The Human Side of Service Engineering, Vol. 143, 2024, 108–117
https://doi.org/10.54941/ahfe1005089

Support Vector Machines Models for Human Decision-Making Understanding: A Different Perspective on Emotion Detection

Paolo Barile[1], Clara Bassano[2], and Aysel Alizada[3]

[1]Department of Computer Science (DI) University of Salerno, Italy
[2]Department of Pharmacy (DIFARMA) University of Salerno, Italy
[3]Department of Management, Sapienza University of Rome, Italy

ABSTRACT

The high integration of artificial intelligence (AI) into our daily life has led to get much research on the technology's potential, particularly with its ability to understand human emotions. Our study mainly focuses on the potential of Support Vector Machine (SVM) models in facial emotion recognition (FER) and examines the possibility to analyse the results of emotion detection through the Viable Systems Approach (VSA) perspective. The understanding of emotions as an important difference between human and machine is an ongoing issue, underlining the necessity for AI to incorporate emotional intelligence. The main objective of the project is to fill the knowledge gap existing between AI and human surroundings, ethics, and social factors. From an experimental point of view, we realized three different SVM models based on the most widely used kernel functions (linear, polynomial, and radial). Then, we used the "JAFFE" dataset to test the models on three different configurations of the initial data, to understand which parameters are most influential for the performance of the classifiers and to investigate the limitations and potential of SVMs for emotion recognition. The next step addressed in the study is to integrate computer science with VSA, providing a fresh perspective on emotion detection. This approach is not just about developing a framework for human-computer interaction (HCI) but delves deeper into understanding the social dynamics underlying decision-making. In conclusion, our paper emphasizes the significance of emotional aspects in HCI and the potential of AI in understanding human emotions. By employing the VSA, it extends the discussion on AI's capabilities in complex decision-making processes, highlighting the necessity for AI systems to resonate cognitively with human users in increasingly digital environments.

Keywords: Emotion detection, Human computer interaction, Viable systems approach, Support vector machines, Facial emotion recognition

INTRODUCTION

The period we are living in is marked by the continuous advent of new techniques and tools based on artificial intelligence (AI) aimed at carrying out the activities that characterize our everyday lives. Recent technological developments, mainly related to the Internet of things (IoT), have led information technology and artificial intelligence to contaminate almost any area of human knowledge (Laghari, 2022). What we might call the "Age of Computerization" decrees the need to establish a daily relationship between technology and those who use it, making the issue of human-computer interaction (HCI) a central topic in the scientific literature (Wobbrock, 2016). By now, machines have achieved the ability to analyse and understand all the material aspects of the world around us: the processing machines capacity is at the complete service of man, to such an extent that the same lifestyle and habits of human beings have undergone a significant transformation, impacting people's daily routine at all social levels. Nevertheless, it has been proven how much the emotional and affective aspect is important in the communication and decision-making processes, in social interaction and, therefore, also in HCI (Denham, 2007). It has been finally understood that the human being is characterized by rationality, like intelligent machines, but also and particularly by emotion, permeating human cognitive and decision-making processes. The emotions relevance appears on the one hand as a defence/-survival barrier for humans against the growing computational potential of machines, and on the other as the human connotation that guarantees balance and not dominance in HCI (Golinelli, 2020). Emotions strongly characterize human decision-making processes and differentiate them from the rational and computational ones of machines, beyond the sophistication of the latter. However, the greatest limitation of artificial intelligence consists in the lack of understanding and awareness of the human context which makes it impossible for machines to act according to ethical and social considerations (Shanahan, 2018). The current scenario suggests AI scholars to focus on aspects of knowledge that go beyond the material world, in order to allow the possibility of imbuing machines with a form of emotional intelligence, enabling computers to simulate human empathy.

Our study presents a twofold objective: the first is to analyse possibilities and limitations of Support Vector Machines (SVMs) for facial emotion recognition (FER); the second is to provide a new perspective on emotion detection, based on Viable Systems Approach (VSA), aimed at the interpretation of human decision-making dynamics.

The remainder of the paper is structured as follows: section 2 provides an analysis of the relevant literature; section 3 describes the experimental setup; section 4 presents and discusses the experimental results; section 5 focuses on the new perspective given by VSA; section 6 draws conclusions based on the findings.

LITERATURE REVIEW

The computer science field that deals with the study of human emotions is affective computing (AC) (Picard, 1997). The main objective of the AC is to create systems, based on artificial intelligence and ML algorithms, capable of automatically recognizing the state of mind that characterizes an individual in a specific time interval. The analysis is carried out thanks to the use of various types of data linked to physical parameters, physiognomic and proxemic or relating to language. Among various AC techniques, the most direct and widespread method is FER, consisting in the classification of human emotions through the analysis of images or videos containing people's faces. The emotions distinguishable by the AI, on the basis of objective factors related to the combination of contractions and relaxations of the muscles of the face, are the seven primary emotions of Ekman (Ekman, 1987): neutral, fear, anger, happiness, sadness, disgust, and surprise. The most commonly used training techniques for FER involve the use of Deep-Learning (DL), which is widely used for the task of image classification. However, several difficulties related, above all, to the unavailability of a sufficient amount of data for training neural networks have recently led to a preference for different machine learning techniques (Balasubramanian, 2019). In this regard, the approach presented in our paper, based on SVM models for FER, is seen as a viable alternative for emotion classification to the use of deep learning; the main reason is that SVMs have much better suitability for small amounts of data than DL algorithms, especially when the data is well-structured, and the classes are separable (Lyons, 2020). The study of HCI is often limited to emotion classification as it is; however, to investigate the consonance (structural compatibility between two or more agents) between human and machine, confirming both components to be capable of developing cognitive resonance and managing decision-making processes, there is the need to reinterpret the results in a different key. This possibility is offered by the theoretical contributions of VSA and "absentee management" (Maggioni, 2014; Piciocchi, 2018). Emotions and their regulation belong to the human species, making it possible to differentiate the human cognitive process from the machine ones; regulating emotional experiences is fundamental for the well-being and productivity of the individual, as this process expresses the degree of adaptive compatibility (consonance) all kinds of interactions. According to several studies, the decision-making processes of individuals are different and context-based; emotional stability can be "modulated" according to the degree of coherence and adaptability to the interacting interlocutors (Barile et al., 2012; Petruzzelli, 2010). If the VSA admits the understanding of phenomena according to a logic of subjective interpretation - rationality and emotional expectation - then the concept of absence becomes relevant for understanding how much the growth of knowledge is inextricably linked to the non-knowledge/emptiness that persists. There is, therefore, an antagonistic and complementary relationship between knowledge and ignorance which, according to Morin (Morin, 2018), justifies the incompleteness and uncertainty that stimulate creative actions. The uncertainty of our cognitive variety represents the so-called absence which, despite the progress, distinctively characterizes living beings, not machines. From the point of view of management and balance in HCI, it is possible to state an evolutionary dimension of the "unknown", representing the potential for further evolution and knowledge that must be shared.

EXPERIMENTAL SETUP

Dataset

The dataset selected for our work is the "Japanese Female Facial Expression" (JAFFE) (Lyons, 2021). JAFFE consists of grayscale images depicting the faces of Japanese women, all captured frontally to the camera and against a neutral back-ground. The dataset contains a total of 213 high- resolution images, 256x256 pixels in size, labelled on the basis of Ekman's primary emotions. The dataset is balanced and does not suffer from the presence of minority classes (Goodfellow, 2016). Since SVM algorithms are known to work more smoothly on small sets, the rather small number of samples compared to other facial expression dataset makes JAFFE a particularly appropriate choice for our investigation. In addition, the high resolution of the images greatly simplifies the feature extraction steps.

Feature Selection

Features selection plays a crucial role in machine learning, as it goes to determine the data representation on which the learning model will be built. The task of emotion recognition using SVM is comparable to a generic image classification problem: the different output classes can be discriminated based on various types of features, derived from shape, texture or colour characteristics of the images themselves. In our study, given the high resolution of the images, we decided to use a very common method of feature selection based on the numerical representation of the pixels of which each image is composed. Specifically, each pixel was represented, on the basis of its shade of grey, with a value between 0 and 255; as a result, after the feature selection phase, each image in the JAFFE dataset was characterized by a sequence of 65536 numerical values, one for each pixel.

Kernel Functions and Hyperparameter Tuning

The classification mechanism of SVM algorithms is based on the identification of a hyperplane capable of linearly separating samples, belonging to different classes, in the feature space. Kernel functions are the key element that enables this operation (Shawe-Taylor, 2004). In order to evaluate the behaviour of SVM models with respect to the facial emotion recognition task, we chose to use three of the most common kernel functions (linear (1), polynomial (2) and radial (3)):

1. $K_{lin}\left(x,y\right) \;=\; x^T y$
2. $K_{poly}\left(x,y\right) \;=\; \left(\gamma\, x^T y\right)^d$
3. $K_{rbf}\left(x,y\right) \;=\; exp(-\gamma\; \left|\left|x-y\right|\right|^2)$

where x and y represent two different observations of the dataset, γ , d and r are numerical coefficients, called hyper- parameters, that determine the configuration and dimensionality of the new vector space.

The optimal value of each hyperparameter is not possible to know a priori, but varies according to the problem. Techniques of hyperparameter tuning, based on grid search and cross validation, were used to find the best values employed in the continuation of the study for training the three SVM models.

RESULTS AND DISCUSSION

Evaluation of the Original Dataset

The experimental setup for our study involved dividing the images in the dataset, as usual, into training set (75% of the original samples) and test set (remaining 25%). Accordingly, of the total 213 images, 159 were used for model training and 54 for its evaluation. The test accuracy for the linear and polynomial kernel (89% and 85%, respectively) appears to be satisfactory and in line with the expected results based on the literature study (Adeyanju, 2015). In contrast, the accuracy value for the radial kernel model, turns out to be significantly lower (48%).

Several case studies report better (or at least comparable) performance of the radial kernel than the others for image classification tasks. However, there is no scientific evidence that the higher complexity of the radial kernel should necessarily guarantee a better performance of the classification process. A possible explanation for the underperformance of the radial kernel could be related to an overfitting condition: where the number of features is much greater than the number of samples in the training set, the higher the complexity of the algorithm, the larger the tendency to fit the training data (Alam, 2016).

To evaluate the influence of the number of features on the testing accuracy of the models, tests were carried out by going to reduce the size of the original images from time to time. Specifically, the three SVM models were tested by keeping the same subdivision of training and test set using different sizes for the images.

The performance of the three models, in terms of classification accuracy on the test set, is shown in Table 2.

Table 1. Test accuracy as a function of the dimension of the images in the dataset.

Test Accuracy	SVM Kernels		
Dimension (Pixels)	Linear	Polynomial	Radial
256x256	0.89	0.85	0.48
200x200	0.85	0.79	0.21
150x150	0.81	0.74	0.47
100x100	0.66	0.65	0.24

The linear and polynomial kernels show a very similar trend: as the number of features decreases, the accuracy of the model diminishes. As for the radial kernel, on the other hand, random fluctuations in the test accuracy value are observed, showing no correlation with the size of the images in the dataset. Figure 3, provides a more immediate visualization of the anomalous behavior of the radial kernel SVM, appearing to confirm the overfitting hypothesis advanced previously.

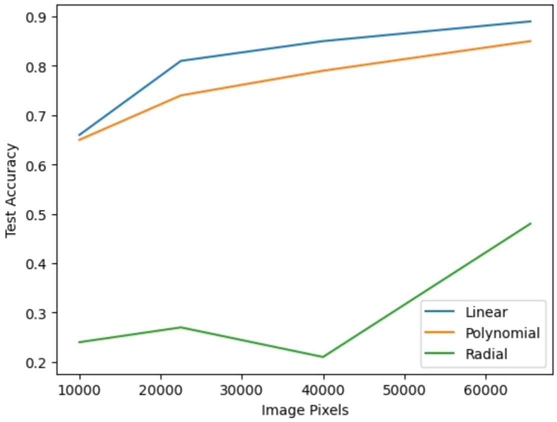

Figure 1: Graph of test accuracy as a function of the number of features (image pixels).

Evaluation of the Augmented Dataset

To evaluate the algorithm's performances variation as the number of images in the training set changes, we used data augmentation techniques. The training time and the test accuracy for each of the different SVM models was evaluated on four different sizes of the training set. The aim of the investigation was to understand whether the poor generalization ability on new data evidenced by the radial kernel SVM could be addressed by introducing new images into the training set.

While it is not possible to come up with a rigorous mathematical formulation, several studies (Bordes, 2005) demonstrate the existence of a lower bound for the training complexity of an SVM model: this is $N \cdot s$ for the linear kernel and between N^2 and N^3 for the nonlinear ones (with N the number of samples in the training set and s the number of features).

Table 2. Training complexity as a function of number of training samples.

Training Time (s)	SVM Kernels		
Training Samples	Linear	Polynomial	Radial
~750	24	22	20
~1500	76	76	86
~3000	332	329	331
~6000	1124	1247	1087

Table 3 shows the performance in terms of training computational time for each of the three SVM models as the size of the dataset changes, following the data augmentation operations.

The training complexity, as expected, showed to be highly dependent on the size of the dataset. It can be seen, however, that the run times for the linear model do not differ significantly from those for the radial and polynomial kernel. Given the large number of features that characterize each of the samples, it applies to all the examined cases:

$$N^2 < N \cdot s < N^3.$$

The trends highlighted by the table are, therefore, justified by the fact that the number of operations performed by the models in the training phase is of the same order of magnitude for both the linear kernel and the radial or polynomial one.

As for the test accuracy, the performance of the three models for each of the modified datasets is summarized in Table 4.

Table 3. Test accuracy as a function of the number of training samples.

Test Accuracy	SVM Kernels		
Training Samples	Linear	Polynomial	Radial
~750	0.13	0.13	0.16
~1500	0.16	0.21	0.20
~3000	0.15	0.17	0.21
~6000	0.18	0.25	0.23

What is clear, in addition to a lack of improvement in the performance of the radial kernel, is a general degradation in accuracy for all three models. One possible explanation is that the use of synthetic images generated a new case of overfitting; in other words, the new features are not causally correlated with the model classification function, resulting in random outputs. The hypothesis is also confirmed by the confusion matrices shown in Figure 5 for the 750-sample training set: it can be seen that the individual predictions for each of the three models differ significantly, a clear sign that the algorithms simply guessed.

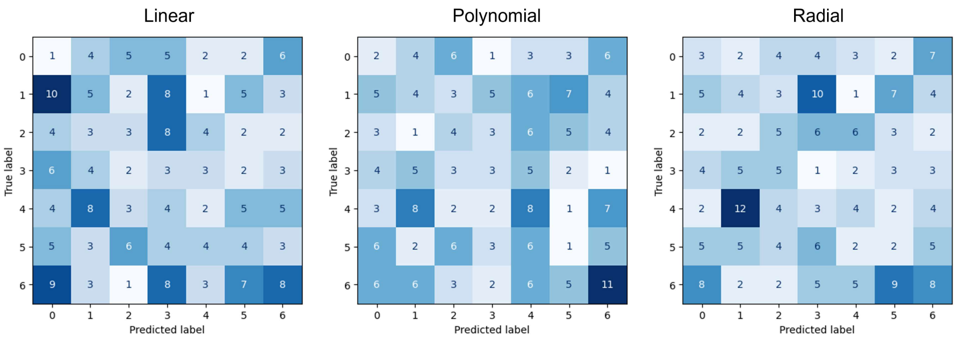

Figure 2: Confusion matrices for the three models trained on the 750-sampled dataset.

VSA INTERPRETATION OF EMOTIONS

The study of emotional space is a rather widespread topic in the computer science community; as can be deduced from the analysis of the literature, AC provides all the necessary means for the analysis of emotional states through techniques and algorithms that allow software to work in a completely automatic manner; however, the connection between the emotional state of subjects and their behaviour in decision-making is still completely unexplored territory (Barile et al., 2020). The meaning of empathy, which can be understood as a distinguishing element between humans and machines, has deep roots that are not limited solely to the recognition and emulation of emotions. AI's ability to detect empathy depends on being able to reinterpret emotion detection results in a new key, allowing it to observe the decision-making patterns underlying the acquisition of new information. In other words, emotions can be viewed as a "chemical" response to mental stimuli: in order to generate empathy, it is necessary to know according to what criteria one's mind absorbs such stimuli.

The concepts of value categories (or general schemes) and consonance, typical of VSA, allow the two sides to be related: when observed through the lens of VSA, the human emotional sphere eventually becomes an expression of the decision-making processes that influence each person's actions. Value categories are the building block at the basis of each individual's behaviour: every behavioural dynamic, reaction or decision depends on strong beliefs, representing the first response to the mental and emotional stimuli that arise in the various decision-making contexts (Badinelli, 2012). The analysis of emotions by means of AI techniques, when explored through the VSA theory can therefore lead to the study of consonance, which would allow relational dynamics to be investigated at a higher level. The impact of the research could be, given the premises, very significant both in the field of HCI and in the psycho-sociological field related to the study of human behaviour, since the study of emotions through AI is not only aimed at the development of a framework for HCI but also allows for a deeper investigation of the social dynamics underlying decision making.

CONCLUSION

Our work highlights an aspect of undoubted interest: in the digital age, researchers are going beyond the "simple" erudition of machines to recognize Ekman's emotions. New research directions are the projection towards verticalizations which, in our case, impact on the system consonance to develop ever greater system resonance degrees. In particular, contextualizing in HCI, means that there are some investigation and application fields that would confirm syntropic man-machine processes - such as the study of emotions -, capable of increasing and optimizing decision-making processes in the complexity and variability scenarios. On the application side, it is important to point out that the results of our study are strongly related to the characteristics of the dataset: the choice of different data on which to train the SVM algorithms could likely have led to a different outcome; in general, whatever recognition method is used, the result of FER carries with it an inherent dependence on the quality of the images used. Nonetheless, our research demonstrates the possibility of using different SVM models for FER, even using a rather narrow dataset and highlights the main application limitations and parameters that can affect, negatively or positively, the performance of classification algorithms. Ultimately, the experimental part once again demonstrates the ability of AI to detect human emotions, the theoretical part suggests a new interpretive point of view that can make people think about the possibility of extending the research to models for recognizing consonance and empathy in decision-making dynamics.

REFERENCES

Adeyanju I. A., Omidiora E. O. and Oyedokun O. F., "Performance evaluation of different support vector machine kernels for face emotion recognition," 2015 SAI Intelligent Systems Conference (IntelliSys), London, UK, 2015, pp. 804–806, doi: 10.1109/IntelliSys.2015.7361233.

Alam S., Kang M., Pyun J.-Y. and Kwon G.-R., "Performance of classification based on PCA, linear SVM, and Multi-kernel SVM," 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN), Vienna, 2016, pp. 987–989, doi: 10.1109/ICUFN.2016.7536945.

Badinelli, R., Barile, S., Ng, I., Polese, F., Saviano, M., Di Nauta, P. (2012). Viable service systems and decision making in service management. Journal of Service Management, 23(4), 498–526.

Balasubramanian B., Diwan P., Nadar R. and Bhatia A., "Analysis of Facial Emotion Recognition," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 945–949.

Barile S., Bassano C., Lettieri M., Piciocchi P., Saviano M., Intelligence Augmentation (IA) in Complex Decision Making: A New View of the VSA concept of Relevance. In (Eds.) Spohrer J. C, Leitner C. AHFE 2020, AISC 1208, Springer Nature Switzerland AG 2020, pp. 1–8, https://doi.org/10.1007/978-3-030-51057-2_36.

Barile, S., Pagliuca, I., Vito, P., "Dal management essenziale al management assenziale", Risma, 2021; Deacon, T. W. Natura incompleta, Le Scienze, 2012.

Bordes, A., & Ertekin, S., & Weston, J., & Bottou, L., (2005). Fast Kernel Classifiers with Online and Active Learning. Journal of Machine Learning Research. 6. 1579–1619.

Denham, S. A., Bassett, H. H., & Wyatt, T. (2007), The Socialization of emotional competence. In J. E. Grusec, & P. D. Hastings (Eds.), Handbook of socialization: theory and research (pp. 614–637), New York: Guilford Press.

Ekman, P., Friesen, W. V., O'sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., Lecompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. Journal of Personality and Social Psychology, 53(4), 712–717. https://doi.org/10.1037/0022-3514.53.4.712

Golinelli G., M., L'approccio sistemico vitale (ASV) al governo dell'impresa - Vol. 1–2–3, 2017–2020.

Goodfellow, I., (2016). Deep learning. Yoshua Bengio, Aaron Courville. Cambridge, Massachusetts. pp. 524–534. ISBN 0-262-03561-8. OCLC 955778308.

Laghari, A. A., Wu, K., Laghari, R. A. et al., A Review and State of Art of Internet of Things (IoT). Arch Computat Methods Eng 29, 1395–1413 (2022). https://doi.org/10.1007/s11831-021-09622–6.

Lyons, Michael & Kamachi, Miyuki & Gyoba, Jiro. (2020). Coding Facial Expressions with Gabor Wavelets (IVC Special Issue).

Lyons, M., (2021). "Excavating AI" Re-Excavated: Debunking a Fallacious Account of the Jaffe Dataset. SSRN Electronic Journal. 10.2139/ssrn.3900990.

Maggioni, V., Barile, S., Calabrese, M., & Iandolo, F. (2014). Emerging paradigms: how time affects decision making. L'industria, 35(1), 29–48.
Martinez A, Du S., A Model of the Perception of Facial Expressions of Emotion by Humans: Research.
Morin E., Conoscenza, ignoranza, mistero, Raffaello Cortina, Milano 2018.
Petruzzelli M., A. (2010). The Future of Innovation is Openness, in von Stamm, B., Trifilova, A. (eds.) The Future of Innovation. Gower Publishing.
Picard, Rosalind (1997). Affective Computing. Cambridge, MA: MIT Press.
Piciocchi, P.; Bassano, C.; Pietronudo, M. C.; Spohrer, J. Digital Workers in Service Systems: Challenges and Opportunities. doi: 10.1007/978-3-319-98512-1_18. pp. 409–432. In Handbook of Service Science - ISBN:9783319985121, 97... vol. II.
Shanahan, M. (2018). La rivolta delle macchine. LUISS University Press.
Shawe-Taylor, J.; Cristianini, N. (2004). Kernel Methods for Pattern Analysis. Cambridge University Press.
Wobbrock, J. O.; Kientz, J. A. Research contributions in human-computer interactions, 2016, 23.3: 38–44.