

# Performance Comparison of the KNN and SVM Classification Algorithms in the Emotion Detection System EMOTICA

Kone Chaka<sup>1\*</sup>, Nhan Le-Thanh<sup>2</sup>, Remi Flamary<sup>3</sup> and Cecile Belleudy<sup>1</sup>

<sup>1</sup>Universit e C te d'Azur, Laboratoire LEAT CNRS UMR 7248, Sophia-Antipolis, France

<sup>2</sup>Universit e C te d'Azur, Laboratoire I3S CNRS UMR 7271, Sophia-Antipolis, France

<sup>3</sup>Universit e C te d'Azur, Observatoire de la C te d'Azur, CNRS, Laboratoire Lagrange, Nice, France

## Abstract

Emotica (EMOTIon CAPture) system is a multimodal emotion recognition system that uses physiological signals. A DLF (Decision Level Fusion) approach with a voting method is used in this system to merge monomodal decisions for a multimodal detection. In this document, on the one hand, we describe how from a physiological signal Emotica can detect an emotional activity and distinguish one emotional activity from others. On the other hand, we present a study about two classification algorithms, KNN and SVM. These algorithms have been implemented on the Emotica system in order to see which one is the best. The experiments show that KNN and SVM allow a high accuracy in emotion recognition, but SVM is more accurate than KNN on the data that was used. Indeed, we obtain a recognition rate of 81.69% and 84% respectively with KNN and SVM algorithms under certain conditions.

**Keywords:** Classification; KNN; SVM; Multimodal recognition of emotions; Physiological signals; Emotica

## Introduction

Communication specialists agree that more than 70% of communications are non-verbal. Nonverbal communications rely on behaviors, gestures, facial expressions, and the intensity of a person's voice. Indeed, as well as non-verbal communication, emotions enable us to communicate with our environment. It is from this idea and the evolution of new technologies benefiting people's health that affective computing was born. Affective computing is the study of interactions between technology and emotion to give machines the ability to understand, to interpret our emotions or even to express emotions. Affective computing offers many advantages such as the battle against depression, interactive games, E-Learning, etc.

A lot of research based on video applications or speech analysis and physiological signals analysis has emerged to analyze emotions, with the aim, amongst other, to provide a real-time, aggregate view of users feelings, and in general to identify customer dissatisfaction. The solution proposed in this paper targets the healthcare domain in that it monitors biological signals, but in a non-intrusive manner for the benefit of patients. In the future, emotion detection tests will be very challenging because they constitute a key point to analyze the impact of all medical treatments, and the resulting device market will probably be substantial.

Our goal is to collect the physiological signals of a person under different real life conditions to automatically detect emotions. This paper's contribution is a system design for emotion detection, which includes signal processing, feature extraction, and decision. To design the emotion detection system:

- We firstly address in this paper proposals for signal processing algorithms designed to analyze physiological signals. Our methodology to recognize an emotional activity is based on peak detection of the analyzed signals. This peak detection is done after having acquired and filtered the physiological signal. The detection consists in calculating the gradient of the signal and then, finding the sign changes in the gradient.
- After extracting and isolating a peak, in order to find typical characteristics for each emotion, we calculate seven essential

parameters, thus constituting a feature vector for each detected peak. These seven parameters make it possible to properly characterize a given peak, assuming that each measured signal is generated by a Gaussian process, with independent and identically distributed samples. The two functions that can be used to characterize a raw physiological signal are mean and variance. The average of the filtered signal, the average of the filtered signal gradient's as well as the variance of the filtered signal are calculated to evaluate the signal trend. The width and intensity of a signal can also provide relevant information.

- Another contribution of this paper is on detection algorithms. We propose to compare the performances of 2 classification algorithms in terms of their accuracy in the emotion recognition and their rapidity.
- Finally, we have developed a multimodal approach for the recognition of emotions. This approach is based on a method of merging monomodal decisions (an emotional vector calculated for each type of signal). We have defined our multimodal approach with a weighted voting mechanism, where the weights were determined.

The reminder of this paper is structured as follows. In Section 1, a brief state of the art on multimodal emotion recognition and how emotions can be detected with physiological signals is described. In Section 2, the Emotica system is explained briefly. In Section 3, we present the KNN and SVM algorithms that we have implemented in the Emotica system. In Section 4, a comparison is made in terms of performance between the results obtained with KNN and SVM. Furthermore, a comparison is made between our results and those

**\*Corresponding author:** Kone Chaka, Universit e C te d'Azur, Laboratoire LEAT CNRS UMR 7248, Sophia-Antipolis, France, Tel: 0492942804; E-mail: chaka.kone@unice.fr

**Received** November 29, 2017; **Accepted** February 08, 2018; **Published** February 13, 2018

**Citation:** Kone C, Le-Thanh N, Flamary R, Belleudy C (2018) Performance Comparison of the KNN and SVM Classification Algorithms in the Emotion Detection System EMOTICA. Int J Sens Netw Data Commun 7: 153. doi: 10.4172/2090-4886.1000153

**Copyright:**   2018 Kone C, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

obtained in the state of the art. Finally, our conclusion and future work are reported in Section 5.

## State of the Art

### Monomodal versus multimodal

Several physiological signals can be used for emotion detection. The monomodal approach consists in taking the physiological signals from a single sensor and then passing through two essential phases for the emotion recognition. The first is the learning phase, which aims at designing a learning base by applying different signal processing techniques. In the second phase, the learning base constructed during the first phase is used to detect instantly and automatically an emotional activity. Human emotions by nature are multimodal. Indeed, when a person is suddenly seized with fear, not only his heartbeat accelerates, he also trembles, sweats a little more, blushes sometimes (depending on the origin). Moreover, his voice and his gestures are affected. In order to get closer to the nature of the emotion, it is necessary to propose a multimodal system of emotion recognition. In order to interpret the emotions of the people which are around us, we use all the sources of information at our disposal (the pitch of their voices, their sweats, have they become pale? Do they tremble?...). Thus, in addition to have better recognition rates, the multimodal approach allows more robustness when one of these signals is acquired in a noisy environment [1]. In our approach, we chose multi-modality using several sources of information at the physiological level. That is to say we will use different types of physiological signals in order to propose a multimodal system of emotion recognition. To propose this system, first of all, it is necessary to merge the information of the various physiological signals. To do this, three techniques (signals level fusion, features level fusion and decisions level fusion) have been proposed in the literature for fusion. The interested reader for more information on these techniques will be able to refer to these few works [2-4]. At feature level, there is a semi-supervised method which outperforms the other state of the art methods. It is called the global-label-consistent classifier [5]. The voting method is the most used at decision level fusion and the concatenation method is used at signal level and feature level fusion. In this paper, we use decision level fusion with weighted voting technique. In the future, the quantity of data will be important, so it will be interesting to test feature level fusion with the global label consistent classifier method, which provides good classification rate and is not too time consuming.

### Emotion detection with physiological signals

Physiological signals such as instantaneous heart rate, heart frequency variability and sweating are some important parameters to monitor a person's health status and to get information on the presence of cardiovascular dysfunction [6,7]. These markers reflect the state of the Autonomic Nervous System (ANS), and are observed to highlight an increase of the sympathetic activity responsible for controlling certain unconscious activities of the organism (such as heart rhythm or smooth muscle contraction) or a reduction of the vagal activity [8]. The limbic system responsible for the regulation of emotions, of behaviors calling for long-term memory includes cerebral structures, namely the amygdala, the hypothalamus and the limbic cortex. The latter two, have functions related to memory and control of the spatial orientation. The cerebral amygdala, the nucleus of which is located in the temporal lobe, is responsible for analyzing the degree of threat and emotional significance of all the information both internal and external. It responds to visual, audible, tactile, gustatory and olfactory senses. It also allows the storage of emotional traumas without conscious control and is linked to phobias [9]. The ANS, being a part of the nervous system

responsible for the functions that are under involuntary control, is involved in the functions of the organism such as thermoregulation, or even regulation of blood pressure by controlling the vasomotor center, which controls the diameter of the blood vessels to maintain an arterial pressure at a tolerable level. Physiological signals acquired through the use of less and less invasive sensors provide information on the state of the autonomic nervous system. Therefore, the acquisition of such signals allows the monitoring of the emotional and behavioral state of the human being.

In our work, we will use four physiological signals (Galvanic skin response GSR, Blood Volume Pulse BVP, Electromyogram activity EMG and Respiratory Volume VR). Healey [10] used these signals during her thesis work to set up a database of physiological signals for eight emotions.

Table 1 lists some studies on the detection of emotions using physiological signals.

### Emotica System

The Emotica system [4] allows an automatic and instantaneous detection of emotions using physiological signals based on the monomodal and the multimodal approach. This system is based on two phases, the learning and detection phases. The learning phase (Figure 1) is composed of the successive steps of splitting, filtering, and features extraction which is designed to create a basis for learning. The second phase is the automatic and instantaneous detection of emotions. The different steps in this detection phase (Figure 2) are the same as for the first phase, except that we no longer go through the splitting step. In the Emotica system, each emotion  $e$  can be written as a linear combination of the 8 basic emotions ( $B$ ). ( $B$ )= (No emotion, anger, hate, grief, love, romantic love, joy and reverence).

$$(e) = \alpha_1 * No\ Emotion + \alpha_2 * Anger + \dots + \alpha_8 * Reverence \quad (1)$$

$$(e) = (\alpha_1 \alpha_2 \dots \alpha_8)_B \quad (2)$$

Where  $(\alpha_1 \alpha_2 \dots \alpha_8)$  are the probabilities of the feature vector extracted belonging to each emotional class of our base. The detection phase in the Emotica system aims at constructing an emotional vector for each emotional activity detected. To do so, it is required to determine the different values of  $\alpha_i$ .

### Learning phase

This phase consists in four steps (signal splitting, filtering, feature extraction, creation of the basis for learning) in order to provide a learning base which will then be used in the detection phase for the automatic detection of emotions. Figure 1 shows the synoptic of the learning phase.

**Signal splitting:** In this step, after having acquired the physiological signal, we isolate the part of the signal corresponding to a given emotion because we have information on the period in which each of the eight emotions is expressed. Therefore, this step divides the input signal into eight portions of signal corresponding to eight emotions.

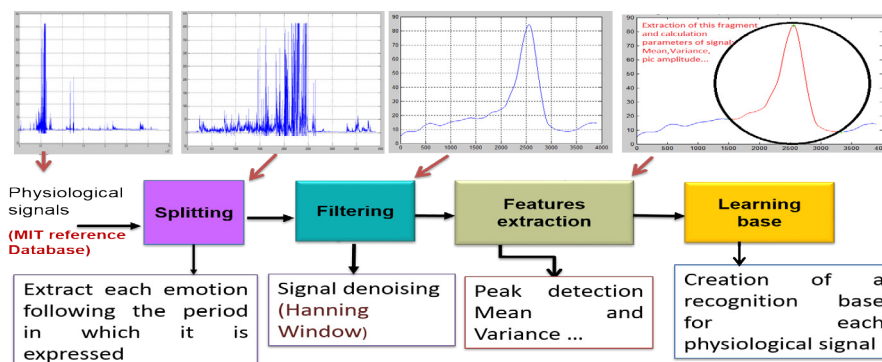
**Filtering:** After having isolated the signal, we filter it to remove the noise of the useful signal, which will facilitate the extraction of the features. We have opted for the convolution method for filtering, which consists in convoluting the signal in the spatial domain with different filters (for which we chose the Hanning filter). This method is less computationally expensive in calculations.

**Extraction of features:** For each isolated and filtered signal,

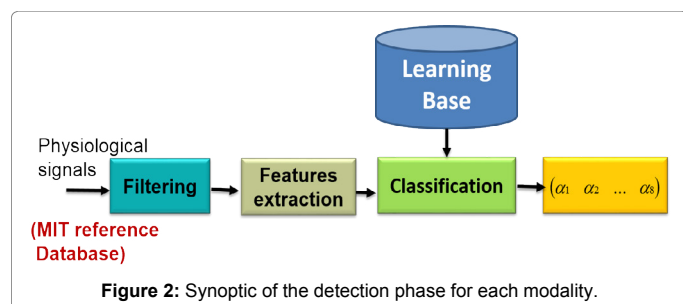
Authors information	Signals	Part.	Feat.	Sel.	Classifiers	Target	Result (%)
[11]	C, E	24	5	B-W	HMM, Viterbi	frustration/not	63%
[12]	C, E, R, M	1	11	Fisher	QuadC, LinC	3 emotions	87%-
							75%
						anger/peacefulness	99
						2 arousal levels	84
						2 valence levels	66
[13]	C, E, R, M	1	12	SFS	kNN	4 stress levels	87
[14]	C, B	10	12		NN, SVM	2 valence levels	62
[15]	C, E, M, S	1	18		FL, RT	3 anxiety levels	59-91
[16]	C, E, R, M, S	1	30	reg, LDA	kNN	5 emotions	24
[17]	C, E, B	12	18		SVM	5 emotions	42
						3 emotions	67
[18]	C, E, S, and others	32	?		kNN, NN	2 fear levels	92
[19]	C, R	15	18	ANOVA, PCA	LDA	2 emotions	65
						4 emotions	72-83
[20]	C, E, M	3	54		SVM	3* 2 levels	85-80
							84
[21]	C, E	40	28		regression model	5 emotions	63-64
[22]	C, E, M, S	5	18		FL, RT	anxiety scale?	57-95
[23]	C, E, R, M, S	24	4*50		LDA, GMM	2 levels of stress	94-89
[24]	C, E, R, M, O	34	23		ANOVA PDA	fear, sadness, neutral	69-85
[25]	C, E, M	6	54		SVM	3*2 levels	81
[26]	C, E, M, S	6	?		SVM, QV-learning	3* 2 levels	83
[27]	M	1	12	DWT	NN, TM	4 emotions	75
[28]	C, E, R, S	1	225	SFS, Fisher	LDA, kNN, NN	4 emotions	90

Signals: C : cardiovascular activity; E : electrodermal activity; R : respiration; M : electromyogram; S : skin temperature; O : Expiratory pCO<sub>2</sub>. Classifiers: HMM: Hidden Markov Model; RT: Regression Tree; NN: Artificial Neural Network; SVM: Support Vector Machine; LDA: (Fisher) Linear Discriminant Analysis; kNN: k-Nearest Neighbors; FL: Fuzzy Logic System; TM: Template Matching classifier; QuadC: Quadratic classifier; LinC: Linear classifier; Viterbi: Viterbi decoder  
Selection: B-W: Baum-Welch re-estimation algorithm; PCA: Principal Component Analysis; SFS: Sequential Forward Selection; ANOVA: Analysis of Variance; DWT: Discrete Wavelet Transform; Fisher: Fisher projection; PDA: Predictive Discriminant Analysis. part.: number of participant; feat.: number of feature; result: classification rate; sel.: selection/reduction

**Table 1:** Some studies on systems of classifications of affective states using physiological signals.



**Figure 1:** Synoptic of the learning phase.



we proceed to the detection of peaks, which is done by calculating the gradient of the signal and then, finding the sign changes in the gradient, because it is rare to find points in discrete signals where the gradient is zero. A maximum is shown by the passage of a positive gradient to a negative gradient, a minimum by the passage of a negative gradient to a positive gradient. To detect and isolate a peak, our method should detect a minimum followed by a maximum followed by a minimum. Once a peak is isolated, we calculate a feature vector composed of five features that are: the mean, the variance, the mean of the filtered signal, the variance of the filtered signal, and the amplitude of the peak.

**Creation of the learning base:** After extraction of the features vectors, we create a learning base for each modality. Thus, at the end of this step, we get 4 learning bases. The conception of this learning base is well explained in the section 3.3 and the section 5.

## Detection phase

This phase consists of two steps. The first step consists in the features extraction, requiring the same steps as in the learning phase, without going through the splitting step since in this phase; there is no information beforehand on the period at which every emotion is expressed. The remainder of our process will be based on this features extraction step. It is necessary to detect a peak (an emotional activity) before pass to classification step. Indeed, thanks to this condition on the necessity of detecting an emotional activity, our method allows an instantaneous recognition of emotions.

The second step is classification, the purpose of which is to predict the emotional class of the features vector extracted using our learning base, which was developed in the learning phase. We have developed and implemented 2 classification algorithms which are introduced and explained in Section 4.

## The database used in this study

In addition to the hardware architecture of the sensing system composed of sensors, microcontroller, RF module and battery, the sensing system has a test protocol for the induction of emotions. A simple and effective protocol to stimulate the person on whom the acquisition is made is required. This protocol stimulates the person on whom the data is acquired to feel and express the different emotions and thus, it is of considerable help for the construction of a physiological signal database. The signal acquisition method is essential as it strongly influences the efficiency of the detection. Indeed, the quality of the signals (thus the detection) is not the same when the acquired data is simulated or real, hence the problem of the definition of an adequate and effective protocol. These protocols call upon methods of induction of emotions to acquire a database in which the influence of each emotional state has been

faithfully reflected. Several protocols for induction of emotional states exist in the literature. As an example, we can cite the one proposed by Clynes Manfred [29] which consists of stimulating the subject with different musical melodies and/or images to produce eight emotional states. This method, called the Sentic Cycle, is defended by its initiator as a method for training to feel each of the emotional states considered. The international emotional picture system (IAPS) developed by Lang et al. [30], is another protocol for induction of emotional states. The IAPS system was adopted by many studies. In particular, it was used by Faiza Abat in his thesis work [31] in view of the simplicity of his manipulation. Finally, another protocol based on films is proposed in the literature for the induction of emotions [film, 2017]. The emotional reactions provoked by the latter are often too weak when we talk about complex emotions [31].

In our work, we used a physiological signal database established by Jennifer Healey during her PhD [10] held at MIT (Massachusetts Institute of Technology). This database serves as a reference for the vast majority of people working on emotion detection. It allows a comparison of different algorithms for emotion recognition on the same data.

During 32 days, an actor on whom 4 physiological signal sensors (EMG, RESpiration, GSR, BVP) sampled at a frequency of 20 Hz were connected, was stimulated to express and feel eight emotions: No-emotion, Anger, Hate, Grief, Love, Romantic Love, Joy and Reverence. Due to problems such as the detachment of an electrode or the malfunction of a sensor, on the 32 days of acquisition, only 20 have been perfectly acquired and are therefore the only data we use.

The Sentic cycle method developed by Manfred Clynes [32] was used, which involves stimulating the actor with music or images so that he feels and expresses a given emotional activity. Every day, the actor sitting on a chair with a backrest and connected to the 4 physiological signal sensors was stimulated to express each emotion for a period of 3 to 4 minutes. The interested reader by this acquisition protocol is invited to refer to these different works for more details [10,29,32]. The Table 2 regroups the images used in the MIT protocol to encourage the actor to express the eight emotions taken into account as well as the description, intensity and valency of these different emotional states.

Thus, as shown in Table 2, each image is chosen according to the emotion that the actor must express. For example, images of children suffering from famine cause grief in the actor. When church or prayer images are presented to him, he would feel the reverence emotion, while images of angry people cause the actor to feel the same angry emotion.

## Classification Algorithms

### K nearest neighbors (KNN)

The KNN [33] method is an intuitive supervised learning method that is easy to implement. In a way, one can say that it is based on this proverb: "Tell me who your friends are and I will tell you who you are". Meaning that a man identifies himself with the people around him. These people forming his network have a more or less considerable impact on his character. By analogy with this, the KNN algorithm consists in calculating the distances between the features vector of a new observation to be classified and the features vectors of the reference signals which form the learning base. The signal to be classified is then assigned to the majority class among the k closest classes according to the chosen distance calculation. There are different types of distance calculation, but, we have



Emotion	Images		Description		Intensity	Valence
No emotion	white paper, typewriter		boredom, job offer		low	neutral
Anger	angry people		desire fight	to	very high	very negative
Hate	injustice, cruelty		passive anger		low	negative
Grief	children suffering, loss of mother		lost, ness	sad-	high	negative
Love	family, mer	sum-	joy, peace		low	positive
Romantic love	dating romantic		excitation, desire		very high	positive
Joy	music		a lot of joy		moderately high	positive
Reverence	church, prayer		calm, peace		very low	neutral

**Table 2:** Images used in the MIT protocol to provoke different emotions in the actor.

opted for the method of calculating the Euclidean distance. Let  $X_p = (x_{p1}, x_{p2}, \dots, x_{pN})$  be a features vector to be classified, N the number of features in this vector and Y the base of reference.

$$d(X_p, Y) = \sqrt{\sum_{i=1}^N (X_{pi} - Y_i)^2} \quad (3)$$

In the KNN method, the coefficients are determined by using the following formula

$$\alpha_i = \frac{k_i}{k} \quad (4)$$

where  $k_i$  represent the number of the i nearest neighbors found.  $\sum_{i=1}^M \alpha_i = 1$  where M represents the number of the emotional class. The difficulty in the KNN algorithm is to find the optimal value of k. The higher k is, the greater the risk of over-learning is, and the boundaries between the emotional classes are less distinct. For this reason, it is necessary to make a validation of the k value in order to find the value which reduces the risk of over-learning and maximizes the emotion recognition rate.

The k nearest neighbors' classification algorithm is quite intuitive and simple to implement. After the implementation of the latter, we have developed and implemented a new algorithm called support vector machine (SVM) which, on the classification systems literature, performs better than the KNN cite. The next paragraph introduces this classification algorithm.

## Support Vectors Machines (SVM)

SVM [34] is a supervised learning method for binary classification. The principle consists in defining a hyper plane separating the 2 classes while maximizing the margin. The latter is the distance between the separation boundary and the support vectors that are the nearest samples. In the case where the data are linearly separable, an optimal separating function is defined (maximizing the margin). Otherwise, the data is projected into a space of larger dimensions where it can be linearly separable.

Let  $x_i, x_j$  be the data of the two classes. The SVM relies on the scalar product of these input vectors  $\langle x_i, x_j \rangle$ . In the case where the data has been projected in a larger space using the transformation,  $\phi: x \rightarrow \phi(x)$ , then, the scalar product becomes: with  $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$  with  $K(x_i, x_j)$  as the kernel function. In our case, we choose the Gaussian as kernel function.

$$K(x_i, x_j) = \exp \frac{-\|x_i - x_j\|^2}{2 * \sigma^2} \quad (5)$$

The implementation of the SVM classical algorithm does not make it possible to go back to the coefficients  $\alpha_i$ . To keep our vectorial representation of emotions, we used the LIBSVM [35] toolbox to retrieve the posteriori probabilities which are the coefficients  $\alpha_i$ . Our work focuses on the multimodal detection of emotions. Two approaches exist to pass from the classical SVM (binary classification) to the multi-class SVM.

**The one against all approach:** The principle of this approach is based on the transformation of the problem to k classes into k binary classifiers. We learn a model to discriminate between the groups  $y = k$  and  $y \neq k$ . Each model must be capable of separating a specific class from the others. The estimates of the posteriori probabilities found for each model give the coefficients  $\alpha_i$  from which the emotional vector is formed.

**The one against one approach:** In this approach, the principle consists in considering all the pairs of possible classes ( $C_k^2$  combinations possibilities) and to constitute as many models as pairs of classes. Let us take an example of 8 classes, given that  $C_8^2 = 28$ , therefore there will be 28 possible combinations. The coefficients  $\alpha_i$  are calculated from the number of duels won by each of the emotional classes.

The difference in performance between these two approaches is not so obvious. However, some contradictory studies have been done on these approaches as in [36] and [37]. Therefore, the use of one approach or another depends on the intended application and its performance depends on the fineness of the models pre-established in the binary classification step. That is, each model should be as efficient as possible in classifying a given class in order to increase the accuracy of the multiclass algorithm. As a result, we have chosen to implement the one against all approach. It is the most obvious approach, which requires fewer models to be defined, so the validation step requires less time and thereafter the detection itself requires less computing time than in the one-to-one approach.

The interested reader is invited to refer to this work [38] for more information's on how the support vector algorithm works.

## Results and Comparisons

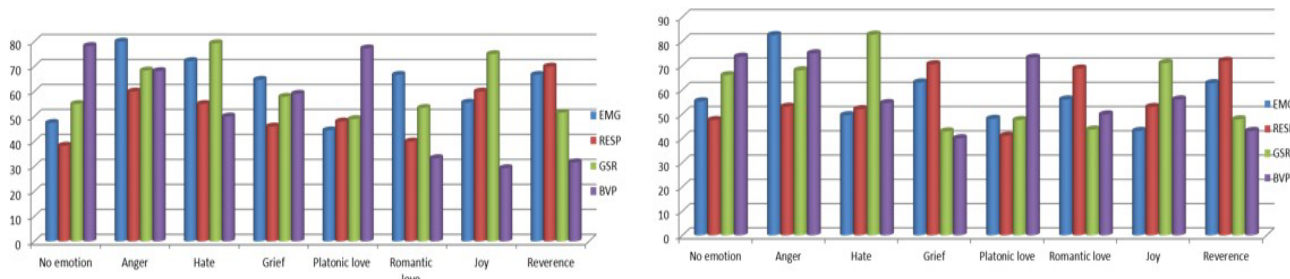
After regrouping all these physiological signals, we used the k

cross validation method with  $k=3$  to validate the hyper-parameters of our classifiers. We have chosen to validate our parameters using the pessimistic method which consists in subdividing the database into three bases: the learning base  $\varepsilon$ , the validation base  $v$  and the test base  $\tau$ . In this pessimistic method, one learns the prediction function on  $\varepsilon$ , validates its parameters on the  $v$  data, and tests the performance on the data  $\tau$ . This method eliminates the bias on the optimistic method because the data on which the performance of the prediction function was estimated was unknown for the function at the time of its implementation and the validation of its parameters. This method is therefore a generalization of the classification, because we can expect the same order of performances when we change the test data.

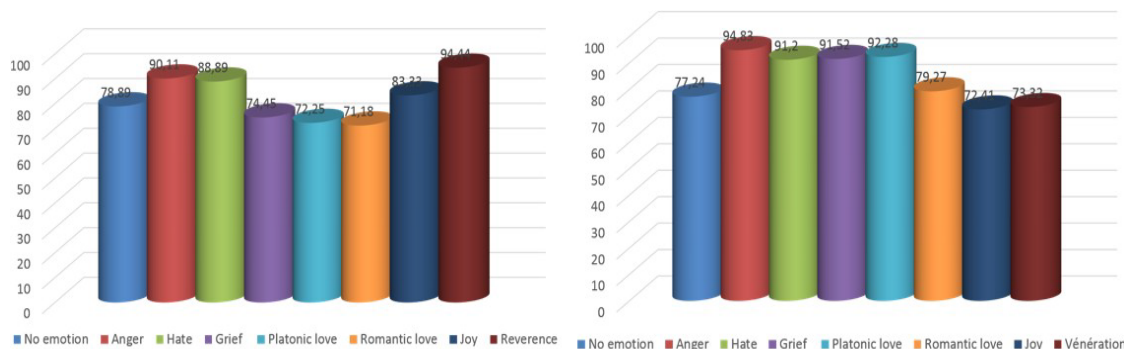
Thus, 84 feature vectors were extracted from the physiological signals for each emotion. Since we have 8 emotions, 672 feature vectors have therefore been extracted for each kind of physiological signal. On these feature vectors,  $33 \times 8 = 266$  (about 40%) were used as a test basis to test our algorithms and the remaining 60% ( $50 \times 8 = 400$  feature vectors) were used as the learning and validation base. 60% of these 400 feature vectors, i.e.  $30 \times 8 = 240$  feature vectors constitute the learning base and finally  $20 \times 8 = 160$  for the validation base. In short, the learning, validation and test bases are respectively composed of 240, 160 and 664 feature vectors.

After subdividing the base of signals provided in three bases: learning base, validation base and test base; we applied the KNN algorithm and the SVM one on this data. The parameter  $k$  has been validated on the validation base. The optimal value of the  $k$  parameter found is  $k = 10$  for the KNN algorithm. Thus, the results obtained by our algorithm of the KNN with the uni-modal approach are shown in the figure on the left of Figure 3. This monomodal approach allows having an emotion recognition rate of 57.24%. As shown in Figure

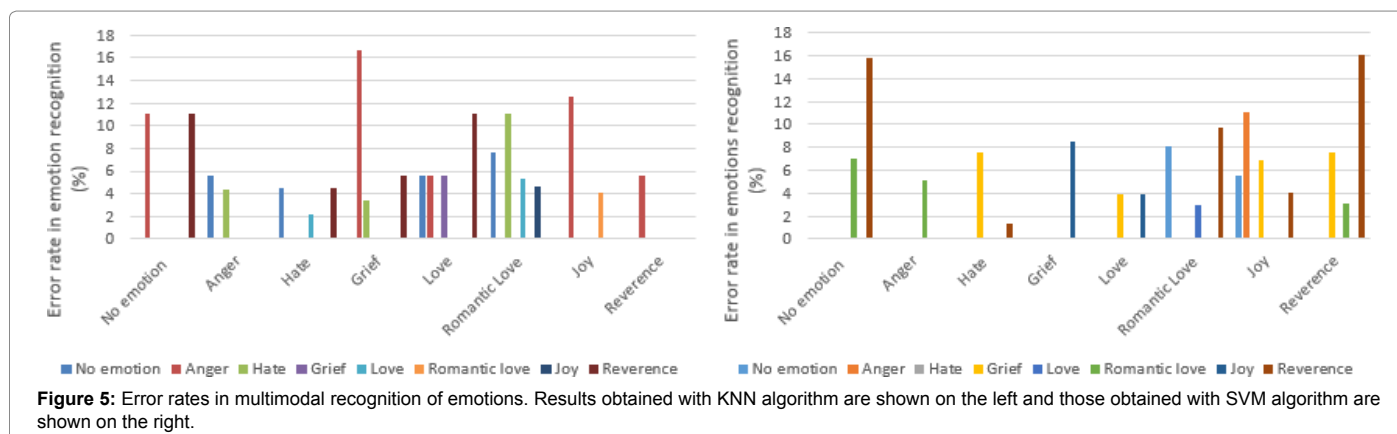
3, some emotions are better detected with some physiological signal sensors than with others. Indeed, we can notice that the emotions “no emotion” and “platonic love” are better detected with the BVP modality, while the GSR modality allows a better detection of the emotions “hate And joy”. The EMG modality meanwhile allows us to better recognize “the emotions” anger and “romantic love”. It is interesting to have general practitioners in a society, however, it is also very important to have some of them who are specialized in different branches of medicine. In the same way, this characteristic of the modalities that allows them to be specialized in the recognition of certain specific emotions is very important because it will allow weighing each of the modalities, depending on whether it is better to detect an emotion or not in the more efficient detection purpose. Thus, we will use this feature in our multimodal approach to considerably improve the recognition rate. Then, we validated the parameters  $c$  and allowing adjusting to its maximum the Gaussian for the classification of each emotional class for the SVM algorithm. We obtained as many optimal values of  $c$  and as emotional classes. The results obtained with the SVM algorithm are shown in the figure on the right of Figure 3. For this monomodal approach, SVM allows to have an average recognition rate of 57.87%. From these findings, two remarks can be made. The first shows that the specialization characteristic of each physiological signal sensor in the recognition of certain specific emotions is confirmed with the exception that the EMG sensor is no longer the specialist for emotion “romantic love”. As a second observation, we see that the recognition rate obtained by using the SVM classifier is slightly higher than that obtained with the KNN classifier. However, the difference is not very large, so, we can say that in the monomodal approach the recognition rate obtained with the two classifiers is almost identical.



**Figure 3:** Monomodal recognition of emotions. Results obtained with KNN algorithm are shown on the left and those obtained with SVM algorithm are shown on the right.



**Figure 4:** Multimodal recognition of emotions. Results obtained with KNN algorithm are shown on the left and those obtained with SVM algorithm are shown on the right.



For the multimodal approach, once the parameters ( $k$ ,  $c$ ,  $\gamma$ ) are validated in the Monomodal approach, on the test basis, we have verified the reliability of our two classification algorithms. The results obtained for the KNN and SVM classification algorithms are presented in Figure 4.

On the basis of these results, we can conclude that this multimodal approach has considerably improved the emotion recognition rate in relation to the monomodal approach. Moreover, our algorithms allow to detect each emotional activity and to distinguish it from the others with very good rates with minimum rate at about 71%. With the KNN algorithm, we pass from an average recognition rate of 57.24% with the monomodal approach to a recognition rate of 81.69% for the multimodal approach. With the SVM algorithm, we obtained an average recognition rate of 84%. In the monomodal approach, and with the two classification algorithms, we obtain fairly close recognition rates. However, in the multimodal approach, the SVM algorithm has a better recognition rate. We can say that on this data the SVM is more efficient than the KNN because the recognition rate is improved by more than 2% which is especially not negligible within the framework of emotions detection.

Figure 5 allows to study the error rates on the recognition of the emotions for our two algorithms.

We grouped the emotions in 3 classes according to their valences:

- Neutral class: No emotion, reverence
- Negative class: anger, hate, grief
- Positive class: platonic love, romantic love, joy

Taking the misclassification rates for each emotion and summing these different rates on each class of valences, we find the error rates grouped in Table 3.

Table 4 makes it possible to make a comparison between our results and the different results of the state of the art methods, which allow an instantaneous detection of emotions. These results show that the method that we proposed with each of the two classification algorithms allows to have better recognition rates than the methods cited in the state of the art. Indeed, the recognition rate is improved by more than 10%.

Our emotions detection algorithms have been designed in such a way so that they are simple, in order to reduce complexity, since we plan to implement them in mobile devices like smartphones. The

KNN algorithms			
	Neutral	Negative	Positive
Neutral class	11	16.67	0
Negative class	19.9896	24.34	2.2204
Positive class	24.36	34.76	14.12
SVM algorithms			
	Neutral	Negative	Positive
Neutral class	31.77	7.58	10.09
Negative class	1.28	7.52	13.65
Positive class	27.41	21.82	6.81

**Table 3:** Error rates following the emotional valences.

Methods	Recognition rate %
Method of Kim [39]	61.2
Fusion based HHT features [40]	62
Baseline [41]	71
Proposed method with KNN	81.69
Proposed method with SVM	84

**Table 4:** Comparison of different methods.

average processing time consisting of the time taken by the application on Matlab to filter the signal, to analyze the signal, to detect peaks, to calculate and classify the feature vector of the extracted peak, to merge the different emotional decisions and finally to display the emotional state resulting from this analysis is 2.1482 seconds for the K Nearest Neighbors algorithm and 1.7835 seconds for the support vectors machine algorithm. Thus, our algorithms require a maximum of 2.1482 seconds to analyze these signals to detect a human's emotional state. Yu and al. introduce [42] introduce some applications of the machine learning classification algorithms for human behavior analysis.

To propose generalized methods which could be used with all types of data, transfer learning from one system to another is a key problem. In our future work, we will study this problem in detail. In the literature, a lot of methods have been proposed to deal with this problem. For example, Lei Zhang et al. proposed a discriminative kernel transfer learning method. The proposed method is independent from the classification targets. It realizes a robust domain transfer by simultaneously integrating domain-class-consistency metric based discriminative subspace learning, kernel learning in reproduced kernel Hilbert space, and representation learning between source and target domain [43]. Another method proposed in the literature aims to reconstruct the target data with a few source data points by using

a sparse coefficient matrix in some low-dimensional latent space. In this method, the sparse reconstruction coefficient matrix and the low-dimensional latent space projection are learned simultaneously [44]. Extreme Learning Machine (ELM) based Domain Adaptation (EDA) method has been proposed as a cross domain learning method. In this method, there is a network classifier and category transformation are done by using labeled source data, a limited number of target data and unlabeled target data. Authors also have extended EDA for structural information sharing of multiple local features with different feature representations [45]. These 3 methods outperform other state of the art proposed methods.

## Conclusion

In this paper, we have presented a novel method of multimodal recognition of emotions based on the processing of physiological signals. Two classification algorithms have been developed and implemented in the Emotica system. The different results show a marked improvement in the recognition rate of emotions. In our future work, on the one hand, we will study physiological signals acquisition platforms in order to generate our own recognition base and improve our classification method. We obtain excellent emotion recognition rate using SVM classification method. To improve this rate, we plan to implement an extension ELM classification method. This extension called Adaptive Extreme Learning Machine (AELM) is proposed for handling cross-task

(domain) learning problems [46]. Lei Zhang et al. [43-46] demonstrated that ELM method is faster for learning and application and more accurate than SVM and KNN. On the other hand, we will set up a complete system from the acquisition of physiological signals for the detection of emotions. Moreover this system will allow creating a more appropriate recognition base for a wide range of people.

## References

- Pantic M, Rothkrantz LJM (2003) Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE* 91: 1370-1390.
- Wagner J, Andre E, Lingenfelser F, Kim J (2011) Exploring fusion methods for multimodal emotion recognition with missing data. *IEEE Transactions on Affective Computing* 2: 206-218.
- Me'nard M, Richard P, Hamdi H, Dauce' B, Yamaguchi T (2015) Emotion recognition based on heart rate and skin conductance. In *PhyCS* pp: 26-32.
- Kone' C, Tayari MI, Thanh NL, Belleudy C (2015) Detecting and classification of the basic emotions using a multimodal approach for emotions detection. In *The Fourth International Conference on Global Health Challenges. Global health*.
- Zhang L, Zhang D (2016) Visual understanding via multi-feature shared learning with global consistency. *IEEE Transactions on Multimedia* 18: 247-259.
- Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS (2006) Heart rate variability: a review. *Med Biol Engg Comp* 44: 1031-1051.
- Neto EPS, Neidecker J, Lehot JJ (2003) Comprendre la variabilite' de la pression arte'rielle et de la fre'quence cardiaque. In *Annales franc,aises d'anesthe'sie et de re'animation* 22: 425-452.
- Fre'de'ric Bousefsaf (2014) Mesure sans contact de l'activite' cardiaque par analyse du flux vide'o issu d'une came'ra nume'rique: extraction de parametres physiologiques et application a' l'estimation du stress. PhD thesis, Universite' de Lorraine.
- Denis Jacques, Nicolas Zdanowicz, Christine Reynaert. Les bases anatomiques et physiologiques du syste'me limbique et leur implication dans le traitement du be'galement. Les be'gaiements de l'adulte. Wadre: Mardaga. pp: 45-49.
- Healey JA (2000) Wearable and Automotive Systems for A ect Recognition from Physiology. PhD Thesis, Citeseer.
- Fernandez R, Picard R (1997) Signal processing for recognition of human frustration *IEEE Conferences, Seattle, USA*.
- Healey J, Picard R (1998) Digital processing of affective signals. In *Acoustics, Speech and Signal Processing. Proceeding of IEEE International Conference* 6: 3749-3752.
- Healey J, Picard R (2000) Smartcar: detecting driver stress. In *Pattern Recognition, Proceedings of 15th International Conference* 4: 218-221.
- Takahashi K, Tsukaguchi A (2003) Remarks on emotion recognition from multi-modal bio-potential signals. In *Systems, Man and Cybernetics. IEEE International Conference* 2: 1654-1659.
- Rani P, Sarkar N, Smith CA, Adams JA (2003) Affective communication for implicit human-machine interaction. In *Systems, Man and Cybernetics. IEEE International Conference* 5: 4896-4903.
- Herbelin B, Benzaki P, Riquier F, Renault O, Grillon H, et al. (2005) Using physiological measures for emotional assessment: a computer-aided tool for cognitive and behavioural therapy. *Int J Disab Human Develop* 4: 269-278.
- Takahashi K (2004) Remarks on emotion recognition from multi-modal bio-potential signals. In *Industrial Technology, IEEE International Conference* 3: 1138-1143.
- Zhou JZ, Wang XH (2005) Multimodal affective user interface using wireless devices for emotion identification. In *IEEE Engineering in Medicine and Biology 27th Annual Conference* pp: 7155-7157.
- Rainville P, Bechara A, Naqvi N, Damasio AR (2006) Basic emotions are associated with distinct patterns of cardiorespiratory activity. *Int J Psychophysiol* 61: 5-18.
- Liu C, Conn K, Sarkar N, Stone W (2007) Affect recognition in robot assisted rehabilitation of children with autism spectrum disorder. In *Proceedings IEEE International Conference on Robotics and Automation* pp: 1755-1760.
- Villon O, Lisetti C (2007) A user model of psycho-physiological measure of emotion. In *International Conference on User Modeling* pp: 319-323.
- Rani P, Sarkar N, Adams J (2007) Anxiety-based affective communication for implicit human-machine interaction. *Advanced Engg Informatics* 21: 323-334.
- Ho'nig F, Batliner A, No'th E (2007) Real-time recognition of the affective user state with physiological signals. *Proceedings of the Doctoral Consortium, Affective Computing and Intelligent Interaction*.
- Kreibig SD, Wilhelm FH, Roth WT, Gross JJ (2007) Cardiovascular, electro dermal, and respiratory response patterns to fear-and sadness-inducing films. *Psychophysiol* 44: 787-806.
- Liu C, Conn K, Sarkar N, Stone W (2008) Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE Transactions on Robotics* 24: 883-896.
- Conn K, Liu C, Sarkar N, Stone W, Warren Z (2008) Towards affect-sensitive assistive intervention technologies for children with autism. *INTECH Open Access Publisher*.
- Cheng B, Liu GY (2008) Emotion recognition from surface emg signal using wavelet transform and neural network. In *Proceedings of the 2nd international conference on bioinformatics and biomedical engineering (ICBBE)* pp: 1363-1366.
- Benovoy M, Cooperstock JR, Deitcher J (2008) Biosignals analysis and its application in a performance setting. In *Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing* pp: 253-258.
- Clynes M (1988) Generalised emotion how it may be produced and sentic cycle therapy. In *Emotions and Psychopathology* pp: 107-170.
- Lang PJ (2005) International affective picture system (iaps): Digitized photographs, instruction manual and affective ratings. Technical Report.
- Abdat F (2010) Reconnaissance automatique des e'motions par donne'es multimodales: expres-sions faciales et signaux physiologiques. PhD thesis, UNIVERSITE' DE METZ.
- Manfred C (2013) Music, mind, and brain: The neuropsychology of music. Springer Science & Business Media.
- Larose DT (2005) An introduction to data mining. Traduction et adaptation de Thierry Val-laud.
- Lipo, Wang (2005) Support Vector Machines: Theory and Applications Series: Studies in Fuzziness and Soft Computing Springer Science & Business Media 177: 1-431.



35. Chang CC, Lin CJ (2011) LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)* 2: 1-27.
36. Milgram J, Cheriet M, Sabourin R (2006) "One Against One" or "One Against All": Which One is Better for Handwriting Recognition with SVMs?? Tenth International Workshop on Frontiers in Handwriting Recognition, Oct 2006, La Baule (France), Suvisoft.
37. Allwein EL, Schapire RE, Singer Y (2000) Reducing multiclass to binary: A unifying approach for margin classifiers. *J Machine Learning Res* 1: 113-141.
38. Khan SM, Yu M, Feng P, Wang L, Chambers J (2015) An unsupervised acoustic fall detection system using source separation for sound interference suppression. *Signal processing* 110: 199-210.
39. Kim KH, Bang SW, Kim SR (2004) Emotion recognition system using short-term monitoring of physiological signals. *Med Biolo Eng Comp* 42: 419-427.
40. Zong C, Chetouani M (2009) Hilbert-huang transform based physiological signals analysis for emotion recognition. In 2009 IEEE International Symposium on Signal Process- ing and Information Technology (ISSPIT) pp: 334-339.
41. Tayari IM, Le-Thanh N, Amar CB (2012) Detecting depression using a multidimensional model of emotional states. *GLOBAL HEALTH 2012: The First International Conference on Global Health Challenges*.
42. Yu J, Zhou H, Gao X (2015) Machine learning and signal processing for human pose recovery and behavior analysis. *Signal Processing* 110: 1-4.
43. Zhang L, Yang J, Zhang D (2017) Domain class consistency based transfer learning for image classification across domains. *Information Sci* 418: 242-257.
44. Zhang L, Zuo W, Zhang D (2016) Lsd: Latent sparse domain transfer learning for visual adaptation. *IEEE Transactions on Image Processing* 25: 1177-1191.
45. Zhang L, Zhang D (2016) Robust visual knowledge transfer via extreme learning machine- based domain adaptation. *IEEE Transactions on Image Processing* 25: 4959-4973.
46. Zhang L, He Z, Liu Y (2017) Deep object recognition across domains based on adaptive extreme learning machine. *Neurocomputing* 239: 194-203.