IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 24, 2020, accepted February 6, 2021, date of publication February 10, 2021, date of current version February 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3058443

Emotional State Recognition Performance Improvement on a Handwriting and Drawing Task

JUAN A. NOLAZCO-FLORES[1], (Senior Member, IEEE), MARCOS FAUNDEZ-ZANUY[2], OLIVER A. VELÁZQUEZ-FLORES[1], GENNARO CORDASCO[3], AND ANNA ESPOSITO[3]
[1]School of Engineering and Science, Tecnologico de Monterrey, Monterrey 64849, Mexico
[2]Escola Superior Politecnica, TecnoCampus Mataro-Maresme, 08302 Mataro, Spain
[3]IIASS, Università della Campania Luigi Vanvitelli, 81100 Caserta, Italy

Corresponding author: Oliver A. Velázquez-Flores (a01380501@itesm.mx)

This work was supported in part by the Consejo Nacional de Ciencia y Tecnología (CONACyT).

ABSTRACT In this work we combine time, spectral and cepstral features of the signal captured in a tablet to characterize depression, anxiety, and stress emotional state recognition on the EMOTHAW database. EMOTHAW contains the emotional states of users represented by capturing signals from sensors on the tablet and pen when the user is performing 3 specific handwriting and 4 drawing tasks, which had been categorized into depressed, anxious, stressed, and typical, according to the Depression, Anxiety and Stress Scale (DASS). Each user was characterized with six time-domain features, and the number of spectral-domain and cepstral-domain features for the horizontal and vertical displacement of the pen, the pressure on the paper, and the time spent on-air and off-air, depended on the configuration of the filterbank. As next step, we select the best features using the Fast Correlation-Based Filtering method. Since our dataset has 129 users, then as next step, we augmented the training data by randomly selecting a percentage of the training data and adding a small random Gaussian noise to the extracted features. We then train a radial basis SVM model using the Leave-One-Out (LOO) methodology. The experimental results show an average accuracy classification improvement ranging of 15%, and an accuracy classification improvement ranging from 4% to 34% compared with baseline (state of the art) for specific emotions such as depression, anxiety, stress, and typical emotional states.

INDEX TERMS Data augmentation, emotional state recognition, emotional states, feature extraction, SVM.

I. INTRODUCTION

Biometrics can be used on e-Security and e-Health [1]. Morphological biometrics, such as fingerprint or eyes pupils, are based on direct measurements of physical traits of the human body [2], [3]. Behavioral biometrics, such as handwriting and drawing, are based on specific drawing and handwriting task performed by the subjects involved in the data collection t [4]. From a health condition perspective, online handwriting biometrics are more appealing and informative on indicating states of mental disorders and diseases, such as dementia, than other popular biometrics traits such as fingerprints or iris [3], [4] because they are required in routinely functional activities.

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

Mental illnesses are one of the major causes of disabilities among the worldwide population and can be considered a public health priority. Depression, stress, and anxiety are among the most prevalent worldwide, and stress is almost always present as comorbidity. The number of people living with depressive disorders and anxiety are estimated to be the 4.4% and 3,6% of the global population, respectively, [5]. The manifestation of such disorders is commonly accompanied, and occasionally preceded, by the deterioration of social behavior mainly exemplified by the inability to code and decode oneself and other emotions [6] and loss of motivations such apathy, reduced feelings, lack of self-esteem, reduced social contacts. There are no established cures for depression and anxiety and treatments can last for the entire life with a consistent impoverishment of patients' quality of life and substantial increase of public cares' costs. The research for detecting and diagnosing these mental disorders is driven by the need of identifying early indicators able to predict early signs of their development to provide tempestive cures before they become chronic and hard to be treated. Among these indicators there are changes in physical biomarkers such as voice, facial expressions, and body postures, changes in behaviors (psychometrics), as well as changes in functional abilities such as handwriting and drawing [7]. Identifying early indicators is especially important for implementing early interventions, and reduce multiple in-office clinical visits, and the costs of public cares [8]. Depression, stress, and anxiety are distinguished by emotional states because they are lasting for long time periods and are considered a negative state of mind (mood) worsening the quality of life of patients.

Depression is an emotional state that causes an irregular loss of interest in activities enjoyed by people with this disease. Energy and willingness tend to disappear, and sadness is a constant emotion, which causes the brain to consume more energy than intended. If depression is not cured, in the worst cases may end up in suicide [9].

Anxiety is, according to Schlenker and Leary [10]: ''A cognitive-affective response characterized by physiological arousal (nervous system activation) and apprehension regarding a potentially negative outcome that the individual perceives as impending''. Symptoms presented are sweating, palpitations, tachycardia, rapid breathing, among others [9]. This state causes tiredness and impact the working activities of patients [11].

Stress is a natural self-defense process in which the body of a living creature undergoes when faced with a complicated or dangerous situation [12]. Stress becomes a mental disorder when it is experienced frequently since it involves the releasing of a large amount of several hormones which in turn may overflow the body and impact the functionality of organs and cognitive processes alike, causing diabetes or cardiovascular diseases [13]. In such cases behavioral biometrics, such as writing and drawing, determine the emotional state, such as depression, anxiety, or stress [14].

The incremental usage of social media and the constant text interaction between users suggest an increasing information pool, where important data could be gathered for future analysis.

The field of emotion and mood detection and classification through handwriting and drawing (and its applications) is new, and literature is in its initial stages. There are not too many literature studies nor public data bases to exploit as benchmark. The current paper is taking upon a study published in 2017 by Likforman-Sulem, *et al.* [15] which proposed a first attempt to exploit is handwriting and drawing features for discriminating among depressed, stressed, and anxious patients from healthy control subjects. By using time features such as the time the pen is kept on the pad, on the air, and total time to complete the assigned task, as well other features such as the number of strokes and pressure of the pen on the tablet, and a random forest classification algorithm, the authors were able to obtain measure of accuracy, sensitivity, and specificity for depression, anxiety, and stress, well above the chances.

In the current paper, we extended the research of additional handwriting and drawing features, adding complex transformations of the initial measurements such as the Log Energy Spectrum (E), the Log Energy of a linear Filterbank (LEFB) and Filterbank Cepstral Coefficients (FBCCs).

Section II reviews the SVM data modeling concepts. Section III describes the features of the EMOTHAW database. Section IV describes the time-domain features obtained directly from an INTUOS WACOM series 4 tablet. Section V describes the Feature Extraction process used in this paper and describes the type of features obtained. Section VI presents a brief explanation of the Fast Correlation-Based Filtering feature selection methodology [16] and a modified version that yielded better results in experiments. Section VII describes the experiments conducted and their results. Finally, in Section VIII remarks some comments and conclusions.

II. SUPPORT VECTOR MACHINE

SVM is a widely known machine learning and statistical model that can be used as a classifier or regressor [17]. It is based on the concept of statistical learning theory and capitalizes on the creation of a feature space with respect to maximizing margin distance (creating an optimal hyperspace), in other words, given a n-feature database and an instance-label pair of the form: (Xu, yu), where u =1,2, 3,..., l, Xu ∈ Rn, and yu ∈ {-1, 1} l and can be linear separated if there are a vector w and a scalar b that fulfill the following inequalities:

$$w \cdot X_u + b \leq -1 \; when \; y_i = -1$$

Both equations can be rewritten by multiplying the right-hand side by yu:

$$y_i \left(w \cdot X_u + b\right) \geq 1, i = 1, 2, 3, \ldots, l$$

When these conditions are met, the optimal hyperspace is found when the equation equals 0, for w0 and b0, both stand for the weight vector and the bias (scalar), respectively. To maximize the boundary in-between space the vector weight's module must be minimize:

$$min\left(\frac{1}{2}||\mathbf{w}||^2\right)$$

Now, it becomes an optimization problem, therefore we use Lagrange multipliers ai ≥ 0 for each Xi - yi pair, which now we have a Lagrange function of the form:

$$\mathrm{L}\left(\mathbf{w}, b, \mathbf{a}\right) = \frac{1}{2}||\mathbf{w}||^2 - \sum_{\mathrm{u}=1}^{1} a_u\{y_u\left(\mathbf{w}\cdot\mathbf{X_u} + b\right) - 1\}$$

Furthermore, a kernel function is defined as:

$$k\left(\mathbf{X_u}, \mathbf{X_u^{'}}\right) = \phi\left(\mathbf{X_u}\right)^{\top}\phi\left(\mathbf{X_u^{'}}\right)$$

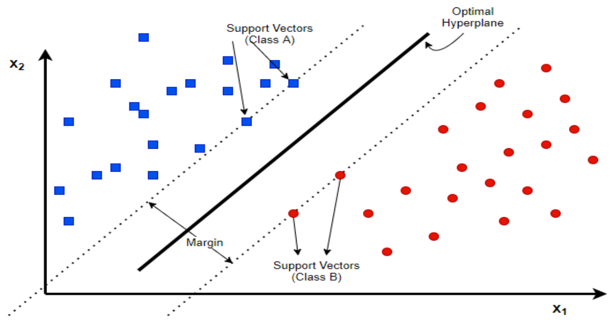


FIGURE 1. Representation of the support vectors for classes A and B and the created margins in-between for the optimal hyperplane.

where $\mathbf{X_u}$, $\mathbf{X_u^{'}}$ are samples in the input space, and $\phi(X)$ is a feature-space transformation function, with both equations we can define the classification function as:

$$f\left(\mathbf{X}\right)=\sum_{i=1}^{l}a_i y_i k\left(\mathbf{X_u},\mathbf{X_u^{'}}\right)+b$$

There are different kernels (so-called kernel trick) that SVM can work with: linear, polynomial, Radial Basis Function (RBF), and sigmoid. For this article, we worked with the RBF kernel, which is used when databases cannot be linearly separated. which is defined by the equation:

$$k\left(\mathbf{X_i},\mathbf{X_j}\right)=exp\left(-\gamma||\mathbf{X_i}-\mathbf{X_j}||^2\right),\gamma>0$$

where γ is a regularization hyper-parameter.

III. EMOTHAW DATABASE

In the mentioned article, the importance of early emotional state recognition and the implication of it in healthcare were mentioned. Thanks to the well assessed testing available in the literature, anxiety, stress, and depression were the emotional states to study. The DASS Scale (Depression-Anxiety-Stress Scale) [18] provides a reliable questionnaire to relate to handwriting tasks. They created a database called EMOTHAW (EMOTion recognition from HAndWriting and drAWing) from seven handwritten tasks picked from cognitive and psychological tests that measure mental state, cognitive capabilities, and even brain damage: CDT (Clock Drawing Test), MMSE (Mini-mental State Examination), and HTP (House-Tree-Person test) [19], [20]. They added two writing tasks to understand how impactful different tasks are in emotion recognition.

A. THE DASS SCALE

The Depression Anxiety and Stress Scale was first proposed by [18] and the Italian version consisting of 42 questions (I-DASS-42) was assessed by Severino [21]. The scale measures whether the subject is suffering of depression, anxiety, and/or stress through an emotional score scale as reported in Table 1, through the DASS it was possible to establish a bridge between tasks and emotions, since the scores received by each subject were indicating whether her/his stress, depression, anxiety were at a normal, mild, moderate, severe or extremely severe degree. Range values are shown in Table 1 for the three emotional states. The database only has information of 129 subjects, so the classification in Table 1 will generate outliers for most of the labels in it. What they did was to cluster the score into a binary classification according to the following procedure. A subject was considered:

a) Depressed when her scores to DASS for the questions associated with the detection of depression were higher than 9;
b) Anxious when her scores to the questions associated with anxiety were higher than 7;
c) Stressed when her scores to the questions associated with stress were higher than 14.

In doing so, they created for each emotional state a coarse binary classification of it, according to the values reported in Table 1, neglecting the different degree of severity of the disorder. A subject therefore can be only classified as depressed or not, stressed or not, anxious, or not.

This database has a binary-label output. Table 2 shows the threshold for the Normal and Above Normal range in the DASS Score range.

TABLE 1. Dass score range (taken from Likforman-Sulem, *et al.*[15]).

	Depression	Anxiety	Stress
Normal	0-9	0-7	0-14
Mild	10-13	8-9	15-18
Moderate	14-20	10-14	19-25
Severe	21-27	15-19	26-33
Extremely Severe	28+	20+	34+

TABLE 2. Binary labeling.

	Depression	Anxiety	Stress
Normal (0)	0-9	0-7	0-14
Above Normal (1)	10-28+	8-20+	15-34+

B. SUBJECTS

In total, 129 participants with similar demographic information to avoid differences in handwriting features according to their article [15]. The subjects were all master's and B.S. students of an age range of 21-32 years from the Università della Campania "Luigi Vanvitelli" in Italy. They did the DASS test first and then the seven-task test to measure the same emotional state on them.

C. TASKS

The test consists of seven tasks picked from different cognitive tests as mentioned in Section III. The tasks are as follows: two-pentagon drawing, house drawing, 4-word writing, left-hand Circle drawing, right-hand Circle drawing, clock drawing, and sentence writing.

For the data acquisition, they tested the difference between online and offline data collection and concluded that collecting online data had more impact for the research because information like, pressure, pen position, and time stamp could be obtained this way, and the relevance of these characteristics were proven in their article. They used an INTUOS WACOM series 4 and an Intuos Inkpen for data collection. These files include: X and Y position of the drawing in the tablet, time stamp, pen status (if it is on paper (1) or in air (0)) altitude and azimuth pen angles, and pressure, which are consider the raw characteristics.

In their analysis, they extracted four time-features: Time in air during task, time on paper during task, total duration of the task, and number of strokes, then they selected five tasks and extracted the mentioned features, having in total twenty new features per emotional state. They wanted to see the contribution of writing-based tasks (two tasks = eight features), drawing-based tasks (three tasks = twelve features), and the combination of both (20 features), so in the experimentation phase, they used a Random Forest model for emotional state recognition with a Leave-One-Out approach and repeating the experiments ten times for each type of tasks.

The results show that writing-based tasks have worse performance than the others in all three emotional states, especially in stress. For depression, drawing-based tasks have the most impact of the three, and the combination of tasks perform better on stress and anxiety.

IV. TIME SIGNALS CAPTURED BY THE TABLET

Fig. 2 shows the signals captured in real time by the software, when the user writes or draws on the tablet. These signals are position, or horizonal displacement, of pen tip in X-axis, $\boldsymbol{x}(\boldsymbol{n})$; position, vertical displacement, of pen tip in Y-axis, $\boldsymbol{y}(\boldsymbol{n})$; on-surface/in-air pen position information or status (touch/no-touch tablet's surface), $\boldsymbol{sq}(\boldsymbol{n})$; altitude of the pen with respect to the tablet's surface, $\boldsymbol{al}(\boldsymbol{n})$; pressure applied by the pen tip, $\boldsymbol{p}(\boldsymbol{n})$; azimuth angle of the pen with respect to the tablet's surface, $\boldsymbol{az}(\boldsymbol{n})$; timestamp, $\boldsymbol{Ts}$.

V. FEATURE EXTRACTION

Given the time signals captured by the tablet, we calculate the time-domain, spectral-domain, and cepstral-domain features.

A. TIME-DOMAIN FEATURES

Time-domain features are the features that can be obtained directly from the discrete time signal. The time features, for class C, used in this work are

$$TDF^C = \left\{F_1^C, F_2^C, T^C, NSt^C, \overline{P}^C\right\},$$

where, $T^C = F_1^C + F_2^C$ is the duration to complete task of class C, $F_1^C = \sum_{h=1}^{H-1} d_{2h}$ is the on-air pen duration

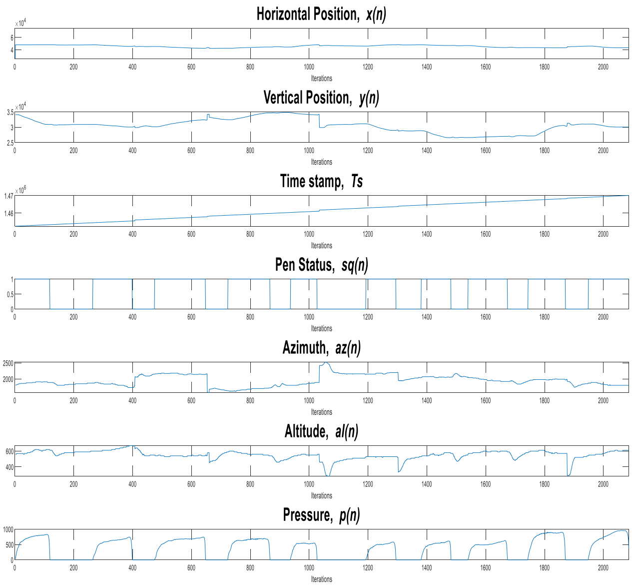


FIGURE 2. First task plot (pentagon drawings).

for task of class C, $F_2^C = \sum_{h=0}^{H} d_{2h+1}$ is the on-paper pen duration for task of class C, H^C is the total number of strokes in class C in $A^C = \{St^{1,C}, St^{2,C}, \ldots, St^{H,C}\}$, $St^{i,C}$, is the i-th stroke. On a given set of strokes, pen alternates contact with air and paper; by definition, $St^{1,C}$ and $St^{H,C}$ are the first and last strokes on paper, respectively, d_i is the duration of stroke i; note that d_i when i%2 == 0 is on-air, otherwise on-paper, t_i is the duration of the stroke i, NSt^C = is the number of on-paper strokes, this corresponds to $length\left(St_O^C\right)$, where $St_O^C = \{St^{1,C}, St^{3,C}, \ldots, St^{H,C}, \overline{P}^C\}$ is the non-zero value mean of the signal, and $C = \{pentagons, house, clock, words, sentence\}$.

Then, the time-domain feature set is the union of the features of each class, and it is expressed as follows:

$$TDF = \bigcup_{c \in C} TDF^c$$

B. SPECTRAL-DOMAIN FEATURES

The frequency features are the ones we obtain from the spectrum, $S(k)$, of the signal, $s(n)$. The Spectral feature vector for each signal, $s(n)$, and class C, is as follows:

$$SDF_\Psi^C = LEFB_\Psi^C(m),$$

where

$$LEFB_\Psi^C(m) = \underset{\theta=1,2,\ldots,M}{filterbank}\{E_\Psi(k)\}, for\ m = 1, \ldots, M,$$

$E_\Psi^C(k) = log_2\left(\left|S_\Psi^C(k)\right|^2\right)$ is the log energy spectrum, $S_\Psi^C(k) = \sum_{n=0}^{N-1} s_\Psi^C(n)\, e^{-\frac{2\pi i}{N}kn}, for\ k = 0, \ldots K$ is the Discrete Fourier Transform (DFT) and $s_\Psi^C(n) = \{x^C(n), y^C(n), p^C(n)\}$.

Then, the spectral domain feature set is the union of the features of each of the signal used, $s_\Psi^C(n)$, for all class C, and

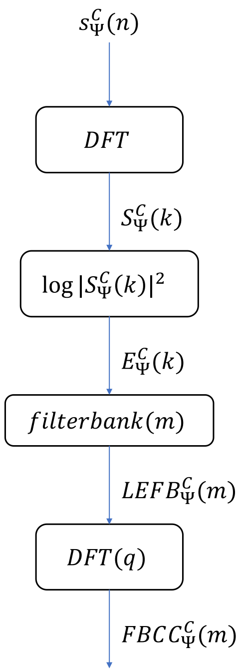


FIGURE 3. **This figure shows the process in which the input signal ($s_\Psi^C(n)$) is transformed. It undergoes first a DFT process, to obtain $S_\Psi^C(k)$, then a logarithmic spectrum calculation to obtain the energy ($E_\Psi^C(k)$), after that, we set a linear filterbank for all energies, resulting in logEnergy filterbank ($LEFB(m)$), and at the end we apply DFT again to calculate the cepstrum of the logEnergy filterbank ($FBCC(q)$).**

it is expressed as follows:

$$SDF = \bigcup_{c \in C} \bigcup_{\psi \in \Psi} SDF_\psi^c$$

C. CEPSTRAL-DOMAIN FEATURES

The cepstral domain features are the ones obtained directly from the cepstrum. In this work, we directly use the Cepstral coefficients, using the DFT of the LEFB, as follows:

$$FBCC_\Psi^C(q) = \sum_{m=0}^{M-1} LEFB_\Psi^C(m)\, e^{-\frac{2\pi i}{N} qm}, \text{ for } q = 1, \ldots, Q$$

where, M is the number of Filterbanks, $Q = M/2$ is the length of the signal divided by 2.

Then, the cepstral features set is the union of the features of each of the signal used $s_\Psi^C(n)$, for all class C, and it is expressed as follows:

$$CDF = \bigcup_{c \in C} \bigcup_{\psi \in \Psi} FBCC_\psi^c(q)$$

D. USERS FEATURE VECTOR

The feature vector of each user is the union of the Time, Spectral and Cepstral features sets, represented in the following formula

$$\mathrm{FV}_u = TDF_u \cup \mathrm{SDF}_u \cup CDF_u.$$

Now, let us define the observation for each emotional state, as the set of users with the same emotional state, as follows:

$$O_E[u, FV_u] = \bigsqcup_{u=1}^{U} FV_u^E$$

where $\sqcup$ is the append operator for the feature vector of all users, and $E = \{depression, anxiety, stress\}$. Observe that, $O^E[u, |FV_u|]$ is a 2-D array, where 1 dimension is the number of users, and the other is the number of features.

For example, if Q= 15 (linear filters in the filterbank) for 30Hz, M= 8, the number of spectral and cepstral features for each user is 23 per signal. If we use three signals, say $x(n)$, $y(n)$ and $p(n)$, to obtain spectral and cepstral features, then the number of features per user is 69. Adding the five time-domain features per class, C, we have 74 features. Since the number of classes are 5, then the number of the features per user is 370 per subject, 148 of them are from the writing-tasks features and 222 are from the drawing-tasks feature.

VI. FEATURE SELECTION

Feature selection is a popular and common pre-modeling step in machine learning, especially in high-dimensional databases. Each of our databases present a considerable number of features, so in this work we reduce the dimension of the number of features using a modified version of the Fast Correlation-Based Filtering (FCBF) [16].

FCBF is based on two correlation factors: correlation between each feature and output, and correlation among features. As first step, the selected features are the one with the highest correlation, higher than a threshold, with the output. As second step, these selected features are correlated, and the final set of selected features are the ones with lowest correlation, lower than a threshold, between them. The original FCBF, in the second step, when each correlation of each pair of features is compared to select the features, the algorithm selects any of them. In our modified version, mFCBF, we select the one with higher correlation with the output.

Algorithm 1 shows a pseudo code for the mFCBF process. The mFCBF algorithm allows to set the minimum correlation threshold (oTh) and the maximum correlation threshold between features (iTh). Then the selected features set, given oTh and iTh is defined as follows

$$\hat{O}^E_{oTh,iTh} = \mathrm{mFCBF}_{oTh,iTh}\left(O^E\right). \tag{1}$$

Observe that in $\hat{O}^E_{oTh,iTh}$ is a 2-D array, where 1 dimension is the number of users, and the other is the number of selected features.

Feature selectivity is controlled with oTh and iTh values. For example, given 370 user feature vectors f, and by using iTh= 0.15 and oTh= 0.7 the number of selected features is reduced to 26, or the depression state; reduced to 28 for anxiety; reduced to 20 for stress.

Algorithm 1 The mFCBF algorithm receive the Users Feature Matrix ($\boldsymbol{O^E}$), minimum correlation threshold ($\boldsymbol{oTh}$) and the maximum correlation threshold ($\boldsymbol{iTh}$) and returns the selected set of features.

function mFCBF($\boldsymbol{O^E}$, $\boldsymbol{oTh}$, $\boldsymbol{iTh}$)
    Calculate corr($\boldsymbol{O^E}$)
    $\boldsymbol{O_{tmp}} \leftarrow$ Select columns whose correlation with the output is $> \boldsymbol{oTh}$
    Calculate corr($\boldsymbol{O_{tmp}}$)
    $\hat{O}^E_{\boldsymbol{oTh},\boldsymbol{iTh}} \leftarrow$ Select columns whose correlation with the input is $< \boldsymbol{iTh}$ and with the highest correlation with the output.
    return($\hat{O}^E_{\boldsymbol{oTh},\boldsymbol{iTh}}$)
end function

TABLE 3. The eight highest ranked features selected by modified-FCBF for the depression-writing database.

Highest Ranked Features
• Number of Strokes – Pentagon Drawings
• Cepstral Coefficient #6 - Horizontal Position – Pentagon Drawings
• Cepstral Coefficient #8 - Horizontal Position – Pentagon Drawings
• Cepstral Coefficient #3 – Pressure – Pentagon Drawings
• Cepstral Coefficient #6 - Pressure – Pentagon Drawings
• Cepstral Coefficient #3 – Horizontal Position – House Drawing
• FilterEnergy #8 – Pressure – House Drawing
• Cepstral Coefficient #4 - Pressure – House Drawing

Table 3 shows the eight highest ranked features in the depression-writing observations after processing them using mFCBF. We can observe, the features with higher importance are, mostly, Cepstral Coefficients in both Pressure and Horizontal position. Nonetheless, this selection is bounded to the database type, number of total features in it, correlational thresholds (mFCBF methodology), and the parameters listed in all four batches of experimentation.

VII. RESULTS

Leave-One-Out (LOO) was used for testing. In LOO the data model is trained with all database registers but one, and we test with the register that was out. We repeated this testing, such that we leave each of the user out one time, and we averaged the accuracy of all tests.

Since we introduced different concepts to test, the purpose of the different experiment was to show the effectiveness of each of this concepts, and to tune the parameters. The purpose of the first experiment was verifying if the inclusion of the spectral coefficients has a positive effect in the accuracy of the system. Fig. 4.a shows the accuracy's results when the bandwidth varies from 0Hz to 40Hz, for the cases of filters bandwidth 2Hz and 3Hz. Notice that 0 is the case where no spectral features are used. From this figure, we can observe that spectral features help to increase the accuracy of the system, and we can also observe that a filterbank's bandwidth larger than 25Hz is a good choice to obtain the best results.

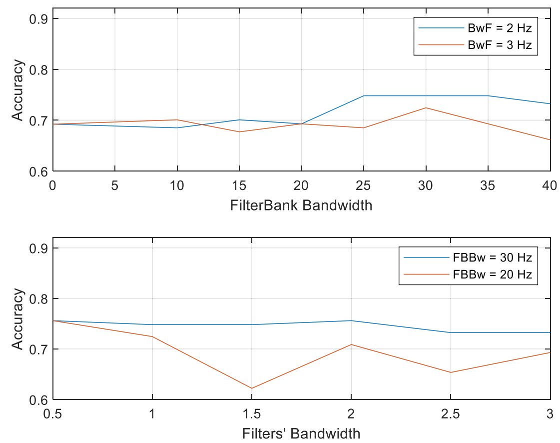


FIGURE 4. a. Accuracy results for depression writing when varying filterBank bandwidth form 0Hz to 40Hz. Notice that 0Hz is an experiment without spectral features, b. Accuracy results when varying the filters' bandwidth from 0.5 to 3.0.

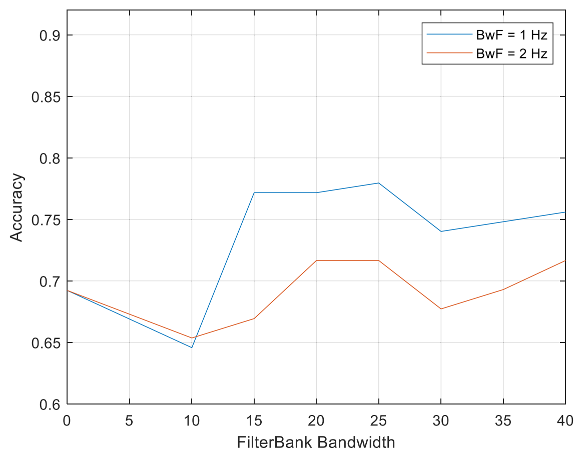


FIGURE 5. Accuracy results for the depression-writing database when varying FilterBank bandwidth. For bandwidth filter of 1 and 2 Hz.

And we can observe that accuracy results were better when the filter bandwidth is 2Hz.

Second experiment had the objective to optimize the filter bandwidth that maximizes accuracy. Fig. 4.b shows the accuracy's results when filter bandwidth is varied from 0.5 to 3.0, for bandwidth 20Hz and 30Hz. From this figure, we can observe that filterbank bandwidth of 30Hz has the best results, and that choosing a filter bandwidth between 0.5 to 2.0 is an appropriate choice.

The filterbank can be configured with different parameters, such as: initial frequency of the filterbank, the center of the filterbank, and the bandwidth of the filterbank. For example, we developed an experiment for a non-continuous filterbank for two cases, the first case was when the filterbank starts in frequency 1 Hz, the centers were separated by 2 Hz and the filters' bandwidth was set to 1Hz. In the second case, the filterbank starts in frequency 1.5 Hz, the centers were separated by 3 Hz and the filters' bandwidth was set to 2Hz. Fig. 5 shows the results of these two cases, and

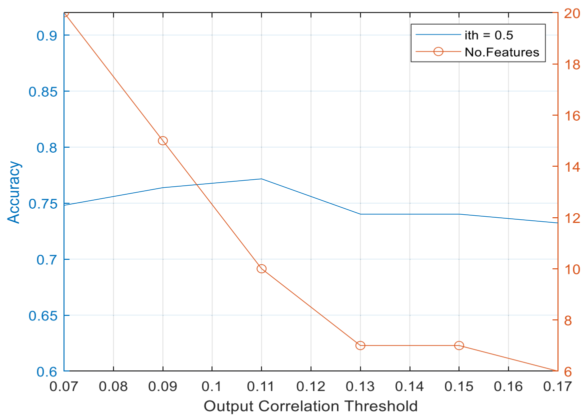


FIGURE 6. **Accuracy results and number of selected features after applying mFCBF algorithm and varying** iTh**, and fixed** oTh **for depression-writing database.**

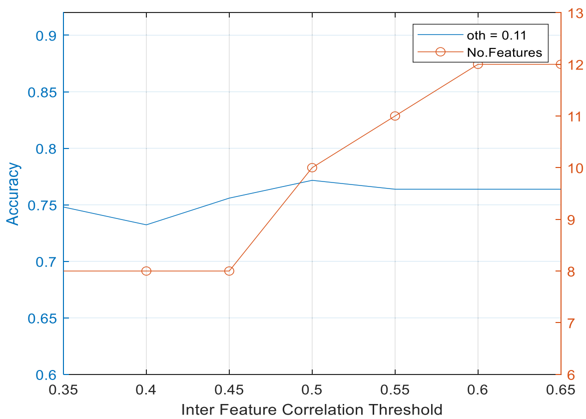


FIGURE 7. **Accuracy results and number of selected features after applying mFCBF algorithm and varying** iTh**, and fixed** oTh **for depression-writing database.**

we can observe that best results are obtained for the first case.

The fourth experiment was designed to optimize the Feature-Output Correlation Threshold (oTh). Fig. 6 shows the experiments accuracy for SVM(rbf) classifier, and we can observe that oTh values around 0.11 are appropriate choice to obtain the higher accuracy. In this experiment, the rest of the parameters where keep fixed. The same figure shows the number of features when varying oTh and keeping fix the inter-Feature Correlation Threshold to 0.5. We can observe that by increasing the value of oTh the number of coefficients decreases.

The fifth experiment was designed to optimize the Inter-Feature Correlation Threshold(iTh). Fig. 7 shows the experiments accuracy for SVM(rbf) classifier, and we can observe that iTh values around 0.5 are good choices. The same figure shows the number of features when varying iTh and keeping fix the Feature-Output Correlation Threshold to 0.5. We can observe that by increasing the value of iTh the number of coefficients increases.

Next, we show the accuracy results for all tasks and emotion states in Table 4. We can observe that depression was the

TABLE 4. **Results.**

Emotional state	Task type	Baseline (%)	Max (%)	Max Performance (%)
Depression	Writing	69.22	80.31	16
	Drawing	72.48	75.59	4
	Both	71.47	74.01	4
Average		**71.06**	**76.64**	**8**
Anxiety	Writing	56.59	68.50	21
	Drawing	58.68	67.71	14
	Both	58.53	72.44	24
Average		**57.93**	**69.29**	**20**
Stress	Writing	50.54	67.71	34
	Drawing	59	67.71	15
	Both	61.24	70.07	14
Average		**56.93**	**68.5**	**20**
Overall Average		61.97	71.47	15

emotion state with better accuracy, followed by anxiety, and the emotional state with lower accuracy is stress. We can also observe that depression for written tasks was the one with the best performance. Moreover, the results obtained when the spectral and cepstral features are included results in an average improvement of 15% compared to the baseline work done in EMOTHAW paper.

The classifier model used is the SVM 'rbf' kernel model, with a Leave-One-Out (LOO) Cross Validation technique, and a Gaussian Noise Data augmentation.

VIII. COMMENTS AND CONCLUSION

This work showed the performance results of depression, anxiety, and stress emotional state recognition on the EMOTHAW database. Since our data is small, first, we augmented the training data by randomly selecting twenty percentage of the training data and adding a small random Gaussian noise to the features. As next step, we calculate time, cepstral and spectral features, followed by a selection of these features using a modified version of the Fast Correlation-Based Filtering method. Next, we train a radial basis SVM model using the Leave-One-Out (LOO) methodology.

To tune the parameters, we develop several experiments. In our first experiment, we confirmed that adding the spectral coefficients of the horizontal and vertical displacement and pressure, has a positive effect on the accuracy of the system with an increase of 9.2% for depression writing. In the first experiments, we limited to non-overlapped, linear filters starting in 0.5Hz, and we found that 2Hz filters' bandwidth is a good choice to optimize accuracy. In this work, we found that a value of around 25Hz is a good choice for the filtebank's bandwidth.

In the next experiments, we limited to 1Hz-separated-non-overlapped, 1Hz-bandwidth-linear filters starting in 1.5Hz and we found that for this modification of the filterbank configuration the accuracy improves another 3%.

The number of coefficients was significantly reduced with the modified Fast Correlation-Based Filtering (FCBF).

We also found that for this task the maximum accuracy for the threshold values of both the features-correlation-with-the-output and correlation-between-features, both required in mFCBF, were around 0.5 and 0.11, respectively. We observed that in most of the experiments, most of the selected features were cepstral, and then number of spectral and time features had more less the same proportion.

Using this parameter tunning, we developed experiments for stress, anxiety, and depression and for all cases, accuracy improved the baseline. The highest improvement, compared with the baseline, was 34%, and the best accuracy was for depression writing achieving 80.31%.

We can conclude that if we want to detect Depression, writing is the best, if we want to detect Anxiety and stress, we need to combine writing and drawing information.

As a future work, we can explore other filterbank configurations and we can also explore adding spectral and cepstral features of altitude of the pen with respect to the tablet's surface and azimuth angle of the pen with respect to the tablet's surface.

REFERENCES

[1] M. Faundez-Zanuy, J. Fierrez, M. A. Ferrer, M. Diaz, R. Tolosana, and R. Plamondon, "Handwriting biometrics: Applications and future trends in e-security and e-health," *Cogn. Comput.*, vol. 12, no. 5, pp. 940–953, Sep. 2020.

[2] R. Plamondon and S. N. Srihari, "Online and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 63–84, Jan. 2000.

[3] S. Z. Li and A. K. Jain, *Encyclopedia of Biometrics*. New York, NY, USA: Springer, 2015.

[4] C. De Stefano, F. Fontanella, D. Impedovo, G. Pirlo, and A. S. di Freca, "Handwriting analysis to support neurodegenerative diseases diagnosis: A review," *Pattern Recognit. Lett.*, vol. 121, pp. 37–45, Apr. 2019.

[5] World Health Organization. (2017). *Depression and Other Common Mental Disorders: Global Health Estimates*. Accessed: Dec. 14, 2020. [Online]. Available: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf

[6] F. Scibelli, A. Troncone, L. Likforman-Sulem, A. Vinciarelli, and A. Esposito, "How major depressive disorder affects the ability to decode multimodal dynamic emotional stimuli," *Frontiers ICT*, vol. 3, p. 16, Sep. 2016.

[7] A. Esposito, G. Raimo, M. Maldonato, C. Vogel, M. Conson, and G. Cordasco, "Behavioral sentiment analysis of depressive states," in *Proc. 11th IEEE Int. Conf. Cogn. Infocommun. (CogInfoCom)*, Sep. 2020, pp. 209–214.

[8] A. I. Turner, N. Smyth, S. J. Hall, S. J. Torres, M. Hussein, S. U. Jayasinghe, K. Ball, and A. J. Clow, "Psychological stress reactivity and future health and disease outcomes: A systematic review of prospective evidence," *Psychoneuroendocrinology*, vol. 114, Apr. 2020, Art. no. 104599.

[9] A. T. Beck and A. Beamesderfer, "Assessment of depression: The depression inventory," in *Modern Trends in Pharmacopsychiatry Psychological Measurements in Psychopharmacology*, vol. 7, P. Pichot, Ed. Basel, Switzerland: Karger, 1974, pp. 151–169.

[10] B. R. Schlenker and M. R. Leary, "Social anxiety and self-presentation: A conceptualization model," *Psychol. Bull.*, vol. 92, no. 3, pp. 641–669, Nov. 1982.

[11] C. D. Spielberger, "State-trait anxiety inventory," in *The Corsini Encyclopedia of Psychology*. Hoboken, NJ, USA: Wiley, 2010.

[12] H. Selye, *The Stress of Life*. New York, NY, USA: McGraw-Hill, 1984.

[13] A. V. Christensen, J. K. Dixon, K. Juel, O. Ekholm, T. B. Rasmussen, B. Borregaard, R. E. Mols, L. Thrysøe, C. B. Thorup, and S. K. Berg, "Psychometric properties of the danish hospital anxiety and depression scale in patients with cardiac disease: Results from the DenHeart survey," *Health Qual. Life Outcomes*, vol. 18, no. 1, pp. 1–13, Dec. 2020.

[14] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "Benchmarking desktop and mobile handwriting across COTS devices: The e-BioSign biometric database," *PLoS ONE*, vol. 12, no. 5, 2017, Art. no. e0176792.

[15] L. Likforman-Sulem, A. Esposito, M. Faundez-Zanuy, S. Clemencon, and G. Cordasco, "EMOTHAW: A novel database for emotional state recognition from handwriting and drawing," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 2, pp. 273–284, Apr. 2017.

[16] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in *Proc. 20th Int. Conf. Mach. Learn. (ICML)*, 2003, pp. 856–863.

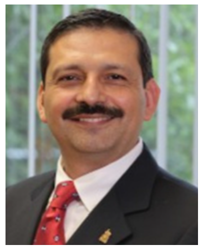
[17] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[18] P. F. Lovibond and S. H. Lovibond, "The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the beck depression and anxiety inventories," *Behav. Res. Therapy*, vol. 33, no. 3, pp. 335–343, Mar. 1995.

[19] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-mental state," *J. Psychiatric Res.*, vol. 12, no. 3, pp. 189–198, 1975.

[20] P. Kline, *Handbook of Psychological Testing*. Florence, Italy: Taylor and Francis, 2013.

[21] G. Severino. (Jul. 26, 2018). *Severino Italian Translation*. Accessed: Dec. 13, 2020. [Online]. Available: http://www2.psy.unsw.edu.au/dass/Italian/Severino.htm



JUAN A. NOLAZCO-FLORES (Senior Member, IEEE) was born in Gral. Terán, Nuevo León, Mexico. He received the B.Sc. and M.Sc. degrees from the Tecnológico de Monterrey, Monterrey, México, in 1986 and 1988, respectively, and the M.Phil. and Ph.D. degrees from the University de Cambridge, Cambridge, U.K., in 1992 and 1995, respectively.

From 2017 to 2020, he was the Dean of the School of Engineering and Science, South Region, Tecnológico de Monterrey, Mexico. From 2011 till 2017, he was the Head of Information Technology with the Tec of Monterrey, Monterrey campus. He was Visiting Scholar with Marburg University and Mannheim University, Germany, Carnegie Mellon University, Pittsburgh, PA, USA, Zaragoza University, Spain, and Cambridge University, U.K. He is currently the Head of the Data Science Hub, Tec of Monterrey. He is author of more than 50 articles indexed in Scopus citation report. His research interests include in the fields of signal processing and data science applied to speech, security, and health.

Dr. Nolazco is member of the Mexican Academy of Science, a member of the Mexican Research System, and an ACM Senior Member. He was awarded in 2005 and 2009 with the ITESM Teaching and Research Award. He is certified ABET Evaluator, and he has participated as an Evaluator of the Computer Science programs with University of West Florida, Grand Canyon University, Nevada, USA, and The University of Iowa, Iowa, City, IA, USA.

MARCOS FAUNDEZ-ZANUY was born in Barcelona, Spain. He received the Ph.D. degree from the Polytechnic University of Catalunya, in 1998.

From 2009 till 2018, he was the Dean of the Escola Superior Politecnica Tecnocampus (the Polytechnic University of Catalunya, till 2015, Pompeu Fabra University, after 2015). From 2010 till 2019, he was the Head of research with Tecnocampus. He is currently a Full Professor with ESUP Tecnocampus Mataro, where he also heads the Signal Processing Group. He is an author of more than 100 articles indexed in ISI Journal citation report, more than 100 conference papers, around ten books, and responsible of ten national and European research projects. He is a Spanish liaison of EURASIP. His research interests include in the fields of biometrics applied to security and health. He was the Initiator and Chairman of the European COST action 277 Nonlinear Speech Processing, and the Secretary of COST action 2102 Cross-Modal Analysis of Verbal and Non-Verbal Communication.

OLIVER A. VELÁZQUEZ-FLORES was born in Saltillo, Coahuila, México. He received the B.Sc. degree in mechatronics engineering from Tecnológico de Monterrey, Saltillo Campus, México, in 2019. He is currently pursuing the M.Sc. degree in computer science with Tecnológico de Monterrey, Monterrey Campus, México, researching through his thesis the impact of handwriting in emotion recognition.

He is currently a Research Assistant with the School of Engineering and Science, Tecnológico de Monterrey, Monterrey. His research interests include e-security and data science applied to health and psychology.

GENNARO CORDASCO received the Laurea degree *(cum laude)* in informatica (computer science) from the University of Salerno, Italy, in 2002, discussing a thesis in Distributed algorithms whose title is: Un algoritmo di mapping ottimale per alberi binari completi su memorie parallele (An optimal mapping algorithm for complete binary trees to parallel memories), and the Ph.D. degree in April 2006, discussing a thesis on Peer to Peer Systems whose title is: On Designing Overlay Networks for Peer to Peer Systems.

The thesis was supervised by Prof. Alberto Negro. The thesis was supervised by Prof. Alberto Negro and Prof. Vittorio Scarano. From 2002 to 2006, he attended the Ph.D. Program in computer science with the Dipartimento di Informatica ed Applicazioni, University of Salerno, Italy, (Advisor: Prof. Alberto Negro). He is currently an Assistant Professor with the Dipartimento di Psicologia, Facoltà di Psicologia, Università degli studi della Campania Luigi Vanvitelli, Italy, and an Affiliate of the International Institute for Advanced Scientific Studies, Vietri sul Mare, Italy. From March 2005 to November 2005, he was a Research-Scholar with the Theoretical Aspect of Parallel and Distributed Systems (TAPADS) Laboratory, Department of Computer Science, University of Massachusetts at Amherst, Amherst, MA, USA, doing research activity under the supervision of Distinguished Professor Arnold L. Rosenberg. From May 2006 to March 2011, he was a Postdoctoral fellow (Assegnista di Ricerca) with the Dipartimento di Informatica, University of Salerno, cooperating in the research activity of the ISISLab Laboratory.

ANNA ESPOSITO received the Laurea degree *(summa cum laude)* from the University of Salerno with a thesis published on Complex System, in 1992, and the Ph.D. degree in applied mathematics and computer science from the University of Naples Federico II, with a Ph.D. thesis published on Phonetica, in 2002, and developed at MIT, RLE Laboratory, Cambridge, MA, USA.

She has been a Postdoctoral Researcher with the International Institute for Advanced Scientific Studies (IIASS), and an Assistant Professor with the Department of Physics, University of Salerno, where she taught Cybernetics, Neural Networks, and Speech Processing, from 1996 to 2000. She held a Research Professor position at the Department of Computer Science and Engineering, Wright State University, Dayton, OH, USA, from 2000 to 2002. She has been a Full Professor in computer science with Università della Campania L-Vanvitelli (UVA), since 2003, teaching Cognitive and Algorithmic Issues of Multimodal Communication, Social Networks Dynamics, Cognitive Economy, and Decision Making. She is author of more than 250 peer-reviewed publications and editor/coeditor of more than 30 international books. She is also the Head of the Behaving Cognitive Systems Laboratory (BeCogSys), UVA. The lab is participating to the H2020 funded projects: a) Empathic, and b) Menhir, c) SIROBOTICS, and d) ANDROIDS.