispheric locations (Figure 3). Different EEG electrodes covering various

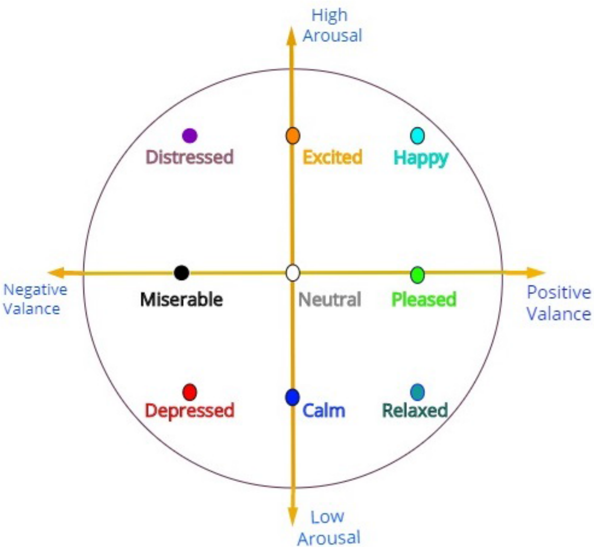


FIGURE 2. Emotion states used in this study based on the valance-arousal score level.

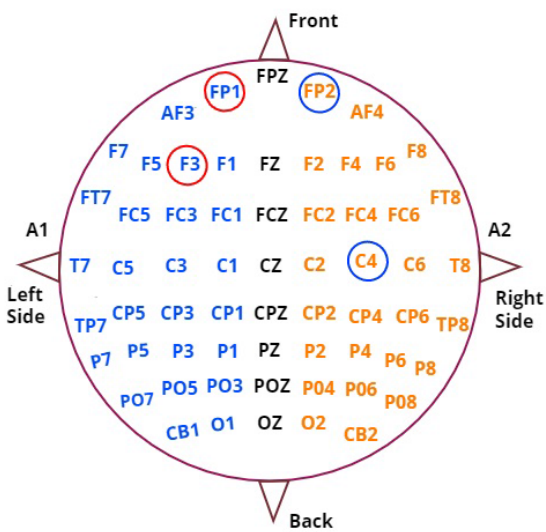


FIGURE 3. The EEG electrode placement according to 10-20 electrode placement system. C: central, O: Occipital, P: Parietal, T: Temporal, Fp: Pre-frontal, even numbers: right hemisphere, odd numbers: left hemisphere, the selected electrodes are encircled.

brain regions are dedicated to specific kind of emotions. In [31], a normalized mutual information method selected the best electrodes for observing both valence and arousal. In [32], it was shown that the left and right hemispheres and cerebral laterality plays a prominent role in enticing emotions, while pre-frontal cortex significantly effects emotion regulations [33]. In [34], five electrodes (P3, FC2, AF3, O1 and FP1) were used and achieved a significant accuracy in emotion recognition. This shows that data from certain brain regions is more significant for emotion classification than others. In our proposed model we have used four EEG channels, giving the best performance, and were selected using a grid search approach. Towards this end, we started with 32 electrodes, removing one electrode at a time and recording the classification performance using the time domain features. After performing grid search, four electrodes were selected including FP1, FP2, F3, and C4. To further verify this, we performed grid search approach using the hybrid feature vector (after feature fusion) and the same four electrodes were found significant.

C. EEG SIGNAL PRE-PROCESSING

The data was collected through an EEG device and contains various artefacts. The data was down-sampled to $128Hz$ and a band-pass filter was used with pass-band frequency of $2Hz - 45Hz$ using EEG-lab toolbox. Following the procedure in [19], blind separation technique was used for removing eye movement artefacts. The signals contained 8064 data points and were windowed to represent 60 seconds of recording. After pre-processing these windowed samples were used for feature extraction.

D. FEATURE EXTRACTION

Emotions are complex brain functions and are evoked by neuronal activation within different brain regions. From the perspective of machine learning, obtaining salient information from EEG data requires appropriate feature extraction. These feature can be extracted from multiple domain (such as time, frequency, and wavelet) and finding the right combination of these features requires extensive experimentation. While it is shown that EEG signals (as a time series) have representative information in both time and frequency (EEG subbands) domains [35] and therefore different signal domains can contain complementary data. To this end, we used time-, frequency-, and wavelet-domain, as well as asymmetric features to be well representative of the underlying EEG data. On top of this, we used features that are well established in these domains for EEG based studies and are known to be more representative. The features extracted from time, frequency, and wavelet domains are discussed below for completeness.

1) TIME DOMAIN FEATURES

We extracted Hjorth parameters and entropy features from the time domain. The Hjorth parameters [36], [37] including activity (A_h), mobility (M_h), and complexity (C_h) were calculated using,

$$A_h = var(s_j), \tag{1}$$

$$M_h = \sqrt{\frac{var(s'_j)}{var(s_j)}}, \tag{2}$$

$$C_h = \frac{M'_h}{M_h}, \tag{3}$$

where activity is defined as the variance of the input signal, $var(s'_j)$ represent the variance of the first derivative of input signal, $var(s_j)$ represents the variance of the signal, and M'_h is the mobility of the first derivative of the input signal (s'_j).

TABLE 1. Average accuracy using wavelet and frequency domain features individually for different frequency sub-bands and cross validation techniques.

Cross-Validation	Features	$\alpha(\%)$	$\beta((\%)$	$\gamma(\%)$	$\alpha,\beta(\%)$	$\alpha,\gamma(\%)$	$\beta,\gamma(\%)$	$\alpha,\beta,\gamma(\%)$
10-fold	Power	60.08	61.24	61.06	61.15	61.45	61.20	61.94
	RASM	63.65	63.82	63.68	63.90	63.64	63.70	63.75
	DASM	63.20	63.24	63.47	63.31	63.09	62.98	62.98
	RASM, DASM	62.22	63.34	63.21	63.31	63.45	63.28	63.88
	Wavelet (energy)	63.05	63.21	63.08	63.25	63.24	63.17	63.50
	Wavelet (entropy)	63.23	63.60	63.57	64.58	64.58	64.98	**65.01**
	Wavelet (entropy, energy)	63.63	63.81	64.67	64.34	64.58	65.05	**65.14**
	Differential Entropy	63.53	63.01	63.24	63.59	63.76	63.50	63.82
LOOCV	Power	63.10	63.30	63.09	64.07	62.88	64.98	65.00
	RASM	63.07	63.81	63.60	63.94	63.67	63.65	63.69
	DASM	63.20	63.44	63.13	63.20	63.27	63.16	63.28
	RASM, DASM	63.02	63.41	63.03	63.31	63.35	63.20	63.53
	Wavelet (energy)	63.06	63.00	63.01	63.05	63.06	63.04	62.95
	Wavelet (entropy)	63.13	63.70	63.28	63.70	63.99	65.15	**65.19**
	Wavelet (entropy, energy)	64.48	64.82	64.67	63.95	64.20	65.15	**65.23**

For extracting entropy feature, we used windowing to represent the EEG signal samples in 10 equal parts (with no overlap) and calculated the entropy value for each sub-window using,

$$Entropy = -\sum_{j=1}^{n_1} p(s_j) log_b p(s_j), \tag{4}$$

where n_1 represents $1/10^{th}$ of the total number of samples (n) and s_j represents the EEG signal.

2) FREQUENCY DOMAIN FEATURES

The EEG signals are non-stationary and non-linear, therefore for better representing these signals we used features in the frequency domain using short-time Fourier transform. Herein, we extracted the frequency sub-bands including alpha ($8-15$ Hz), beta ($16-32$ Hz), and gamma ($>32Hz$) and calculated power values of these frequency sub-bands (P_{freq}) using,

$$P_{freq} = \frac{1}{n}\sum_{j=1}^{n} p(S_j)^2, \tag{5}$$

where S_j represents the signal in frequency domain and power values were calculated for alpha, beta, and gamma sub-bands.

It has been shown that asymmetry ratios play a significant role in emotion classification, since these ratios embodies signal characteristics from both the left- and right-hemispheres of the brain [38]. In particular, some emotions are dedicated to the right hemisphere, while others to the left hemisphere of the brain [39]. We extracted rational and differential asymmetric features from the power of frequency bands. We used alpha, beta, and gamma bands for rational asymmetry (RASM) and differential asymmetry (DASM) features, which were calculated using,

$$RASM = \frac{P_{left}}{P_{right}}, \tag{6}$$

$$DASM = P_{left} - P_{right}. \tag{7}$$

where P_{left} and P_{right} represent the power (for alpha, beta, and gamma bands) from the electrodes on the left and right hemispheres of the brain (e.g., FP1 and FP2). We computed these values for channel pairs including FP1/FP2, AF3/AF4, F3/F4, F7/F8, FC5/FC6, FC1/FC2, C3/C4, T7/T8, CP5/CP6, CP1/CP2, P3/P4, P7/P8, PO3/PO4, and O1/O2. In feature fusion, RASM and DASM values for all these pairs were used as well as from the FP1/FP2 (included in the four electrodes selected) pair. We observed that emotion recognition accuracy was significant when only FP1/FP2 pair was used for RASM and DASM features. Hence, an agreement was found between selecting the electrode channel pair (FP1/FP2) for asymmetry features and electrodes selected using grid search.

3) WAVELET DOMAIN FEATURES

The wavelet features incorporate information from both time and frequency domain [40]. Discrete wavelet transform (DWT) [41] is used to decompose signals in different levels of decomposition. A signal is decomposed to an approximation coefficient (AC) and detail coefficients (DC). The mother wavelet is used for initial decomposition of signals, and AC is further decomposed to AC and DC [42], repeating the process to obtain the required level of decomposition.

Entropy and energy values were calculated as features using DWT from the alpha, beta, and gamma bands. The energy of a frequency band (E$_{freq}$) is given as,

$$E_{freq} = \sum_{j=1}^{n} p(S_j^w)^2, \tag{8}$$

where n is the total number of data samples in each band and S_j^W represents samples in the wavelet domain. We further calculated differential entropy for wavelet-based features using individual and various combinations of frequency sub-bands (results reported in Table 1).

E. FEATURE FUSION

In feature-level fusion, we concatenated features from multiple domains to obtain the final feature vector. We evaluated results obtained for individual features from each domain (Table 1) and hybrid features (Table 2). For possible amalgamations of features, we experimented with all features one by one in any domain and on the basis of results from each

TABLE 2. The average accuracy, precision, recall and F1 score for features in hybrid domain. E: entropy, H_j: Hjorth, W_e: Wavelet energy, W_{en}: Wavelet entropy, R: RASM, D: DASM, F_p: frequency power. The bold-face values represent the best values for each performance metric.

Features	10-fold Cross-Validation				LOOCV			
	Average Accuracy (%)	Precision	Recall	F1 Score	Average Accuracy (%)	Precision	Recall	F1 Score
Time (E, H_j)	64.94	0.60	0.65	0.57	64.95	0.60	0.65	**0.59**
H_j, W_{en}	**65.72**	**0.61**	**0.66**	0.58	**65.92**	**0.62**	**0.66**	**0.59**
H_j, W_e	64.60	0.58	0.65	0.56	64.60	0.58	0.65	0.56
H_j, W_{en}, W_e	65.55	0.60	**0.66**	0.58	65.71	0.61	**0.66**	0.58
H_j, R	64.43	0.59	0.60	0.55	64.42	0.58	0.64	0.56
H_j, D	64.72	0.59	0.65	0.56	64.82	0.60	0.65	0.56
H_j, R, D	64.43	0.59	0.64	0.56	64.57	0.60	0.65	0.56
H_j, F_p	64.39	0.56	0.64	0.55	63.78	0.57	0.64	0.53
$H_j, W_{en}, R(\alpha)$	65.37	0.58	0.65	0.56	65.74	0.58	**0.66**	0.57
$H_j, W_{en}, R(\beta)$	65.56	0.58	**0.66**	0.57	65.80	0.59	**0.66**	0.57
$H_j, W_{en}, R(\gamma)$	65.45	0.59	0.65	0.57	65.63	0.59	**0.66**	0.57
$H_j, W_{en}, R(\alpha, \beta)$	65.27	0.58	0.65	0.57	65.65	0.58	**0.66**	0.57
$H_j, W_{en}, R(\alpha, \gamma)$	65.33	0.58	0.65	0.56	65.39	0.58	0.65	0.57
$H_j, W_{en}, R(\beta, \gamma)$	65.40	0.58	0.65	0.57	65.70	0.59	**0.66**	0.57
$H_j, W_{en}, D(\alpha)$	65.70	**0.61**	**0.66**	0.57	65.79	0.60	**0.66**	0.57
$H_j, W_{en}, D(\beta)$	65.63	**0.61**	**0.66**	0.57	65.85	0.61	**0.66**	0.57
$H_j, W_{en}, D(\gamma)$	65.65	0.59	0.65	0.57	65.86	0.61	**0.66**	0.57
$H_j, W_{en}, D(\alpha, \beta)$	65.71	**0.61**	**0.66**	0.57	65.78	0.60	**0.66**	0.57
$H_j, W_{en}, D(\alpha, \gamma)$	65.67	**0.61**	**0.66**	0.57	65.82	0.60	**0.66**	0.57
$H_j, W_{en}, D(\beta, \gamma)$	65.61	**0.61**	**0.66**	0.57	65.78	0.60	**0.66**	0.57
E, W_{en}	65.46	**0.61**	0.65	**0.59**	65.59	**0.62**	**0.66**	**0.59**
E, W_e	64.04	0.58	0.64	0.58	64.04	0.59	0.64	0.58
E, W_{en}, W_e	65.36	0.60	0.65	**0.59**	65.52	0.61	**0.66**	**0.59**
E, R	64.23	0.58	0.64	0.58	64.00	0.58	0.64	0.57
E, D	63.63	0.57	0.64	0.56	63.57	0.57	0.64	0.58
E, R, D	63.94	0.58	0.64	0.56	63.47	0.57	0.63	0.56
E, F_p	62.90	0.56	0.63	0.56	62.47	0.56	0.62	0.56
H_j, E, W_{en}	64.83	0.58	0.65	0.58	65.46	0.60	0.65	0.58
H_j, E, W_e	64.58	0.59	0.65	0.58	64.41	0.58	0.64	0.57
H_j, E, W_e, W_{en}	65.18	**0.61**	0.65	0.58	65.40	0.60	0.65	0.58
$H_j, E, R(\alpha)$	64.42	0.56	0.63	0.57	64.58	0.56	0.65	0.57
$H_j, E, R(\beta)$	64.39	0.56	0.64	0.57	64.50	0.56	0.64	0.57
$H_j, E, R(\gamma)$	64.56	0.58	0.65	0.57	64.35	0.57	0.64	0.57
$H_j, E, D(\alpha)$	62.02	0.55	0.61	0.56	62.02	0.55	0.61	0.56
$H_j, E, D(\beta)$	62.45	0.55	0.63	0.57	62.45	0.55	0.63	0.57
$H_j, E, D(\gamma)$	62.61	0.56	0.63	0.56	62.61	0.56	0.63	0.56
$H_j, E, R(all)$	64.60	0.58	0.65	0.57	64.37	0.59	0.64	0.57
$H_j, E, D(all)$	64.56	0.58	0.64	0.57	64.68	0.58	0.65	0.57
H_j, E, R, D	64.48	0.58	0.64	0.57	64.13	0.59	0.64	0.56

individual domain, we concatenated features from multiple domains. We observed that in time domain, Hjorth parameters contributed the most, whereas entropy feature performed well in the wavelet domain. These selected features were further used for emotion classification.

F. EMOTION CLASSIFICATION

The multi-domain feature values were used for classification in different combinations- individually and in a hybrid setting. In our proposed model, SVM was used with 10-fold and leave-one out cross-validation technique for classification. Our technique was able to classify nine emotional states (happy, pleased, relaxed, excited, neutral, calm, distressed, miserable, depressed) which were mapped by dividing the valence arousal plane (Figure 2) in three levels i.e., low, medium, and high.

SUPPORT VECTOR MACHINE (SVM)

In our proposed model we used SVM as a classifier. We used Weka 3.8 tool, which provides a wide collection of machine learning classifiers, and SVM with sequential minimal optimization. SVM is a linear algorithm, which is based on statistical learning and creates a hyper-plane with in the data for class discrimination. For nonlinear problems, SVM tends to use different kernels, such as radial basis function and polynomial kernels which has been shown to improve the classification accuracy [36]. We used an SVM with polynomial kernel (degree = 3), with C equal to 1, an epsilon value of $1.0e-12$, random seed equal to 1, and tolerance parameter of 0.0001. The SVM was trained for each fold (with same parameters) and average results were reported from all folds.

IV. EXPERIMENTAL RESULTS

In our experiments, we used time, frequency, and wavelet domains for extracting features. Furthermore, classification was performed using features from individual and hybrid domains to obtain classification of nine emotions. The classification was performed using 10-fold cross validation, where each part of data was used for training and test purpose in an

TABLE 3. Comparison of EEG based emotion recognition models with our proposed model. CV: cross-validation, GELM: graph-regularized extreme machine learning, MLP: multi-layer perceptron, k-NN: k-nearest neighbour.

Method	Electrodes	Stimulus	Classifier	Emotions	Subjects	Accuracy(%)
[26]	32	Video	GELM	4	32	69.67
[27]	1	Audio	MLP	4	30	78.11
[43]	32	Video	Nearest Neighbour	4	32	73.62
[44]	32	Video	Domain-adaptation	5	14	39.05
[45]	14	Video	SVM	valence-dominance	10	63.04
[46]	32	Video	K-NN	2	30	69.50
Proposed (10-fold CV)	**4**	**Video**	**SVM**	**9**	**32**	**65.72**
Proposed (LOOCV)	**4**	**Video**	**SVM**	**9**	**32**	**65.92**

iterative manner. Moreover, LOOCV was also used to further verify the results and perform subject-independent analysis, where each subject was considered as test instance once while the remaining were considered for training. The results were reported using average values for all iterations and the model performance was evaluated using accuracy, precision, recall, and F1 measure.

A. FEATURES FROM INDIVIDUAL DOMAINS

The results obtained from individual features of time, frequency, and wavelet domains are presented here. In time domain, accuracy of nine emotion classification from entropy and Hjorth features was 63.62% and 64.74% respectively. The average accuracy for individual features using frequency sub-bands are given in Table 1. The results are obtained against the power feature from alpha, beta, and gamma bands. In addition, we also considered the features of RASM and DASM corresponding to these frequency bands. The wavelet energy and wavelet entropy features were extracted for these bands as well. The best classification accuracy of 65.01% (10-fold) and 65.19% (LOOCV) was achieved from entropy feature of wavelet domain. The performance improved to 65.14% (10-fold) and 65.23% (LOOCV) by combining entropy with wavelet energy.

We also explored the effect on classification accuracy using individual frequency sub-bands and combination of sub-bands (Table 1). The performance slightly improves when multiple bands were used. A general trend was observed where Hjorth in time domain and entropy in wavelet domain were the preferred features for emotion classification.

B. HYBRID FEATURES

Since the results improved when features were combined with in a domain. For completeness, we performed feature-level fusion and experimented with different combinations of multi-domain features. The idea behind this was to identify stable features which would best represent the underlying EEG signal characteristics and performed better classification. The results significantly changed with this feature-level fusion of multi-domain signal representation. The results for hybrid features are presented in Table 2, using the parameters including accuracy, precision, recall and F1 score. The best result were achieved against the Hjorth features (time domain) and entropy feature (wavelet domain).

In general, our experimental results reveal that individual features (wavelet entropy, wavelet energy and Hjorth) have significant contribution towards the emotion classification problem. In the hybrid domain, combining these individual features as well as the asymmetric features consistently give good classification accuracy. The best classification accuracy (65.92%) was achieved using Hjorth parameters and the wavelet entropy feature with leave-one-out cross-validation.

V. DISCUSSION

In our proposed method, we used DEAP dataset from all 32 subjects. A total of 40 electrodes were used (32 for EEG and 8 for peripheral) to collect data, among which we used data from four electrodes (EEG) in our experiments. These electrodes were chosen after grid search and are shown to be significant for emotion classification. We experimented with features from individual (Table 1) and multi-domains (Table 2) to find the most significant features and their combinations. The classification model achieved the highest average accuracy of 65.72% with 10-fold cross-validation and 65.92% with LOOCV (with Hjorth and wavelet entropy features), which is best for nine class emotion classification on DEAP dataset to the best of our knowledge. Hence, we show that using features from multiple domains in a hybrid way, instead of individually, benefited the classification problem. Moreover, subject-independent analysis was performed using leave-one-out cross-validation, where data from each subject was considered as test data while all other subjects were considered for training. This process was repeated for all subjects and average results were reported.

The results are compared (Table 3) with methods based on EEG signals for emotion classification. We used six different parameters including number of EEG electrodes, stimuli used to evoke emotions, machine learning classifier, number of emotion classes classified, number of subjects participated, and classification accuracy. It must be noted that our selected multi-domain features achieve the highest accuracy considering the number of emotions. A comprehensive analysis was performed to select stable patterns for emotion recognition using EEG data [26]. Information redundancy was shown to exist if all neural patterns and electrodes from the DEAP data were used. After selecting the stable patterns, an accuracy of 69.67 % was achieved for four valance/arousal states. In [27], audio music was used as a stimuli and 4 emotions were classified using multi-layer perceptron with an accuracy of 78.11%. Although, a higher classification accuracy was achieved, but the number of emotions considered was considerably low. Similarly, four emotion dimensions (valance, arousal, dominance, and liking) recorded in the DEAP data were classified using a neural network classifier [43]. A classification accuracy of 73.62% was achieved for the classification of four emotion dimensions. A domain adaptation technique was used to classify four emotion dimensions (valance, arousal, dominance, and liking) as well as familiarity for DEAP data with an accuracy of 39.05% [44]. An ontology based model using DEAP dataset, with selected 8 subjects, was developed [47]. The system identified two classes of valence and arousal and reached an accuracy of 75.19% and 81.74% for valance and arousal, respectively. A real-time model targeting the arousal-dominance using a subset of DEAP data was developed by selecting a single participant and the accuracy of the system was 63.04% [45]. Although some studies have reported higher accuracies but only when limited number of emotions were considered. Hence, we argue that selecting the appropriate features and electrodes is critical for developing an effective emotion recognition system for more emotion classes.

In general, we observed that our proposed method obtained nine classes of emotions while using data from all subjects with an average accuracy of 65.92%. In summary, we used four parameters (accuracy, precision, recall, and F1 measure) to evaluate our model. The best results were obtained with Hjorth parameters, wavelet entropy, and wavelet energy features in individual domains, and wavelet entropy and Hjorth parameters in multi-domain (hybrid). With extensive experiments, for our selected features (Hjorth from time domain and entropy feature from wavelet domain) the accuracy, precision, recall and F1 score values were 65.72%, 0.61, 0.66 and 0.58, respectively using 10-fold cross-validation. The accuracy improved to 65.92% when LOOCV was employed to perform subject-independent analysis with slight improvement in precision and F1 score. Moving forward the classification performance needs a further boost, and since we have established that hybrid features benefit the classification task, we intend to use deep learning in future. We also intend to combine data from various physiological modalities and curate data in un-controlled environments for mimicking real-time emotions. We believe that classifying a higher number of emotions would benefit real-world human-robot interactive systems in particular and affective computing in general.

VI. CONCLUSION

In this study, an emotion recognition model (for nine emotion states) based on EEG data is proposed. The innovative approach of finding multi-domain features from wavelet (entropy and energy), time (entropy and Hjorth) and frequency (RASM, DASM and power of frequency bands $(\alpha, \beta, and \gamma)$) domains was presented. The results of individual and hybrid features were analyzed and we found that entropy feature in the wavelet domain for three frequency sub-bands (α, β, γ) could out perform other features with the highest average accuracy of 65.19% (with LOOCV). We also experimented with time, wavelet, and frequency domain features and were able to select Hjorth parameters due to its high accuracy and F1 score. We also used multi-domain (hybrid) features: a combination of wavelet entropy and Hjorth parameters from time domain and the accuracy improved to 65.92% (with LOOCV). For affective computing, detecting a higher number of emotions to better understand how humans respond when dealing with machines is very important. In future we intend to use physiological data from sources other than EEG to further improve the emotion classification performance.

REFERENCES

[1] B. A. Erol, A. Majumdar, P. Benavidez, P. Rad, K.-K.-R. Choo, and M. Jamshidi, "Toward artificial emotional intelligence for cooperative social human–machine interaction," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 1, pp. 234–246, Feb. 2020.

[2] R. W. Picard and J. Klein, "Computers that recognise and respond to user emotion: Theoretical and practical implications," *Interacting Comput.*, vol. 14, no. 2, pp. 141–169, Feb. 2002.

[3] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

[4] S. Nirenberg. (2016). *What if Robots Could Process Visual Information the Way Humans do?* [Online]. Available: https://www.tedmed.com/talks/show?id=619685

[5] S. Siddharth, T.-P. Jung, and T. J. Sejnowski, "Utilizing deep learning towards multi-modal bio-sensing and vision-based affective computing," *IEEE Trans. Affect. Comput.*, early access, May 14, 2019, doi: 10.1109/TAFFC.2019.2916015.

[6] Y.-J. Liu, M. Yu, G. Zhao, J. Song, Y. Ge, and Y. Shi, "Real-time movie-induced discrete emotion recognition from EEG signals," *IEEE Trans. Affect. Comput.*, vol. 9, no. 4, pp. 550–562, Oct. 2018.

[7] T. Kalsum, S. M. Anwar, M. Majid, B. Khan, and S. M. Ali, "Emotion recognition from facial expressions using hybrid feature descriptors," *IET Image Process.*, vol. 12, no. 6, pp. 1004–1012, Jun. 2018.

[8] X.-W. Wang, D. Nie, and B.-L. Lu, "Emotional state classification from EEG data using machine learning approach," *Neurocomputing*, vol. 129, pp. 94–106, Apr. 2014.

[9] E. L. van den Broek, "Ubiquitous emotion-aware computing," *Pers. Ubiquitous Comput.*, vol. 17, pp. 1–15, Jan. 2012.

[10] A. Raheel, S. M. Anwar, and M. Majid, "Emotion recognition in response to traditional and tactile enhanced multimedia using electroencephalography," *Multimedia Tools Appl.*, vol. 78, no. 10, pp. 13971–13985, 2019.

[11] H. Qayyum, M. Majid, E. U. Haq, and S. M. Anwar, "Generation of personalized video summaries by detecting viewer's emotion using electroencephalography," *J. Vis. Commun. Image Represent.*, vol. 65, Dec. 2019, Art. no. 102672.

[12] A. Mehreen, S. M. Anwar, M. Haseeb, M. Majid, and M. O. Ullah, "A hybrid scheme for drowsiness detection using wearable sensors," *IEEE Sensors J.*, vol. 19, no. 13, pp. 5119–5126, Jul. 2019.

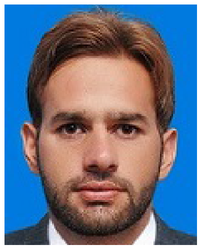
[13] A. Raheel, M. Majid, S. M. Anwar, and U. Bagci, "Emotion classification in response to tactile enhanced multimedia using frequency domain features of brain signals," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 1201–1204.

[14] H. J. Yoon and S. Y. Chung, "EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2230–2237, Dec. 2013.

[15] P. Ackermann, C. Kohlschein, J. A. Bitsch, K. Wehrle, and S. Jeschke, "EEG-based automatic emotion recognition: Feature extraction, selection and classification methods," in *Proc. IEEE 18th Int. Conf. E-Health Netw., Appl. Services (Healthcom)*, Sep. 2016, pp. 1–6.

[16] M. L. R. Menezes, A. Samara, L. Galway, A. Sant'Anna, A. Verikas, F. Alonso-Fernandez, H. Wang, and R. Bond, "Towards emotion recognition for virtual environments: An evaluation of EEG features on benchmark dataset," *Pers. Ubiquitous Comput.*, vol. 21, no. 6, pp. 1003–1013, Dec. 2017.

[17] A. J. Tomarken, R. J. Davidson, R. E. Wheeler, and L. Kinney, "Psychometric properties of resting anterior EEG asymmetry: Temporal stability and internal consistency," *Psychophysiology*, vol. 29, no. 5, pp. 576–592, Sep. 1992.
[18] B. Nakisa, M. N. Rastgoo, D. Tjondronegoro, and V. Chandran, "Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors," *Expert Syst. Appl.*, vol. 93, pp. 143–155, Mar. 2018.
[19] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012.
[20] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using EEG signals: A survey," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 374–393, Jul. 2019.
[21] L. Romeo, A. Cavallo, L. Pepa, N. Berthouze, and M. Pontil, "Multiple instance learning for emotion recognition using physiological signals," *IEEE Trans. Affect. Comput.*, early access, Nov. 19, 2019, doi: 10.1109/TAFFC.2019.2954118.
[22] A. Albraikan, D. P. Tobon, and A. El Saddik, "Toward user-independent emotion recognition using physiological signals," *IEEE Sensors J.*, vol. 19, no. 19, pp. 8402–8412, Oct. 2019.
[23] J. Teo, C. L. Hou, and J. Mountstephens, "Deep learning for EEG-based preference classification," in *Proc. AIP Conf.*, 2017, vol. 1891, no. 1, Art. no. 020141.
[24] S. Pouyanfar and S.-C. Chen, "Automatic video event detection for imbalance data using enhanced ensemble deep learning," *Int. J. Semantic Comput.*, vol. 11, no. 1, pp. 85–109, Mar. 2017.
[25] W. Szwoch, "Emotion recognition using physiological signals," in *Proc. Mulitimedia, Interact., Design Innnovation*, 2015, pp. 15:1–15:8.
[26] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, "Identifying stable patterns over time for emotion recognition from EEG," *IEEE Trans. Affect. Comput.*, vol. 10, no. 3, pp. 417–429, Jul. 2019.
[27] A. M. Bhatti, M. Majid, S. M. Anwar, and B. Khan, "Human emotion recognition and analysis in response to audio music using brain signals," *Comput. Hum. Behav.*, vol. 65, pp. 267–275, Dec. 2016.
[28] S. Tripathi, S. Acharya, R. D. Sharma, S. Mittal, and S. Bhattacharya, "Using deep and convolutional neural networks for accurate emotion classification on deap dataset," in *Proc. 21st AAAI Conf. Artif. Intell.*, 2017, pp. 4746–4752.
[29] K. Guo, H. Candra, H. Yu, H. Li, H. T. Nguyen, and S. W. Su, "EEG-based emotion classification using innovative features and combined SVM and HMM classifier," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 489–492.
[30] L. K. McEvoy, M. E. Smith, and A. Gevins, "Test–retest reliability of cognitive EEG," *Clin. Neurophysiol.*, vol. 111, no. 3, pp. 457–463, Mar. 2000.
[31] Z. Wang, S. Hu, and H. Song, "Channel selection method for EEG emotion recognition using normalized mutual information," *IEEE Access*, vol. 7, pp. 143303–143311, 2019.
[32] C. Niemic, "A theoretical and empirical review of psychophysiological studies of emotion," *J. Undergraduate Res., Univ. Rochester*, vol. 1, pp. 15–18, 2002.
[33] R. J. Davidson, D. C. Jackson, and N. H. Kalin, "Emotion, plasticity, context, and regulation: Perspectives from affective neuroscience," *Psychol. Bull.*, vol. 126, no. 6, pp. 890–909, Nov. 2000.
[34] M. S. Özerdem and H. Polat, "Emotion recognition based on EEG features in movie clips with channel selection," *Brain Informat.*, vol. 4, no. 4, pp. 241–252, Dec. 2017.
[35] A. Mert and A. Akan, "Emotion recognition based on time–frequency distribution of EEG signals using multivariate synchrosqueezing transform," *Digit. Signal Process.*, vol. 81, pp. 106–115, Oct. 2018.
[36] J. Atkinson and D. Campos, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," *Expert Syst. Appl.*, vol. 47, pp. 35–41, Apr. 2016.
[37] J. Chen, B. Hu, P. Moore, X. Zhang, and X. Ma, "Electroencephalogram-based emotion assessment system using ontology and data mining techniques," *Appl. Soft Comput.*, vol. 30, pp. 663–674, May 2015.
[38] S. Kameyama, H. Shirozu, and H. Masuda, "Asymmetric gelastic seizure as a lateralizing sign in patients with hypothalamic Hamartoma," *Epilepsy Behav.*, vol. 94, pp. 35–40, May 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1525505019301593
[39] M. K. Mukul and F. Matsuno, "Feature extraction from subband brain signals and its classification," *SICE J. Control, Meas., Syst. Integr.*, vol. 4, no. 5, pp. 332–340, 2011.
[40] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, D. Hazry, and I. Zunaidi, "Time-frequency analysis of EEG signals for human emotion detection," in *Proc. 4th Kuala Lumpur Int. Conf. Biomed. Eng.* Berlin, Germany: Springer, 2008, pp. 262–265.
[41] M. Swangnetr and D. B. Kaber, "Emotional state classification in patient–robot interaction using wavelet analysis and statistics-based feature selection," *IEEE Trans. Human-Machine Syst.*, vol. 43, no. 1, pp. 63–75, Jan. 2013.
[42] M. Murugappan, N. Ramachandran, and Y. Sazali, "Classification of human emotion from EEG using discrete wavelet transform," *J. Biomed. Sci. Eng.*, vol. 3, no. 4, pp. 390–396, 2010.
[43] V. Rozgic, S. N. Vitaladevuni, and R. Prasad, "Robust EEG emotion classification using segment level decision fusion," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 1286–1290.
[44] Z. Lan, O. Sourina, L. Wang, R. Scherer, and G. R. Muller-Putz, "Domain adaptation techniques for EEG-based emotion recognition: A comparative study on two public datasets," *IEEE Trans. Cognit. Develop. Syst.*, vol. 11, no. 1, pp. 85–94, Mar. 2019.
[45] C. Mühl, B. Allison, A. Nijholt, and G. Chanel, "A survey of affective brain computer interfaces: Principles, state-of-the-art, and challenges," *Brain-Comput. Interface*, vol. 1, no. 2, pp. 66–84, Apr. 2014.
[46] H. Bhavsar and M. H. Panchal, "A review on support vector machine for data classification," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 1, no. 10, pp. 1–5, 2012.
[47] A. Etkin, C. Büchel, and J. J. Gross, "The neural bases of emotion regulation," *Nature Rev. Neurosci.*, vol. 16, no. 11, pp. 693–700, Nov. 2015.



MUHAMMAD KHATEEB was born in Pakistan, in 1994. He received the B.S. degree in software engineering from the Mirpur University of Science and Technology (MUST), and the M.S. degree in software engineering from the University of Engineering and Technology (UET), Taxila. From 2017 to 2019, he was a Research Scholar with the Department of Software Engineering, UET. He is currently working as a Lab Demonstrator with Riphah International University, Islamabad, Pakistan. His research interests include emotion recognition, signal processing, and machine learning.

SYED MUHAMMAD ANWAR received the B.Sc. degree (Hons.) in computer engineering from the University of Engineering and Technology (UET), Taxila, Pakistan, in 2005, and the M.Sc. degree (Hons.) in data communications and the Ph.D. degree in electronic and electrical engineering from The University of Sheffield, U.K., in 2007 and 2012, respectively. He is currently an Associate Professor with the Department of Software Engineering, UET. His research interests include medical imaging, data communication, and human–computer interaction.

MAJDI ALNOWAMI (Member, IEEE) received the B.Sc. degree in electronic engineering from King Abdulaziz University, in 2002, and the M.Sc. degree in medical physics and the Ph.D. degree from the University of Surrey, in 2008 and 2012, respectively. From 2002 to 2008, he was a Project Manager with Siemens. He is currently an Assistant Professor with the Department of Nuclear Engineering, King Abdulaziz University. His research interests include bridge-the-gap between real and computer generated imagery, medical imaging, tracking and radiotherapy to investigate new methods for tracking, and modeling and understanding of the effect of inter- and intra-fraction motion during radiotherapy and applications, include 4D treatment method, tracking internal body motion, radiotherapy dose escalation, and segmentation and respiratory motion compensation.