A S T E S

A New Video Based Emotions Analysis System (VEMOS): An Efficient Solution Compared to iMotions Affectiva Analysis Software

Nadia Jmour*,1, Slim Masmoudi2, Afef Abdelkrim1

1Research Laboratory Smart Electricity & ICT, SE&ICT Lab., LR18ES44. National Engineering School of Carthage, University of Carthage, Charguia, 2035, Tunisia

2Innovation Lab, Naif Arab University for Security Sciences, Riyadh, 14812, KSA

A R T I C L E   I N F O

A B S T R A C T

The Micro-facial expression is the most effective way to display human emotional state. But it needs an expert coder to be decoded. Recently, new computer vision technologies have emerged to automatically extract facial expressions from human faces. In this study, a video-based emotion analysis system is implemented to detect human faces and recognize their emotions from recorded videos. Relevant information is presented on graphs and can be viewed on video to help understanding expressed emotions responses. The system recognizes and analyzes emotions frame by frame. The image-based facial expressions model used deep learning methods. It was tested with two pre-trained models on two different databases. To validate the video-based emotion analysis system, the aim of this study is to challenge it by comparing the performance of the initial implemented model to the iMotions Affectiva AFFDEX emotions analysis software on labeled sequences. These sequences were recorded and performed by a Tunisian actor and validated by an expert psychologist. Emotions to be recognized correspond to the six primary emotions defined by Paul Ekman : anger, disgust, fear, joy, sadness, surprise, and then their possible combinations according to Robert Plutchik's psycho-evolutionary theory of emotions. Results show a progressive increase of the system's performance, achieving a high correlation with Affectiva. Joy, surprise and disgust expressions can reliably be detected with an underprediction of anger from the two systems. The implemented system has shown more efficient results on recognizing sadness, fear and secondary emotions. Contrary to iMotions Affectiva analysis results, VEMOS system has recognized correctly sadness and contempt. It has also successfully recognized surprsie and fear and detect the alarm secondary emotion. iMotions Affectiva has confused surprise and fear. Finally, compared to iMotions the system was also able to detect peak of morbidness and remorse secondary emotions.

1. Introduction

Humans communicate through the expression of their emotions, using various channels (physiological reactions, behaviors, voice, etc.). The emotion is the strong link between cognition and behavior, and between attention, representation and human performance [1]. Facial expression is the most direct and effective way to display human emotional states [2]. In addition, humans use involuntary micro-expressions to unveil their feelings [3]. Nevertheless, they are unable to recognize them accurately.

Indeed, only 47% of recognition accuracy is observed at trained adults [4].

Since the 1960's, several facial expressions and emotions classification models have been psychologically studied. Psychologists such as Paul Ekman and Friesen have studied micro-expressions in their research, showing that these micro-expressions are involuntary, incontrollable, and expressed by humans as a very short reaction to a lived situation. Usually, they last less than 1/25 second which makes it hard to find them in real-time [5, 6].

*Corresponding Author: Nadia Jmour, nadiajmour@gmail.com

The authors showed that micro-expressions can be a lie detection tool, In fact, micro-expressions allow recognizing the real emotions expressed by people being in critical situations as the case of criminal interrogation or job recruitment interview [7].

Thus, psychologists established a method to recognize micro expressions by defining criteria of the facial muscles' movements related to each specific micro-expression. Facial movements are defined into 46 action units and each unit presents a contraction or relaxation of a facial muscle. Thereafter, each micro-expression corresponds to the presence of several action units. This allows, for example, recognizing a true involuntary smile, characterized by the movement of the zygomaticus major and of the orbicularis of the eye muscles, [8], from a false smile activating with the movement of the zygomaticus muscle. The developed system is known as the Facial Action Coding System (FACS) and it has become the most used in scientific research [9].

However, intelligent automated recognition of emotional micro-expressions is still a quite new topic in computer vision. Recently, based on machine learning algorithms, and with the availability of facial expressions databases, emotion analysis systems have become able to automatically recognize micro-facial expressions and emotional states, and give an added value technological undertaking [10, 11].

As defined in [12], these systems operate in three steps: Face detection, feature extraction, facial expression and emotion classification.

In traditional methods, feature extraction and feature classification steps are independent. Compared to recent research [13], these machine learning techniques for facial expression recognition are considered as classical techniques.

Nowadays, researchers tend to highlight the advantage of applying deep learning techniques for facial expression recognition in different applications [14, 15]. They consider that applying deep learning reduces human intervention in the analysis process by automating the phase of extracting and learning features [16, 17]. Also, illumination, identity bias and head pose problems are avoided. The need of a large amount of data is also resolved, [18–20].

Several researchers as [21–24], have studied and developed emotion analysis systems for applications in different fields such as human resources, medical field, etc. Authors in [11], employed in his research a deep learning system based on face analysis for monitoring customer behavior specifically for detection of interest. Authors in [25], applied deep learning-based face analysis system in radiotherapy that monitors the patient's facial expressions and predicts patient's advanced movement during the treatment. Authors in [26], considered that facial expression analysis system can be applied on the field of marketing.

Systems such as EmotioNet, iMotions platform [27], using Affectiva Affdex algorithm, [28], Noldus [29], EmoVu [30], and Emotient employ facial expression recognition technologies to create commercial products using information related to their customers. These systems are able to recognize different metrics as head orientation, facial landmarks and basic emotions and most of them apply deep learning [31–33].

The first contribution of this study is to develop an emotion analysis system that recognizes human emotions from recorded videos.

The second contribution is to challenge the system's performance by comparing measures with iMotions Affectiva analysis software measures obtained by analyzing labeled recorded videos.

This study aims also to employ the deep learning-based image analysis model presented in our recent research [18, 20].

Since micro-facial expressions occur during a split second, the proposed system extracts images frame by frame, and every frame is analyzed by the implemented deep learning model. Recorded information is viewed either in an analyzed video and/or presented on the business analysis Power BI Desktop from Microsoft [34], by using visualization graphs to help understanding expressed emotional responses.

Emotions to be recognized correspond to the six facial expressions defined by authors: anger, disgust, fear, joy, sadness, surprise, and then their possible combinations according to Robert Plutchik's psycho-evolutionary theory of emotions.

We expected that the six basic facial expressions can be correctly recognized by the implemented system, with a relatively high accuracy, and can give as an output person's true emotions. As a criterion/external validity of the implemented system, we further expected its significant correlations with iMotions Affectiva measures. Also, secondary emotions identified by Plutchik were expected to be correctly recognized by the proposed new system.

The remainder of this paper is organized into four sections. Section 2 presents the implemented video-based emotion analysis system with an example of a tested scenario. Section 3 presents comparison and challenging results of the implemented system with iMotions Affectiva software. Finally, section 4 summarizes the conclusions

2. Video-based Emotion analysis system (VEMOS) implementation

2.1. Process analysis description

Micro-expressions cannot be controlled, difficult to hide and occur within a split of second. The proposed system extracts first frames from recording video every 1/25 second.

Then, it analyzes and recognizes automatically frame by frame the six basic emotions [32]. Extracted frames are affected as input in the image-based facial expression analysis system, which was implemented in our recent research [18].

This system used deep learning techniques to recognize facial expression images more accurately. It performed by training a deep convolutional neural network (CNN) on a set of face images. This network is composed of two parts:

- The convolutional part for feature extraction: it takes the image as input of the model and then convolution maps are flattened and concatenated into a features vector, called CNN code.

- The fully connected layer part: this part combines the characteristics of the CNN code to classify the image.

To explore the power of these CNN with accessible equipment and a reasonable amount of annotated data, it is possible to explore pre-trained neural networks available publicly [35].

This technique is called transfer learning, it uses knowledge and information acquired from a general classification task and applies it again to a particular new one.

In our recent research, we compared the use of the VGG-16 and the InceptionV3 pre-trained models on the ImageNet dataset, [36], by applying transfer learning technique using the Kaggle (Facial Expression Recognition Challenge) and Extended Cohn-Kanade CK+ facial expression datasets [37].

We used for our implementation Python libraries such OpenCV and TensorFlow on GPU NVIDIA version 391.25, processor type 16xi5-7300HQ. We trained the model on 100 epochs, with a learning rate of 0.0001 [20]. We leveraged the Keras implementation of VGG-16 and InceptionV3.

The data is split into 90% for training and 10% for testing. Experiments had accuracies of 45.83% and 43.22% using the two pre-trained models on the FER-2013. Then, an accuracy of 73.98% was observed by transferring information from the VGG-16, and by training the new architecture on the Extended Cohn-Kanade facial expression dataset using the top frame from every sequence.

Thus, our recent research approved the effect of the distribution and the quality of the data for an efficient training and application of the facial expression recognition task. So, every fraction of a second, this model assigns a score corresponding to each emotion. Output Information is recorded from every analyzed frame.

Based on the recorded facial expressions information, we proposed a new approach for the determination of other emotional states from video data. This approach is inspired from the Robert Plutchik's psycho-evolutionary theory of emotions [38–40].

According to Plutchik, basic emotions can be combined with each other to form different emotions. Table 1 presents the possible combinations.

Table 1: Plutchik's emotions possible combinations.

Primary dyads	Results
Joy and Trust	Love
Trust and fear	Submission
Fear and surprise	Alarm
Surprise and sadness	Disappointment
Sadness and disgust	Remorse
Disgust and anger	Contempt
Anger and anticipation	Aggression
Joy and fear	Guilt
Trust and surprise	Curiosity
Fear and sadness	Despair
surprise and disgust	horror
sadness and anger	Envy
Disgust and anticipation	Cynism
Anger and joy	Pride
Anticipation and trust	Fatalism
Joy and surprise	Delight

Trust and sadness	Sentimentality
Fear and disgust	Shame
Surprise and anger	Outrage
Sadness and anticipation	Pessimism
Disgust and Joy	Morbidness
Anger and trust	Dominance
Anticipation and fear	Anxiety

The Plutchik theory defines emotions as confidence and anticipation. In our case, we explore only combinations based on the six emotions below: anger, disgust, fear, joy, sadness and surprise. Emotions which Plutchik presents as conflict combinations are not considered in our approach.

This approach allows having more information about the emotional state of the person from the video data type. For example, according to Plutchik's theory, the emotion of delight would be a mixture of joy and surprise. Delight's score is defined as the mean of the two basic emotions.

At the end of the process, two types of output are provided. First, the following output information is extracted and visualized through graphs on the Power BI Desktop [41] :

- Dominant basic emotion occurrence: It gives dominant and non-dominant emotion occurrence.

- Maximum basic emotion peak: It gives maximum peak and the corresponding time.

- Emotions Variability: Its records frame by frame the six emotions score over the time.

- Dominant secondary emotion occurrence: It gives dominant and non-dominant secondary emotion occurrence.

- Maximum secondary emotion peak: It gives maximum peak and the corresponding time.

Then, an analyzed video is extracted where face traking and basic and secondary emotions scores over the time are displayed. We leveraged Pandas, and Csv librairies for recording pertinent information from extracted images. Opencv, Mathplotlib and Dlib librairies were used for a better visualization of the resulting analyzed video.

2.2. Tested scenario and results visualization

We chose Carlos Ghosn intervention as a tested video. Carlos Ghosn is the head of Nissan and the French automaker Renault. In November, 2018, Nissan announced that an internal investigation into Ghosn and another executive found Ghosn's compensation had been underreported. There was also evidence that Ghosn committed other misconduct, including personal use of company assets. Japanese authorities arrested Ghosn in Tokyo. He has already spent a hundred and eight days before his second arrestation. On April 9th, 2019, Ghosn posted a structured YouTube video, [42], that he recorded four days before, in which he presented three messages and where he publicly proclaimed that he was innocent of all the accusations that came around these charges that are all biased, taken out of context, twisted in a way to paint a personage of greed, and a personage of dictatorship. Carlos Ghosn spoke calmly, neutral and looked fixedly at the camera in the recorded video.

We analyzed the first message in the video with our implemented video-based emotion analysis system.

Output Extracted information is presented using visualization graphs on the Power BI desktop. Primary and secondary emotions occurrences are shown in Figure 1.
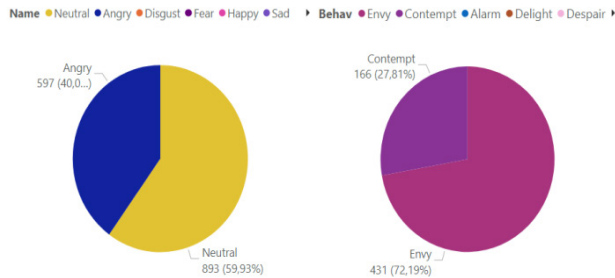

Figure 1: Emotions Occurrence.

Neutral and Anger categories dominated on the whole video with 59.9% and 33.56% respectively. Envy and contempt secondary emotions occurred with 72.1% and 27% respectively.

The variability of each basic emotion over the time is presented in Figure 2. Each color describes an emotion as it is presented on the legend.
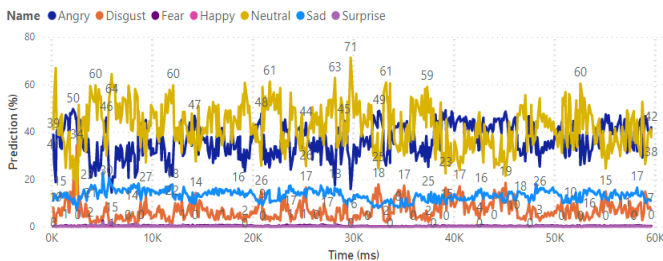
Figure 2: Basic emotions variability.

The figure shows that the two dominant emotions vary with important values. The maximal recorded peak value of the Anger micro-facial expression is 50%. Disgust and Sadness appear with smaller values and they do not exceed 30%.

Secondary emotions are also detected by the system and analysis results are presented in Figure 3.
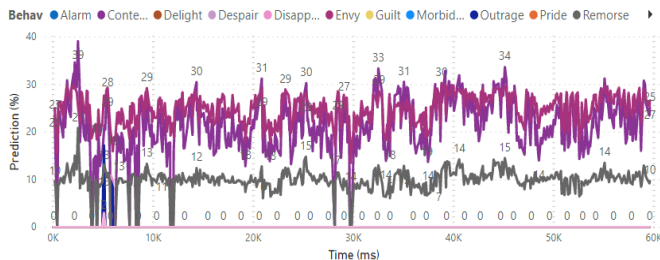
We noticed the appearance of contempt and envy during the time. Peaks values are 30% and 39% respectively.

The authors tried to be calm and keep his emotions stables. The switch between Anger, Neutrality, on one hand, and contempt, envy, on the other hand, explains also how he maybe tried to manage his emotions.

Figure 4 presents a screenshot from the analyzed video. It shows Carlos while announcing his innocence and expressing envy combining anger and sadness emotions.
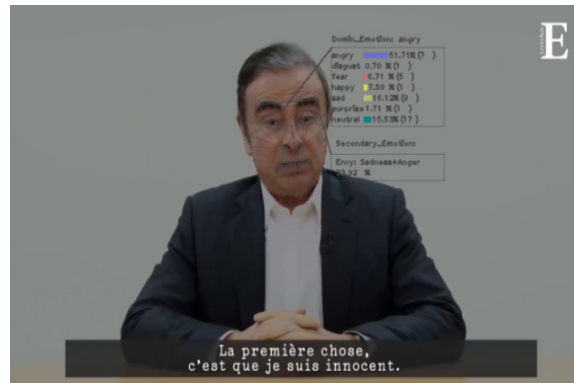

Figure 4: Screenshot of Carlos Ghosn's analyzed video.

Thanks to VEMOS, we unveiled the undetected basic and secondary emotions by a human.

Later, we correlate our results with the 22th April event where the Ex-CEO of Renault-Nissan Carlos Ghosn has been charged again, for breach of trust aggravated by the Tokyo court. This can maybe explain the presence of the detected negative emotions.

3. Comparison with iMotions Affectiva facial expressions analysis software

The performance of the implemented video-based emotions analysis system VEMOS had to be validated. So, we decided to challenge the system by comparing the performance with the iMotions Affectiva AFFDEX emotion analysis software.

3.1. iMotions Biometric Research Platform

iMotions [27, 43], is a software solution helping for data collection, analysis, and quantifying engagement and emotional responses. The iMotions Platform is an emotional recognition software that integrates multiple sensors including facial expression analysis, GSR, eye tracking, EEG, and ECG/EMG along with survey technologies for multimodal human behavior research.

Videos are imported and analyzed in iMotions Platform using the Affectiva Affdex facial expression recognition engine [27, 28].

The Affectiva's Affdex SDK toolkit (in collaboration with iMotions) [28, 44–46], is one of the most known and used emotions analysis tool. It uses the Facial Action Coding system (FACS). The tool detects facial landmark, classify facial action, and recognizes the seven basic emotions.

The system's algorithm uses first a histogram of oriented gradient (HOG) to extract pertinent images regions. Then, an SVM is used for the classification step. The training process is applied on 10000 images and determined a score of 0–100 for each class. The testing process is applied on 10000 images.

3.2. Carlos Ghosn's video correlation results

The video of Carlos Ghosn is analyzed with iMotions Affectiva software. Extracted information are presented in the Table 2.

Table 2: iMotions Affectiva emotions occurence results.

Emotion	Occurrence
Anger	184
Disgust	22
Fear	0
Joy	0
Sadness	29
Surprise	321
Neutral	1429

Through the iMotions Affectiva system, the occurrence of emotions is detected when its corresponding score exceeds the threshold value of 50%.

Taking into account this characteristic, we noticed that results given by iMotions and VEMOS systems lead to the same psychological interpretation of the emotional state of Carlos that we have concluded from the video. Then, we used the Bravais-Pearson correlation coefficient to compare analysis results obtained by the two systems.

Obtained values range between -1.0 and 1.0. A correlation of -1.0 shows a perfect negative correlation, while a correlation of 1.0 shows a perfect positive correlation. A correlation of 0.0 shows no linear relationship between the movement of the two variables. We considered that values less than 0.5 as low values and values greater than 0.5 as high values. Table 3 shows the correlation results.

Table 3: Correlation coefficient with iMotions software

Emotion	Correlation coefficient Initial model	p
Anger	0.051	$p>.05$
Disgust	0.38	$p<.05$
Fear	-0.19	$p<.05$
Joy	0.11	$p<.05$
Sadness	0.06	$p>.05$
Surprise	0.0015	$p>.05$

We obtained low values of correlation for the six basic emotions. As the video is not labeled and is not analyzed by non-verbal professional decoder, results remain open and they do not validate the performance of the system.

To ensure the reliability of our results we proposed to apply our analysis on labeled videos. So, we suggested to prepare two corpus of labeled sequences that were recorded and performed by a Tunisian actor and validated by a professional psychologist. The first corpus contains basic emotions sequences and the second one contains secondary emotions sequences.

3.3. Labeled recorded Sequences

We first collected labeled sequences from YouTube video. The psychologist confirms that emotions expressed in these sequences present muscles' mouvement that correspond to the needed emotions. Then, a Tunisian Actor was hired to perform videos under the monitoring and guidance of a professional psychological doctor.

Before recording the videos, the psychologist gives more details about the muscles to move specifying the facial action coding units per emotion. Then, video recording starts with the actor looking into the camera and showing successively the six facial expressions anger, disgust, fear, joy, sadness, and surprise, repeating the action several times. Thus, many sequences were recorded and only one per emotion was chosen and validated by the psychologist. Table 4 shows the details about the videos characteristics.

Table 4: Basic emotions Sequences characteristics

Attribute	Description
Length of sequences	5000 to 10000 milliseconds
Number of frames	25 fps
Expressions classes annotation	Angry, disgust, fear, happy, sad, surprise
Video format	mp4
Number of subjects	2

Similarly, the actor looked into the camera and expressed the expected secondary emotions. Sequences were recorded and validated by the psychologist. Table 5 below presents characteristics of the prepared sequences.

Table 5: Prepared secondary emotions Sequences characteristics.

Attribute	Description
Length of sequences	5000 to 10000 milliseconds
Number of frames	25 fps
Expressions classes annotation	alarm, contempt, morbidness, despair, remorse
Video format	mp4
Number of subjects	1

3.4. Basic emotions correlation results

To explore the reliability of the Affectiva algorithm for basic emotions recognition, we proposed to take advantages of the constructed corpus of the labeled sequences. So, recorded primary emotion sequences were analyzed by the VEMOS implemented system. They were also analyzed by the iMotions Affectiva Affdex algorithm and temporal information was imported for the 6 sequences.

The analysis results should first correlate with the expected emotion. Then, for a better visualization, the extracted values are presented on graphs using the Power BI Desktop.

By plotting these graphs, facial expression measures are compared to iMotions stored facial expression measures.

Figure 5 shows emotions variability over the time obtained by the VEMOS system for every sequence. Analysis results show that the initial implemented system has difficulties on classifying some facial expressions. Compared to the expected emotions disgust, fear, joy and sadness categories are misclassified. Emotions as angry and surprise are correctly classified and reached maximum values of 87% and 99% respectively.
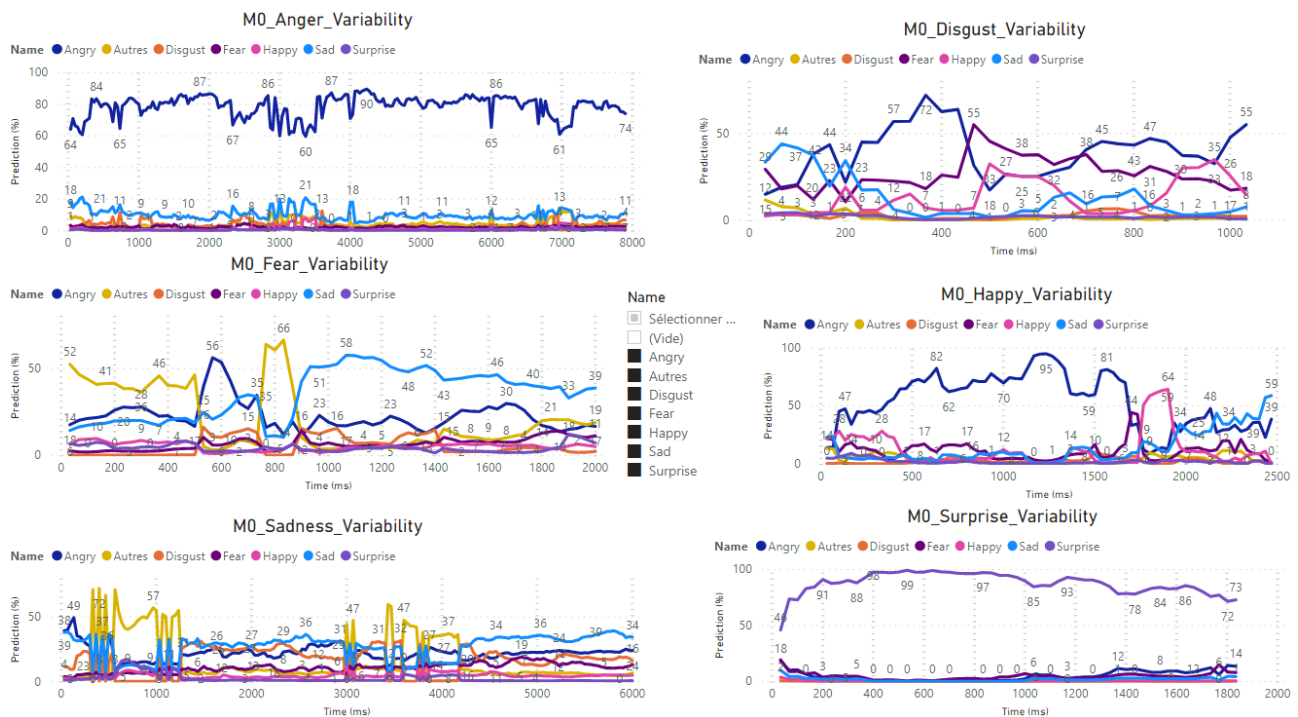
Figure 5: Analysis results with the initial implemented model.

In order to compare results to iMotions platform analysis results, we used the Bravais-Pearson correlation coefficient. Table 6 shows correlation coefficient of each emotion.

Table 6: Correlation coefficient with iMotions software

Emotion	Correlation coefficient Initial model	p
Anger	0.01	$p>.05$
Disgust	0.33	$p<.05$
Fear	-0.41	$p<.05$
Joy	0.19	$p<.05$
Sadness	0.03	$p>.05$
Surprise	0.15	$p<.05$

Obtained results show small values and low correlation between the implemented VEMOS system and iMotions Affectiva results.

Consequently, we proposed to challenge the system and improve its performance. The idea was to challenge the number of data and the distribution of emotions categories used for the training of the image-based facial expression system.

Our recent approach, [18], has shown the importance of the quality of the data in the performance of the implemented model. The Fer-2013 dataset presents images that were collected from google and not properly treated. Sometimes, we found images where there is text on the faces. So, compared to the Fer-2013, our best accuracy was obtained by training and testing the model on the CK+ dataset, which was implemented by exploring the top frame of each sequence from the data.

In this approach, we propose to increase the number of input frames from the CK+ dataset by taking in consideration the data distribution. Then, the model is trained by keeping the same

proposed architecture and its same characteristics. We propose to construct a new corpus based on the CK+ dataset by taking five frames per sequence and to integrate a new category called "others". Neutral faces and all other faces movements that are not recognized by the model as one of the basic emotions are put on the new class "others".

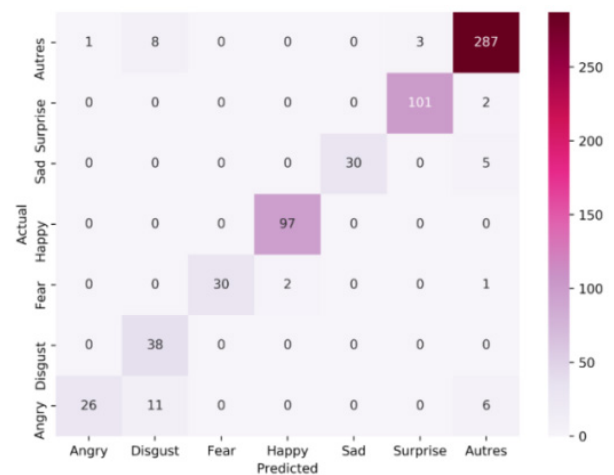The model reached 93.98% test accuracy. Confusion matrix is shown in Figure 6.

Figure 6: Confusion matrix.

To test the reliability and the performance of the challenging system, we used the same labeled sequences. Sequences are already analyzed by the iMotions Affectiva software. So, they were only imported and analyzed by the VEMOS system. The following output information is used for comparison:

- Basic emotions variability of the VEMOS system over the time.

- Basic emotions variability of the iMotions software over the time.

Bravais-Pearson Correlation coefficient is then applied to these measures. Variability over the time of the class anger is presented on Figure 7.
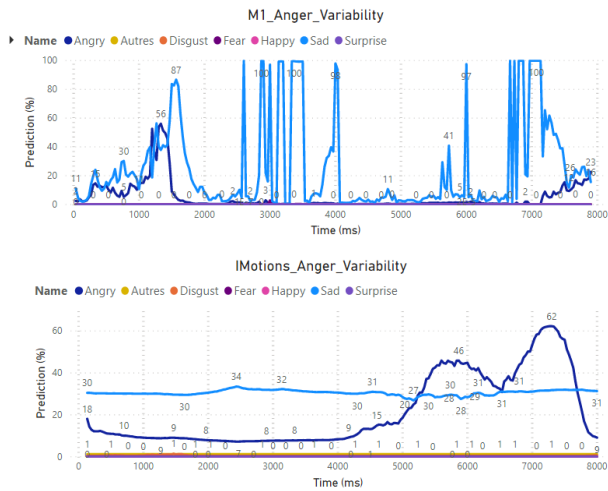

Figure 7: Anger sequence's analysis.

The two systems confirm the presence of the same emotions over the time and recorded correlation coefficient is 0.22 (p<.05). Compared to the initial model, correlation has increased with 0.12 (*p*<.05). But compared to the expected emotion, the two systems show confusion between anger and sadness emotions. It is noteworthy that AFFDEX often confused anger with sadness.

So, we can conclude that VEMOS has this same characteristic. It has an underprediction of anger and overprediction of sadness. Figure 8 presents the analysis results for sadness sequence.
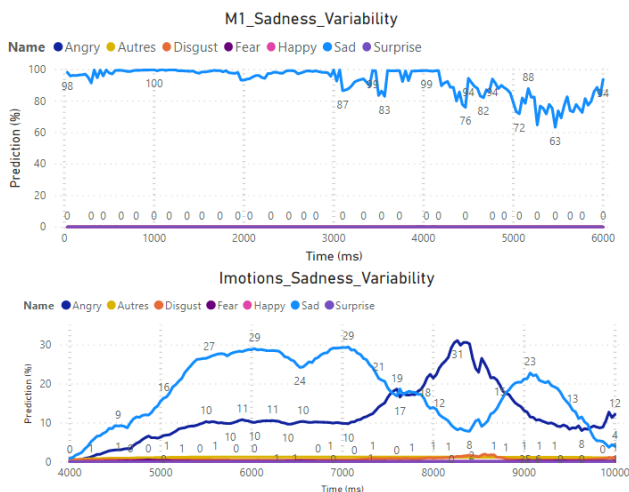
Figure .8: Sadness sequence's analysis.

Sadness is dominating in this sequence and correctly classified by the VEMOS system. Moreover, we recorded satisfying correlation coefficients with values of 0.32 (p<.05) with iMotions Affectiva results.

Although the dominance of sadness shown in the analysis results of iMotions Affectiva, we can see again the same confusion of recognizing anger and sadness. VEMOS recognizes sadness emotion with 100% of occurrence.

For disgust sequence, analysis results are presented in Figure 9.

Disgust is correctly recognized by the VEMOS system and reaches a maximum value of 98% over time. Correlation between the two measures is significant 0.21 (p<.05).

Happy and surprise sequences analysis are presented on Figures 10 and 11.

The corresponding graphs show that these emotions are correctly classified compared to the expected emotion and reach a maximum peak of 100%. We also recorded high significant correlations with 0.81 (p<.05) and 0.54 (p<.05).

Figure 12 shows the analysis results for fear sequence. It shows the dominance of the expected category, and measured correlation coefficient is 0.79 (p<.05).

Our analysis shows that the implemented VEMOS system recognizes correctly joy, fear, surprise, sadness and disgust. But we can notice a rarely confusion of the anger with sadness.
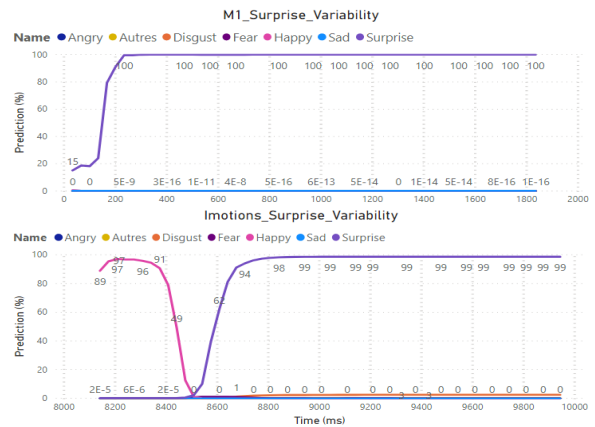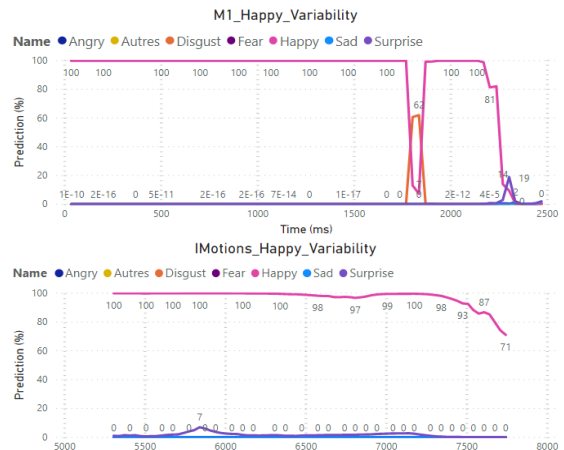

Figure.9: Disgust sequence's analysis.

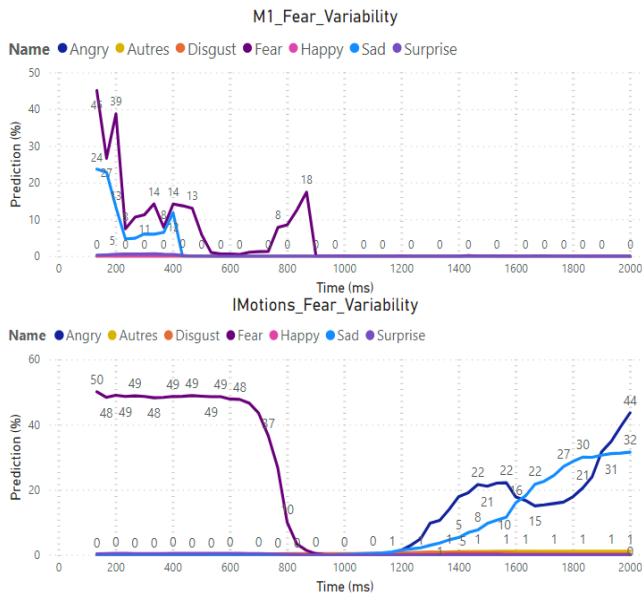Figure.10: Surprise sequence's analysis
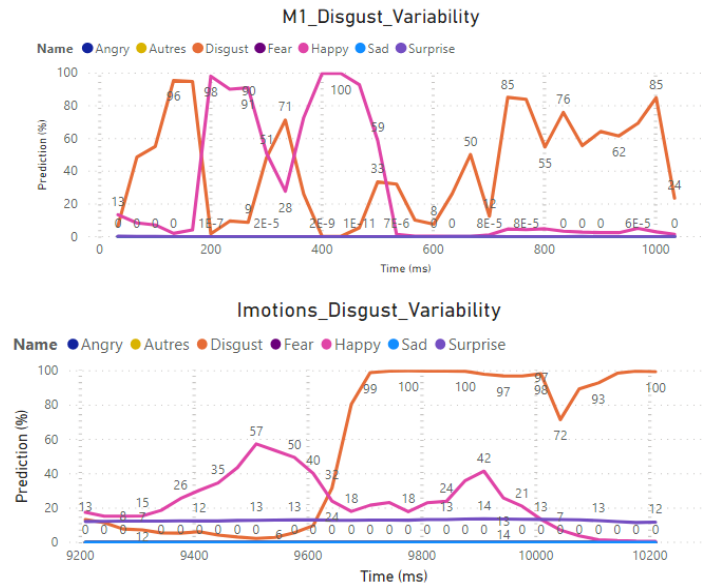
Figure 11: Happy sequence's analysis
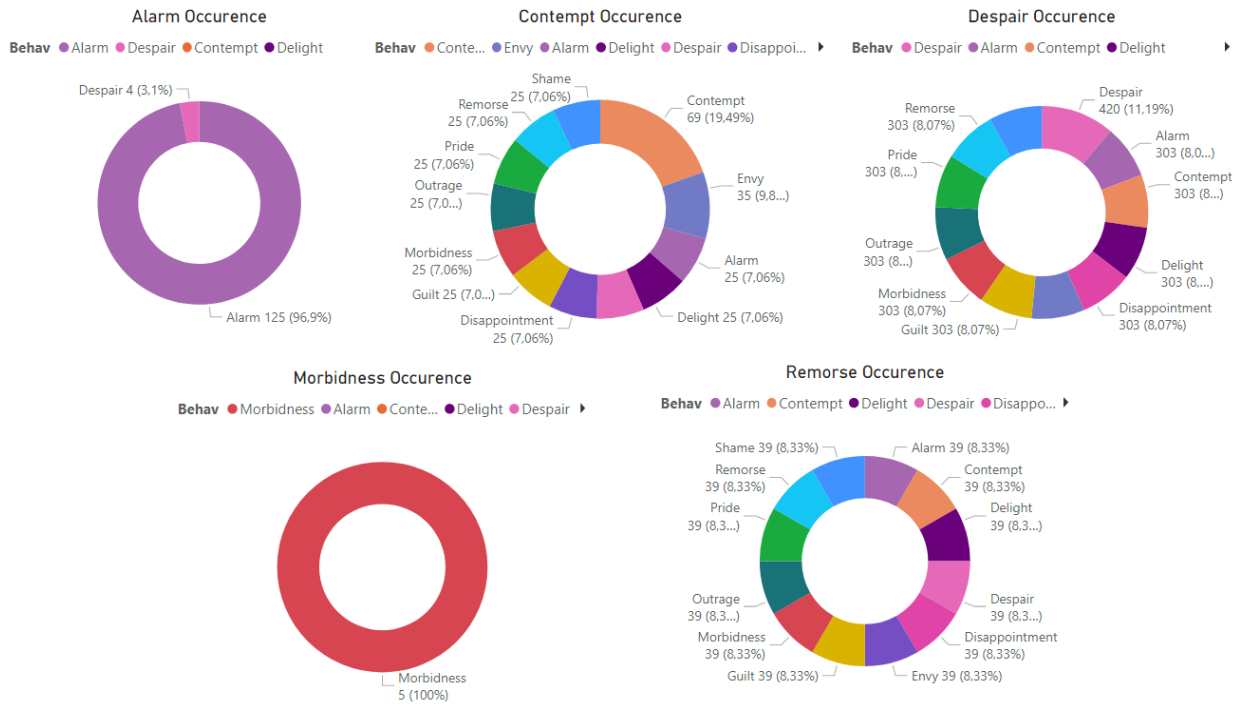

Figure 12: Fear sequence's analysis

Figure 13: Secondary emotions occurrence

3.5. Secondary Emotions correlation results

The authors defined 13 possible combinations based on the 6 primary emotions: anger, disgust, fear, joy, sadness, surprise. Our system was able to analyze the 13 possible combinations.

Tested emotions are: alarm, contempt, despair, morbidness and remorse. As we have mentioned in section 1, the recorded score for each secondary emotion presents the average of the defined combination. We proposed that at a given moment, if one of the basic emotions is zero, the secondary emotion is assigned to 0. In

this case, the detected emotion at this moment will be the primary detected emotion.

To validate and test the performance of the VEMOS system, secondary emotions sequences of the corpus 2 recorded by the actor, are analyzed.

For each sequence, the analysis results are presented on the Power BI Desktop based on the following information:

- Basic emotions variability from the challenging system (VEMOS).

- Basic emotions variability from the iMotions Affectiva software.

- Dominant Secondary emotions occurrence from the challenging system (VEMOS).

Similarly, these sequences were imported and analyzed by the iMotions Affectiva software. The first analysis and validation considered the comparison with the expected emotion.

The graphs describing the dominant secondary emotions for each sequence are presented in Figure 13.

These graphs allow us to compare the dominant emotion with the expected emotion expressed by the actor. Results show relevant information.

We notice that alarm is correctly recognized by the VEMOS system. 96.9% of the extracted frames expressed the expected emotion and only 3% show despair. Contempt is correctly recognized compared to the sequence's label. It dominates with 19.49%.

Figure 14 presents a screenshot of the analyzed video of the actor while expressing contempt emotion. The display of information to be visualized has also been modified and improved. The complete analyzed video of the actor is also available in, [47][1].

Figure 14: Screenshot of the Actor while expressing Contempt[1].

Analyses of despair and morbidness sequences show that the two emotions are correctly recognized and occurred with 11.19% and 100% respectively.

Figure 13 shows also that the remorse is not correctly recognized. Thus, compared to the expected emotion, alarm, contempt, despair and morbidness are correctly recognized by our system.

Then, second analysis and validation considers the comparison with iMotions Affectiva software results. Basically, this comparison concerns primary emotion defined in the combination of each secondary emotion. So, the variability of these primary emotions over the time was recorded for each

---

[1]https://youtu.be/W_0-CuAwDFE

sequence by the two systems and compared by measuring the correlation coefficient.

Consequently, we can see how close the VEMOS results are to that of iMotions Affectiva Software and we can take advantages to confirm again the performance of our system on recognizing primary emotions.

Alarm emotion is a combination of the fear and surprise. Compared to the meaning of fear emotion, Alarm is unexpected and apprehended. It can concern a doubt or lack of self-confidence. Alarm basic emotions variability over the time is analyzed by the two systems. Graphs are presented in the Figure 15.
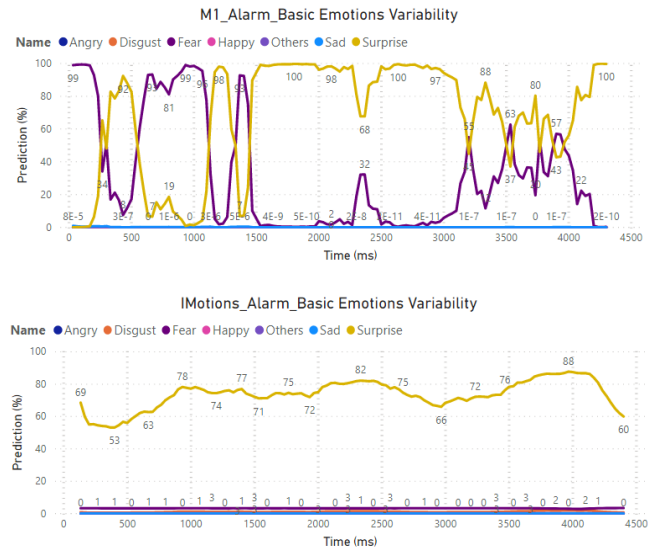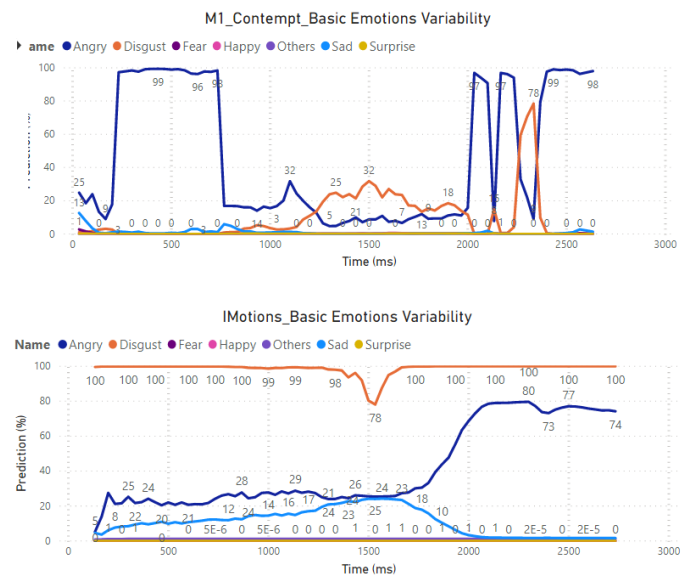

Figure 15: Alarm's sequence analysis.

Figure 16: Contempt's correlation analysis.

Analysis present relevant results compared to iMotions Affectiva. The system shows an appearance of fear and surprise over the time. Obtained correlation coefficients of fear and

surprise are -0.03 and 0.14 respectively. But in this case, iMotions shows only surprise.

Another peculiarity to note is that Affectiva AFFDEX usually confused fear with surprise. This can probably explain the obtained results.

Contempt has the characteristics of anger and disgust. Analyses results are presented in Figure 16.

The two graphs show an appearance of anger and disgust over the time. iMotions analysis also shows the appearance of small variations of sadness. Obtained correlation coefficients of anger and disgust are 0.24 and -0.25 respectively.

Despair emotion combines sadness and fear. This emotion expresses a complete despondency and a loss of hope. We analyzed the sequence with the two systems. Basic emotions variability is recorded and presented on graphs in Figure 17.
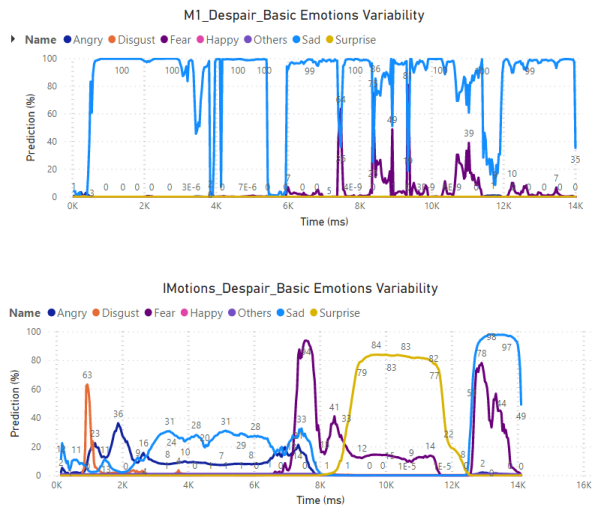

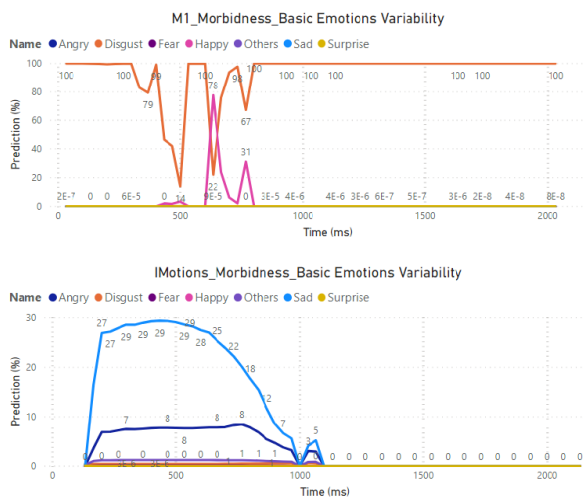
Figure 17: Despair's emotion analysis.

Figure 18: Morbidness's correlation analysis

Our system showed performing results. Sadness and fear appear over the time as it is defined on the Plutchik's emotion

combination. Obtained values of correlation coefficient for fear and sadness with iMotions Affectiva analysis results are 0.65 and 0.11, respectively. The second graph in Figure 17 shows iMotions Affectiva analysis. Results show dominance of sadness and fear. Graph also shows an appearance of the surprise with high values. This appearance can be explained by the fact that iMotions Affectiva confuses surprise and fear. Anger is shown with small variations at the recorded resulting variability of emotions. This can be explained by the fact that iMotions confuses between anger and sadness.

For morbidness, analyzes are presented in Figure 18. According to Plutchik, morbidness combines disgust and joy emotions.

The first graph shows that the two dominant emotions throughout the video are happiness and disgust. It shows that the secondary emotion appears as a peak between 500 and 1000 milliseconds.

iMotions Affectiva analysis shows that the person has expressed sadness and anger. Therefore, the iMotions Affectiva system did not show the expected result. Obtained values of correlation coefficient for disgust and happiness are -0.35 and -0.02 respectively. Figure 19 presents analysis results for remorse emotion.
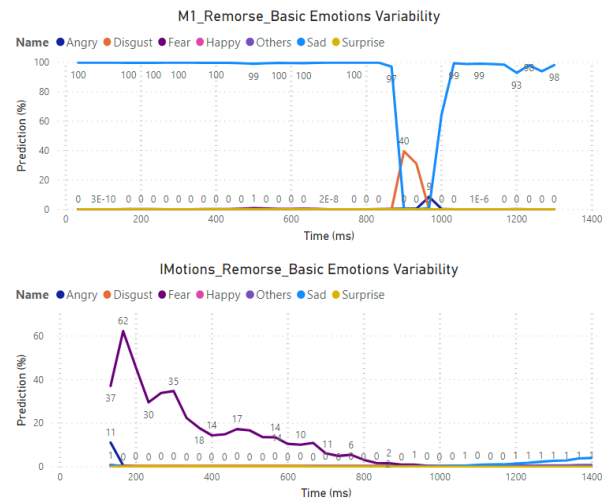

Figure 19: Remorse's correlation analysis.

In the case of the remorse sequence, only sadness has occurred with high scores. The second emotion is disgust, as it is defined in the remorse combination. But disgust appears with one peak all over the video and with low score of 39.53%. iMotions Affectiva system assigned the sequence to the fear category. The correlation values of the basic emotions for remorse combination are 0.05 for disgust and 0.016 for sadness.

We can conclude that our system has successfully recognized secondary emotions. First, results were validated by comparing them to the expected emotion expressed by the actor and validated by the psychologist. These results were also validated according to Plutchik's definition of the basic emotions combination.

Analysis results were also compared to those of iMotions Affectiva results. Joy, disgust and surprise were correctly recognized by the two systems.

More efficient performance was obtained by our implemented system VEMOS for the other emotions.

We notice that iMotions Affectiva confuses anger and sadness emotions. Analysis results of the primary emotions sequences have shown that the system has an underprediction of sadness and overprediction of anger. Because of this confusion, the system was not able to recognize correctly despair and contempt emotions that combine one on the confused primary emotion.

VEMOS system presents an underprediction of anger. But, the system recognizes correctly the sadness emotion. Then, results have shown that the implemented system recognizes correctly despair and contempt emotions.

In addition, VEMOS recognizes more efficiently fear and surprise emotions. Besides, Alarm was correctly recognized. But, we noticed that iMotions Affectiva system has an overprediction of surprise and underprediction of fear. This was shown in the analysis results of Alarm sequence that combine fear and surprise emotions. Only surprise was detected.

Finally, morbidness and remorse secondary emotions were not recognized by iMotions Affectiva system. But, VEMOS was able to detect a maximum peak of the corresponding combination.

4. Conclusion

In this study, a video-based emotion analysis system, VEMOS, was implemented. It detected frame by frame images from video. Then, it employed deep learning techniques for basic emotions recognition. Information about emotions occurrence, emotions variability during the time and maximum peak of each emotions are detected and presented in Power BI Report. Also, an analyzed video showing the variation of pertinent information during the time is obtained.

Then, the system was validated and challenged by comparing correlation with the iMotions Affectiva software on labeled sequences. These sequences were recorded by an actor and validated by a psychologist.

Our validation studies reveal that the implemented system has the potential to recognize basic emotions expressed by individuals. Then, we obtained significant correlations between the implemented system and Affectiva measures. More efficiently than Affectiva, VEMOS system recognizes successfully 5 secondary emotions defined by Plutchik.

With this efficient performance, VEMOS can be exploited in different application. It can help customs officers, known as detection officer, to unveil travelers with illegal items or goods in the customs airport. It is also able to unveil wanted and terrorist people in airport. So, detection officer can stop them from entering and leaving the country. Furthermore, judge can resort to VEMOS to reveal criminal people. Thus, it can assist to know the truth and change the verdict, from a criminal to an innocent.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] S. Masmoudi, D. Yun Dai, N. Abdelmajid, Attention, Representation, and Human Performance : Integration of Cognition, Emotion, and Motivation, Psychology Press, 2012, doi:10.4324/9780203325988.

[2] A. Freitas-Magalhaes, The psychology of emotions: The allure of human face., Uni. Fernando Pessoa Press, 2007.

[3] P. Ekman, Lie Catching and Microexpressions, Oxford University Press, Jan. 2021.

[4] Q. Wu, X. Shen, X. Fu, "The Machine Knows What You Are Hiding: An Automatic Micro-expression Recognition System," in: D'Mello, S., Graesser, A., Schuller, B., and Martin, J.-C., eds., in Affective Computing and Intelligent Interaction, Springer, Berlin, Heidelberg: 152–162, 2011, doi:10.1007/978-3-642-24571-8_16.

[5] P Ekman, WV Friesen, W. Friesen, J Hager, Facial action coding system: A technique for the measurement of facial movement – ScienceOpen, 1978.

[6] P. Ekman, "Facial expression and emotion," American Psychologist, **48**(4), 384–392, 1993, doi:10.1037/0003-066X.48.4.384.

[7] D. Matsumoto, H.C. Hwang, "Microexpressions Differentiate Truths From Lies About Future Malicious Intent," Frontiers in Psychology, **9**, 2018, doi:10.3389/fpsyg.2018.02545.

[8] J.-W. Tan, S. Walter, A. Scheck, D. Hrabal, H. Hoffmann, H. Kessler, H.C. Traue, "Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating various emotions," Journal of Ambient Intelligence and Humanized Computing, **3**(1), 3–10, 2012, doi:10.1007/s12652-011-0084-9.

[9] P. Ekman, J.C. Hager, W.V. Friesen, Facial action coding system: the manual, Research Nexus, Salt Lake City, 2002.

[10] M.S. Hossain, G. Muhammad, "Emotion Recognition Using Deep Learning Approach from Audio-Visual Emotional Big Data," Information Fusion, **49**, 2018, doi:10.1016/j.inffus.2018.09.008.

[11] G. Yolcu, I. Oztel, S. Kazan, C. Oz, F. Bunyak, "Deep learning-based face analysis system for monitoring customer interest," J. Ambient Intell. Humaniz. Comput., 2020, doi:10.1007/S12652-019-01310-5.

[12] A. Krause, Picture Processing System by Computer Complex and Recognition of Human Faces, PhD, Kyoto University, Kyoto, 1973.

[13] B. Reddy, Y.-H. Kim, S. Yun, J. Jang, S. Hong, K. Nasrollahi, C. Distante, G. Hua, A. Cavallaro, T.B. Moeslund, S. Battiato, "End to End Deep Learning for Single Step Real-Time Facial Expression Recognition," in Video Analytics. Face and Facial Expression Recognition and Audience Measurement, Springer International Publishing, Cham: 88–97, 2017, doi:10.1007/978-3-319-56687-0_8.

[14] A. Sun, Y.J. Li, Y.M. Huang, Q. Li, "Using facial expression to detect emotion in e-learning system: A deep learning method," in Emerging Technologies for Education - 2nd International Symposium, SETE 2017, Held in Conjunction with ICWL 2017, Revised Selected Papers, Springer Verlag: 446–455, 2017, doi:10.1007/978-3-319-71084-6_52.

[15] H. Sikkandar, R. Thiyagarajan, "Deep learning based facial expression recognition using improved Cat Swarm Optimization," Journal of Ambient Intelligence and Humanized Computing, 2020, doi:10.1007/s12652-020-02463-4.

[16] R. Ranjith, K. Mala, S. Nidhyananthan, "3D Facial Expression Recognition Using Multi-channel Deep Learning Framework," Circuits, Systems, and Signal Processing, **39**, 2020, doi:10.1007/s00034-019-01144-8.

[17] S. Loussaief, A. Abdelkrim, "Deep learning vs. bag of features in machine learning for image classification," in 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET), 6–10, 2018, doi:10.1109/ASET.2018.8379825.

[18] J. Nadia, Z. Sehla, A. Afef, "Deep Neural Networks for Facial expressions recognition System," International Conference on Smart Sensing and Artifical Intelligence (ICS2'AI), 2020.

[19] Customer Satisfaction Recognition Based on Facial Expression and Machine Learning Techniques, Journal, Jan. 2021.

[20] N. Jmour, S. Zayen, A. Abdelkrim, "Convolutional neural networks for image classification," in 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET), 397–402, 2018, doi:10.1109/ASET.2018.8379889.

[21] Deep Learning Model for A Driver Assistance System to Increase Visibility on A Foggy Road, Journal, Jan. 2021.

[22] M. Edwards, E. Stewart, R. Palermo, S. Lah, "Facial emotion perception in patients with epilepsy: A systematic review with meta-analysis," Neuroscience and Biobehavioral Reviews, **83**, 212–225, 2017, doi:10.1016/j.neubiorev.2017.10.013.

[23] S.H. Lee, "Facial Data Visualization for Improved Deep Learning Based Emotion Recognition," Journal of Information Science Theory and Practice, **7**(2), 32–39, 2019, doi:10.1633/JISTAP.2019.7.2.3.

[24] W. Medhat, A. Hassan, H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams Engineering Journal, **5**(4), 1093–1113, 2014, doi:10.1016/j.asej.2014.04.011.

[25] Kim, "Facial expression monitoring system for predicting patient's sudden movement during radiotherapy using deep learning," Journal of Applied Clinical Medical Physics, 2020.

[26] A.M. Barreto, Application of Facial Expression Studies on the Field of Marketing, Jan. 2021.

[27] iMotions Biosensor Software Platform - Unpack Human Behavior, Imotions, Jan. 2021.

[28] D. McDuff, R. el Kaliouby, T. Senechal, M. Amr, J.F. Cohn, R. Picard, "Affectiva-MIT Facial Expression Dataset (AM-FED): Naturalistic and Spontaneous Facial Expressions Collected 'In-the-Wild,'" in 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 881–888, 2013, doi:10.1109/CVPRW.2013.130.

[29] Noldus | Innovative solutions for behavioral research., Noldus | Innovative Solutions for Behavioral Research., Jan. 2021.

[30] pwtempuser, EmoVu, ProgrammableWeb, 2014.

[31] H.M. Fayek, M. Lech, L. Cavedon, "Evaluating deep learning architectures for Speech Emotion Recognition," Neural Networks: The Official Journal of the International Neural Network Society, **92**, 60–68, 2017, doi:10.1016/j.neunet.2017.02.013.

[32] P. Barros, D. Jirak, C. Weber, S. Wermter, "Multimodal emotional state recognition using sequence-dependent deep hierarchical features," Neural Networks, **72**, 140–151, 2015, doi:10.1016/j.neunet.2015.09.009.

[33] L. Bozhkov, P. Koprinkova-Hristova, P. Georgieva, "Learning to decode human emotions with Echo State Networks," Neural Networks, **78**, 112–119, 2016, doi:10.1016/j.neunet.2015.07.005.

[34] J. Kim, B. Kim, P.P. Roy, D. Jeong, "Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure," IEEE Access, **7**, 41273–41285, 2019, doi:10.1109/ACCESS.2019.2907327.

[35] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, "A Comprehensive Survey on Transfer Learning," Proceedings of the IEEE, **109**(1), 43–76, 2021, doi:10.1109/JPROC.2020.3004555.

[36] A. Krizhevsky, I. Sutskever, G.E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, 2012.

[37] P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 94–101, 2010, doi:10.1109/CVPRW.2010.5543262.

[38] R. Plutchik, Chapter 1 - A GENERAL PSYCHOEVOLUTIONARY THEORY OF EMOTION, Academic Press: 3–33, 1980, doi:10.1016/B978-0-12-558701-3.50007-7.

[39] R. Plutchik, Chapter 1 - MEASURING EMOTIONS AND THEIR DERIVATIVES, Academic Press: 1–35, 1989, doi:10.1016/B978-0-12-558704-4.50007-4.

[40] R. Plutchik, "The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," American Scientist, **89**(4), 344–350, 2001.

[41] Visualisation des données | Microsoft Power BI, Jan. 2021.

[42] CBS News, Carlos Ghosn clame son "innocence" dans une vidéo, 2019.

[43] Facial Expression Analysis: The Complete Pocket Guide, Imotions, Jan. 2021.

[44] In Lab Biometric Solution, Affectiva, Jan. 2021.

[45] L. Kulke, D. Feyerabend, A. Schacht, "A Comparison of the Affectiva iMotions Facial Expression Analysis Software With EMG for Identifying Facial Expressions of Emotion," Frontiers in Psychology, **11**, 2020, doi:10.3389/fpsyg.2020.00329.

[46] S. Stöckli, M. Schulte-Mecklenbeck, S. Borer, A.C. Samson, "Facial expression analysis with AFFDEX and FACET: A validation study," Behavior Research Methods, **50**(4), 1446–1460, 2018, doi:10.3758/s13428-017-0996-1.

[47] Nadia Jmour, VEMOS the new video based emotion analysis system: Actor while expressing Alarm and Contempt, 2021.