HAL
archives-ouvertes.fr

Intelligent pixels of interest selection with application to facial expression recognition using multilayer perceptron

Taner Danisman, Ioan Marius Bilasco, Jean Martinet, Chaabane Djeraba

▶ **To cite this version:**

Taner Danisman, Ioan Marius Bilasco, Jean Martinet, Chaabane Djeraba. Intelligent pixels of interest selection with application to facial expression recognition using multilayer perceptron. Signal Processing, Elsevier, 2013, Special issue on Machine Learning in Intelligent Image Processing, 93 (6), pp.1547-1556. 10.1016/j.sigpro.2012.08.007 . hal-00804171

HAL Id: hal-00804171
https://hal.inria.fr/hal-00804171

Submitted on 29 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Intelligent Pixels of Interest Selection with Application to Facial Expression Recognition using Multilayer Perceptron

Taner Danisman[1,*], Ioan Marius Bilasco, Jean Martinet, Chabane Djeraba

Lille University of Science and Technology, CNRS UMR 8022, Parc Scientifique de la Haute Borne, 50 Avenue Halley, 59655 Villeneuve D'Ascq - France

Abstract

This paper presents an automatic way to discover pixels in a face image that improves the facial expression recognition results. Main contribution of our study is to provide a practical method to improve classification performance of classifiers by selecting best pixels of interest. Our method exhaustively searches for the best and worst feature window position from a set of face images among all possible combinations using MLP. Then, it creates a non-rectangular emotion mask for feature selection in supervised facial expression recognition problem. It eliminates irrelevant data and improves the classification performance using backward feature elimination. Experimental studies on GENKI, JAFFE and FERET databases showed that the proposed system improves the classification results by selecting the best pixels of interest.

Keywords: Facial expression recognition, multi layer perceptron, feature selection.

*Corresponding author

Email addresses: taner.danisman@lifl.fr (Taner Danisman), marius.bilasco@lifl.fr (Ioan Marius Bilasco), jean.martinet@lifl.fr (Jean Martinet), chabane.djeraba@lifl.fr (Chabane Djeraba)

[1]Present address: University of Angers, LISA Laboratory, 62 Avenue Notre Dame du Lac, 49000 Angers - France.

1. Introduction

Facial expression recognition (FER) is a hot research topic and a challenging problem on different domains including face recognition, human computer interaction, facial animation as well as social interaction. In the last decade, researchers from various disciplines focused on efficient, accurate and fast recognition of facial expressions. Emotions can be detected from physical sensors, image and video. Each sensor type has its own challenges such as noisy signals, high dimensionality and quality of selected features. There are many automatic FER studies achieving high accuracy on well-defined datasets. However, these studies still perform poor results under real world situations. Therefore there is still a considerable accuracy gap for realistic classification scenarios. One solution for this problem is to improve the classification results in terms of objective measures. Among others, feature selection is an important step towards better classifiers. Feature selection and reduction strategies are used to select relevant features to create robust models. In this scope, majority of the previous studies on FER considered the face and facial features as a combination of coarse rectangular units [1, 2]. These units are used to locate or extract valuable facial feature information. Although its implementation simplicity, it includes useless and noisy data for the machine learning step. Therefore there is a need to find local pixel of interests (POI) to be used in FER. Group of POI provide non-rectangular masks that can be used to improve the classification performance.

Selection of the best variable and feature become the focus in classification research where there are thousands of different possibilities. Feature selection is the technique for selecting a subset of relevant features from original data to reduce feature size while maximizing the classifier output. Wrapper and filter based feature selection are the most common two approaches in the field. Wrappers evaluates the importance of specific features considering a particular learning algorithm [3] whereas filter based methods reduce the features space using a specific filter. Regardless of the fact that computational complexity of wrapper based methods put aside, the facial area used in FER is a small region that can be represented by regions as small as 20×20 to 50×50 pixels for vision based algorithms.

In this study, we used an analytic approach that performs wrapper based feature selection by exhaustive searching of all possible set of feature windows to find informative pixels to improve the results of FER. For a given emotion class, we created corresponding emotion mask to improve the Multi-

layer Perceptron (MLP) model's performance. Our experiments on different datasets showed that proposed method gives better results than full frame and the best traditional feature window based classification.

The rest of this paper is organized as follows. In Section 2, we briefly overview related works. Database material, mask generation and its application to FER is described in Section 3. Experimental results and discussion are presented in Section 4, followed by the conclusion.

2. Related Works

Considering either analytic or holistic classification problem, there is a need to eliminate redundant and noisy information. Analytic approaches are widely use in face recognition domain and they are based on the detection of specific facial features such as eyes, eyebrows, nose, mouth and the locations of facial fiducial points such as corner positions of eye, mouth and their geometric relationships. Here a system is solved by considering its subparts and how they work together to produce particular results. For this reason, researchers use high level features of the face such as position of eyes, eyebrow, noise and mouth corners [4]. Among others, Nonnegative Matrix Factorization (NMF) is a popular dimension-reduction method that approximates a nonnegative matrix by the product of two other low-rank matrices. NMF distinguishes from others by its use of the non-negativity property. In contrast to holictic methods, NMF is able to learn parts of image. Local Binary Patterns (LBP)[5] is another analytic method used in texture classification and face recognition [6] which summarizes local structure of an image. LBP is a powerful gray-scale invariant texture feature providing more discriminative and invariant features for the recognition phase. Here the input face is divided into non-overlapping regions to compute an enhanced feature histogram.

Holistic approaches consider the encoding of the entire facial image into a point on a high dimensional space. In holistic methods, the problem can not be solved by its subcomponents. Instead, the problem as a whole involves the solution. In the literature, Principal Component Analysis (PCA)[7], Linear Discriminant Analysis LDA [8] and Independent Component Analysis (ICA) [9] are common holistic methods extensively used for finding more informative features which reduce the dimensionality of original data by rejecting low variance features. For example PCA reduces the dimensionality by assuming that variance implies the importance and finds the reduced set of features

that mimics the original data. However, ability to reduce original data is not appropriate for all problem scenarios. For the supervised FER problem we need local representative areas and pixels in a face image that can be used for the recognition.

[10] proposed an efficient online NMF variant called OR-NMF that can be applied to large-scale datasets. Slow convergence problem is adressed by use of NeNMF [11] and MD-NMF [12]. [13] analysed the relationship between the input space and feature space using discriminant analysis and provided an input variable selection method. They selected variables that contain large amount of discriminative information and less discriminative information is discarded. [14] used LBP to obtain more complete description of the face using shifted and scaled sub-windows over the face images. [15] propose a pattern recognition method and its application to face recognition problem. In their method, extracted features are analyzed in the original space using feature feedback and then they identify the important segments of the original data that affects the classification performance. [16] present a holistic method to analyze facial expressions by focusing on the regions such as eyes, mouth etc. whose geometries are mostly affected by variation in facial expressions. They used PCA to recognize different parts of the face.

FER algorithms can be classified as feature-based and appearance-based methods. Feature based methods first detects facial feature points then classification is performed by considering geometric information. Appearance-based methods generally use the texture information for the classification. Hybrid methods like image ratio features [17] combine both the local texture information and the geometric information. [18] studied geometry and gabor-wavelet based FER using MLP and observed that gabor coefficients are more powerful than geometric positions.[19] proposed an efficient local appearance feature extraction method based on Steerable Pyramid(S-P) wavelet transform for face recognition problem.

In a recent study, [20] showed that FER is an analytic rather than holistic process. They used an elliptical mask to remove useless background pixels from the face image as a preprocessing step before their classification method. This type of masks eliminates the appearance of the hair and neck from the face image as seen on Figure 1. However, there exists many other non-informative skin pixels on this masked image which behaves like a noisy pixel in future processing steps. Therefore additional masks or methods are needed to eliminate these pixels during the preprocessing step. Figure 2 shows different masks generated by using different feature selection algorithms for

smiling and non-smiling classification.



Figure 1: Elliptical mask for faces from JAFFE dataset [34]

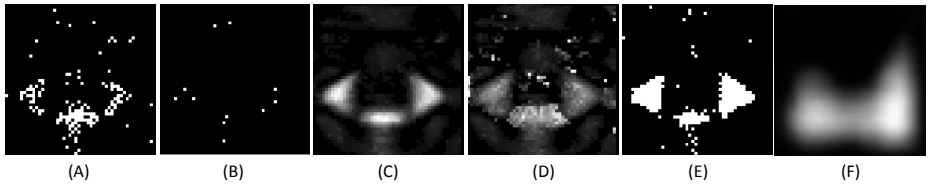


Figure 2: Different feature subsets created using the GENKI dataset for smiling and non-smiling classification. Left to right: (A) Correlation-based Feature Subset Selection. (B) Filtered Las Vegas algorithm. (C) Gain ratio attribute evaluator (200 attributes) using random subsamples. (D) Gain ratio attribute evaluator (100 attributes) using random subsamples. (E) OneR classifier based attribute selector. (F) The mask created by our method.

Correlation-based Feature Subset Selection (CFS) shown on Figure 2(A) uses a search algorithm along with a function to evaluate the merit of the feature subsets [21]. In this method usefulness of individual features depends on the level of inter-correlation among them. Low correlation with the class shows an irrelevant features for the selection. Filtered Las Vegas algorithm shown on Figure 2(B), make probabilistic choices to guide to a correct solution [22]. It uses the randomness property in the search process to find a guaranteed correct solution. This behavior is similar to the random pivot selection in Quicksort algorithm. Attribute evaluators seen on Figure 2 (C) and (D) produces similar results using random subsample of the datasets which evaluates the worth of an attribute by measuring the information gain with respect to the class. OneR classifier based feature selection shown on Figure 2(E) evaluates the worth of an attribute by using the OneR classifier [23]. The classifier selects one-by-one attributes from a dataset and generates a different set of rules based on the error rate from the training set. Attributes having the minimum error rate is selected. Figure 2(F) shows result of our method which uses MLP classifier to select pixels of interest.

[24, 20] studied the contribution of the upper and lower face regions in happy and sad facial expression classification. They used morphed faces to detect the intensity of a particular emotion on either the upper or lower face. In order to highlight important features they used multidimensional signal detection theory. Their study also shows that happy/sad classification is an analytic rather than a holistic process.

[25] defined Visual Context Patterns (VCP) for eye detection problem. Similar to our method, they searched for the smallest region of reference (ROR) rectangle that maximizes a quality function by satisfying maximal stable appearance while preserving less unstable appearance. Similarly in our method we search for the subregions that maximize and minimize the classification performance. On the other hand, in our method instead of selecting the smallest region, we used statistical elimination of regions by considering the variance value as explained in Eq.5. In terms of learning methodology they employed selective sampling based semi-supervised boosting method where false positive samples feed back to update the classifier in each training step. In our case, we create our classifier at the end of the process by considering intermediate classification results.

3. Material and Methods

In this paper, we make use of non-rectangular emotion masks for facial emotion recognition problem to improve overall classification results. Figure 3 shows general flow diagram of our method.

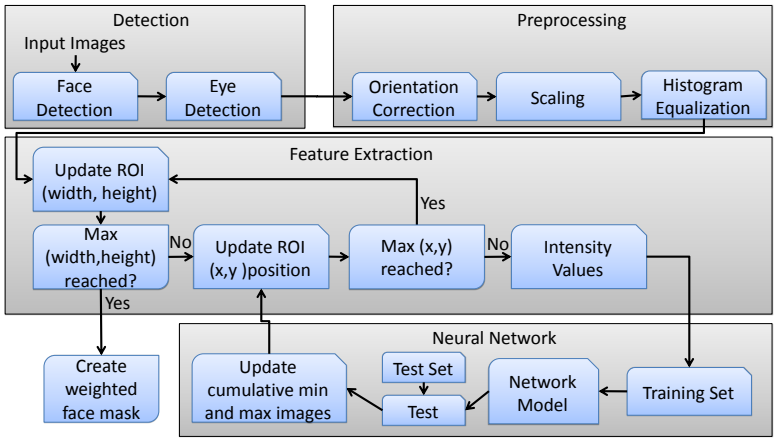


Figure 3: Flow diagram of our method.

Our proposed method exhaustively searches for the best feature window position from a set of static images among all possible combinations using an Artificial Neural Network (ANN) and creates a non-rectangular mask for a given emotion class. As the emotion recognition problem is a non-linear problem, the most common type of ANN, the Multi-Layer Perceptron (MLP) was chosen for this study.

3.1. Preprocessing

We adopted similar preprocessing steps explained in [26]. OpenCV library is used to detect and segment out faces from static images based on frontal face detector [2].

Original image size in GENKI dataset [27] is approximately 180×192 which reduces the face detection rate because of the small facial area. Therefore we performed linear interpolation to increase each image size to 450×480 by preserving the aspect ratio. In this case OpenCV detects 3980 faces out of 4000 individual face images (99.5% accuracy). After that, eye detection is performed to find the locations of the pupil centers. This step is necessary to correct the orientation of the face according to the vertical position of left and right pupil. For this reason, we used the neural network-based eye detector [28] available in the Stacked Trimmed Active Shape Model (STASM) [29] library to locate the positions of the pupils.

After the eye detection, we correct the orientation of the face using the vertical positions of both eyes. After that, the face is scaled down to 50×50 grayscale image considering the Inter Pupillary Distance (IPD). Face image x, y, *width* and *height* values are computed using the following equations.

$$F_x = Eye_{Left_x} - IPD/3 \tag{1}$$

$$F_y = Eye_{Left_y} - IPD/2.5 \tag{2}$$

$$F_w = IPD \times 1.6 \tag{3}$$

$$F_h = IPD \times 1.9 \tag{4}$$

where F_x, F_y, F_w and F_h represents the x, y, *width* and *height* of the face. Eye_{Left_x} and Eye_{Left_y} are the x and y positions of the left eye with respect to upper left origin of the image. Note that initial location of the OpenCV face detection results are updated here according to the IPD distance. Scalar values 1.6 and 1.9 of the IPD for the width and height are selected according to the experimental tests. Finally, histogram equalization

is applied to increase the contrast. Figure 4 shows the initial and cropped face region after the use of Eq. (1), (2), (3) and (4).

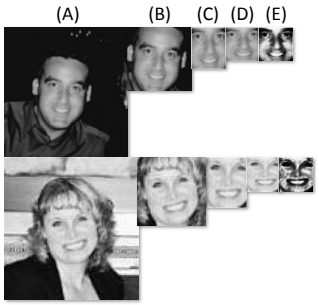


Figure 4: Preprocessing steps for the sample images from GENKI dataset. (A) Original image. (B) OpenCV face detection output. (C) Rotating and reshaping the face dimensions considering IPD distance. (D) Normalizing face size to 50×50. (E) Histogram equalization.

After preprocessing step, our training set consists of 1,976 smiling and 1,746 non-smiling faces.

3.2. Methods

We have performed smiling/non-smiling classification experiments on different datasets. In this study, any class other than pure "smiling" is defined as "non-smiling".

Consider a set of m windows $R_k = \{x, y, w, h\}$, $k = \{1, .., m\}$ where x, y, w and h represents top left x, y, width and height value for the rectangle. Let assume that $MLP(R_k) = A(R_k)$ where $A(R_k)$ is the validation accuracy of the MLP for a given input rectangle R_k. In order to represent the input features, a feature vector F_k is calculated in each window R_k. All possible window positions are feed into a parallel backpropogation neural network having 2 hidden layers as shown in Figure 5 to find the best $A(R_k)_b$ and worst $A(R_k)_w$ window location for a given feature window.

For each R_k, the training set is constructed by selecting equal number of images (50% training, 50% testing) from each emotion class. In order to overcome inherent disadvantages of the gradient-descent method, the batch RPROP (Resilient PROPagation) algorithm [30] which is a direct adaptive method for fast backpropogation learning is used to train the MLP. RPROP algorithm is a learning heuristics which considers only the sing of the partial derivative for the current epoch to determine the next value of the each

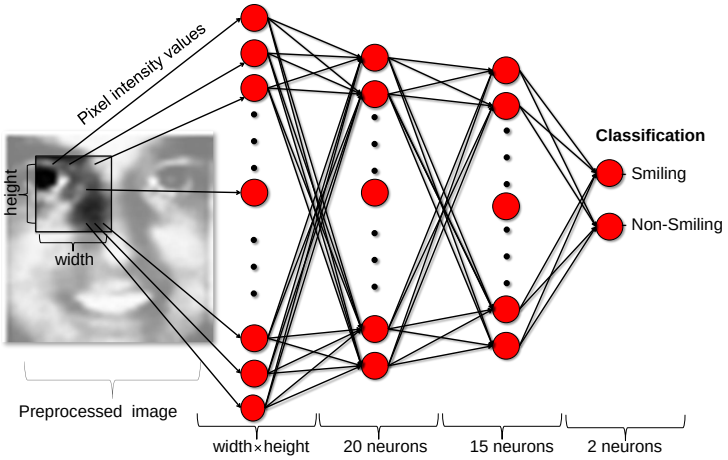


Figure 5: MLP architecture having two hidden layers. The number of neurons in the input layer is dynamically determined by the current size ($width$, $height$) of the search window.

weight. As the weight updates depend only on the sequence of the error signs and not the magnitude of the error, all weights have equal chance to learn. We selected to use the maximum number of allowed iterations and the epsilon value as 200 and 0.001 respectively for the RPROP algorithm.

Since the wrapper based feature selection methods have high computational complexity, we parallelized our code to reduce computational time. In order to speed up the exhaustive search process we employed data-parallel divide-and-conquer method for the neural network training step where each of the 8 physical processor is assigned to specific window size and position. In addition, we horizontally and vertically skipped every 3 pixels to reduce computational time.

Let $\overline{A(R_k)}_b$ and $\overline{A(R_k)}_w$ be the mean of results obtained from best and worst performing windows from a distribution with mean μ_b, μ_w and variance $V(\mu_b)$, $V(\mu_w)$ respectively. Then, the mask C is created considering the variance as shown in Eq. (5). Good results obtained in worst windows and bad results obtained in best windows are eliminated. The prediction power of the R_k is measured in terms of corresponding error rate on the training set. The mask C is then applied to normalized face images obtained in section 3.1. The approach is extendable to serve as a tool to make experimental investigation to find best feature window locations for different set

of emotions.

$$C = (A(R_k)_b > \overline{A(R_k)}_b + V(\mu_b)) \wedge (255 - (A(R_k)_w < \overline{A(R_k)}_w - V(\mu_w))) \quad (5)$$

3.3. Datasets

We used 50% of the GENKI dataset as the training set. JAFFE and FERET dataset has been used for the test and evaluation purposes. Training and test sets are strictly separated.

3.3.1. GENKI Dataset

The MPLab GENKI-4K dataset [27] is a two class database of images containing "smiling" and "non-smiling" faces spanning a wide range of illumination conditions, geographical locations, personal identity, and ethnicity. It contains 4000 labeled face images by human coders. The pose of the faces is approximately frontal.

3.3.2. JAFFE Dataset

JAFFE [31] is the Japanese Female Facial Expression dataset which contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models.

3.3.3. FERET Dataset

FERET [32] is an online Facial Expression Image Database including individuals from different genders, ethnicities and age groups with different facial expressions on their faces. The database contains 1,564 sets of images for a total of 14,126 images that includes 1,199 individuals. We used fafb subset in our experiments.

4. Results and Discussion

For a 50×50 face image we considered ($m = 224$) different search window R_k which yields 14,490 different neural networks. When all possible windows sizes feed into the neural network the outputs are the set of best and worst window locations as shown in Figure 6.

It took approximately 23 hours to process the GENKI dataset. For each R_k, we stored the x and y position giving the highest and lowest accuracy in test phase. Average training accuracy of 14,490 different neural network is 93.6% with a standard deviation of 14.1%. On the other hand, average

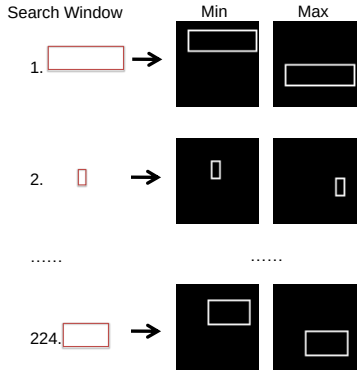


Figure 6: Minimum and maximum accuracy window location for different window sizes.

accuracy on the test set is 71.7% and standard deviation is 8.71%. 46.7% (6,773/14,490) of all neural networks performs error rate of less than 1% .

Figure 7 (A) and (B) shows 224 different R_k giving the worst and best classification results when sigmoid activation function used on 50×50 face images. Note that brighter intensity values represents higher accuracy of MLP classifier for the given R_k. Minimum values usually located on the upper face and forehead region whereas maximum values are located near the mouth and middle area of the face. This is because the positive emotions are likely to persist in the lower face more than in the upper part of the face where the negative emotions occur in the upper part of the face. In Figure 7(A), from left to right and top to bottom the classification accuracy increases from 33.1% to 85.7%. On the other hand in Figure 7 (B), these values are between 74.9% to 87.8%.

Figure 8 presents the result of Eq. (5) on GENKI dataset. By convention, we assume that a high accuracy is an indication of a valuable window and low accuracy is an indication of a worthless window. Sample use of the mask is shown on Figure 9.

According to the Figure 10, it is clear that increased window size leads to better accuracy. In order to quantify the strength of the relationship between the feature size and corresponding best accuracy value, we computed the square of the sample correlation coefficient between the feature size and their corresponding accuracy values. Approximately 64% of the variation in the accuracy can be explained by the feature size and the remaining 36% can be explained by unknown ($r^2 = 0.64$) as shown in Figure 10. Similarly, some of the smaller windows gives better accuracy than large windows which is

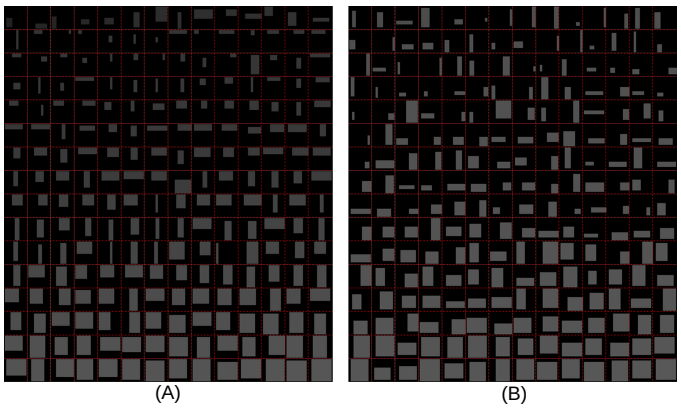


Figure 7: Worst (A) and best (B) classification weighted windows.

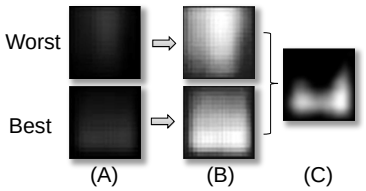


Figure 8: Classification windows weighted representation for sigmoid activation function considering the difference image. Checkerboard effect exists due to the skipped image pixels (3 pixels per step). (A) Mean images of the best and worst weighted feature windows. (B) Contrast enhanced version of (A) for visualization. (C) Substraction of worst (B) from the best (B).

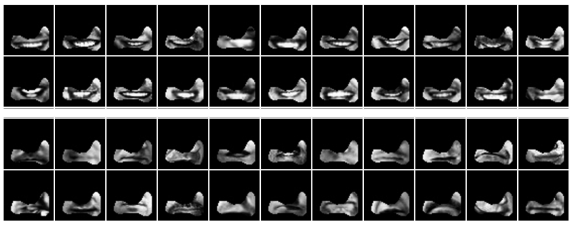


Figure 9: A subset of masked images from GENKI set for smiling and non-smiling classes using our method.

also an important indicator to solve the feature dimensionality problem.

We compared our mask shown in Figure 11 with full frame features (no mask usage) and best traditional rectangular mask obtained from R_k. We measured the performance in terms of accuracy on the test set.

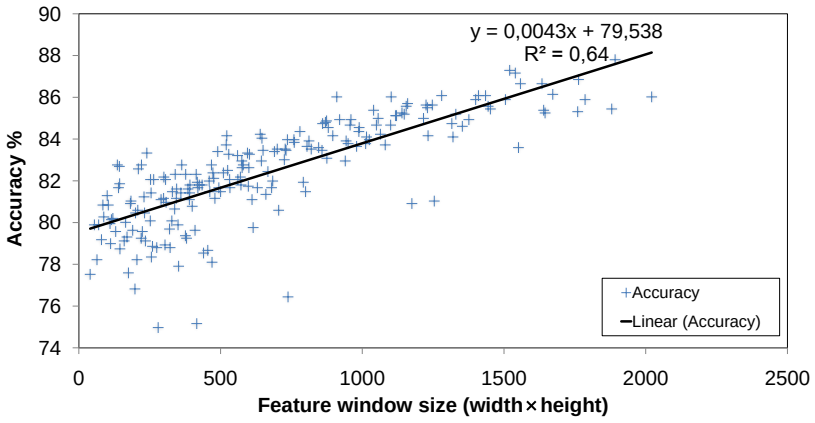


Figure 10: Effect of the window size on classification accuracy for the windows giving the best results for GENKI dataset.

According to the experimentations shown in Figure 12, the best results coming from the GENKI dataset. Since it is the largest dataset, results support our method. On each dataset our mask outperformed traditional rectangle based and full frame based approaches. Since different testing methodologies result noticeable change in classification performance, we also provide the testing methodology of other studies. Table 1 shows state of the art results for GENKI-4K and JAFFE datasets.

Figure 13 shows training size vs. accuracy for GENKI dataset. Predictive capability of the network is stable from 30% to 65% training size and provides approximately 92% of accuracy. Note that the rest is used for testing. Since we are looking for a high degree of confidence that a given set of weights will perform well on unseen data, the range from 30% to 65% training size gives the best results. Higher value (93% accuracy) at 95% training size is ignored since the validation set size is too small (5%).

[33] used graph-preserving sparse nonnegative matrix factorization (GSNMF) algorithm on GENKI-4K and JAFFE datasets. In case of JAFFE dataset they achieved 91% accuracy using 70 features and 86% accuracy for 40 features. In our case, JAFFE is a totally unseen dataset therefore we get lower results on this dataset (82% overall accuracy). Our results also support their findings that recognition results on JAFFE dataset is much lower than other datasets. For GENKI-4K dataset they performed tests using 500 test samples and report 93.1% overall accuracy for GSNMF and 91% for SNMF and

Table 1: State of the art results on JAFFE and GENKI dataset. * For more detail, please refer to referenced article. ** Training performed on GENKI dataset. No training performed on JAFFE dataset. *** Randomly select two images for each expression per person.

Author	Methodology	Dataset	Classification	Acc.	Classes
[35]	leave-one-out*	JAFFE	Gabor filters	88.1%	2
[35]	leave-one-out*	JAFFE	DNMF	85.2%	2
[36]	65% train-35% test	JAFFE	NN+DNMF	56.4%	7
[36]	65% train-35% test	JAFFE	SVM+DNMF	60.0%	7
Our method	100%** test	JAFFE	NN+Mask	82.0%	2
[33]	65%*** train-35% test	JAFFE	GSNMF	92.0%	7
[33]	25% train-75% test	GENKI	GSNMF	93.1%	2
[33]	25% train-75% test	GENKI	SNMF	92.0%	2
[33]	25% train-75% test	GENKI	Laplacian	89.9%	2
Our method	25% train-75% test	GENKI	NN+Mask	88.7%	2
Our method	50% train-50% test	GENKI	NN+Mask	92.0%	2

we obtained 92% overall accuracy using 2000 test samples.

5. Conclusion

In order to find the best feature window position and size, we performed exhaustive search on facial area. Although the bigger windows size have the higher accuracy, our experiments showed that location and size of the windows has a great effect on the emotion classification problem. Experiments showed that, in many cases, smaller feature windows have more accurate results than larger feature windows. In addition, for the same windows size there is a high accuracy difference in very close windows which supports the importance of our study. In this study we showed that proposed method improves the classification results by selecting the best pixels of interest on variety of well known databases. Since we used wrapper based method, the mask creation phase is time consuming task. However, that recognition phase using our mask runs at real-time (30 images per second for GENKI dataset). Therefore, among other methods, proposed method is an alternate solution to real-time facial expression recognition problem. A future work has been planned to create additional masks for the rest of the basic emotion classes.

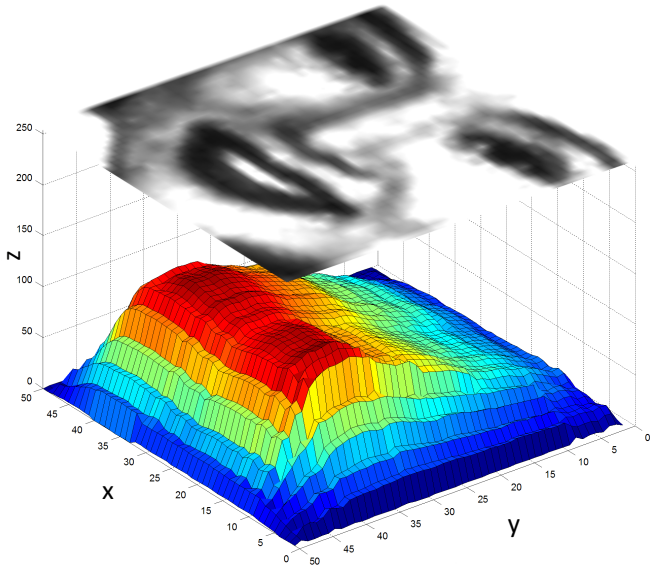


Figure 11: Pixels voting for the smiling class for GENKI dataset, x and y dimension represents the pixel position whereas z represents pixel's vote

6. Acknowledgements

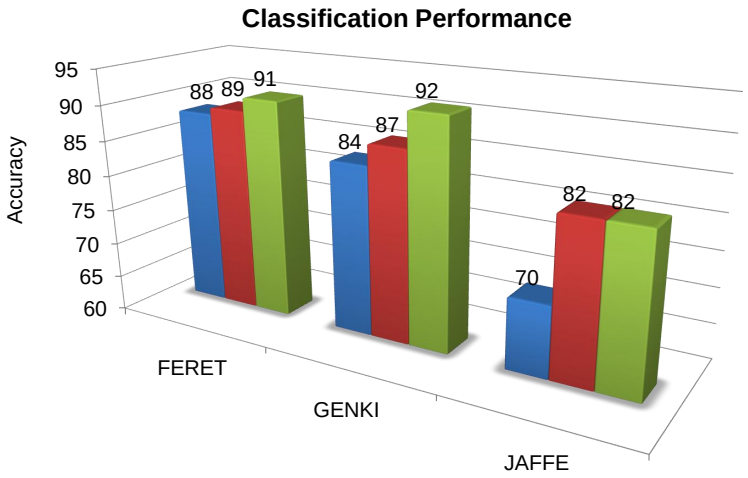
The authors would like to thank the handling editor and reviewers for their constructive comments on this paper. This study is supported by the Multimodal Interfaces for Disabled and Ageing Society (MIDAS) ITEA 2 - 07008 project.

References

[1] A. Hadid, M. Pietikinen, Combining appearance and motion for face and gender recognition from videos, Pattern Recognition 42 (2009) 2818 – 2827.

[2] P. Viola, M. J. Jones, Robust real-time face detection, International Journal of Computer Vision 57 (2004) 137–154. 10.1023/B:VISI.0000013087.49260.fb.

[3] D. Ververidis, C. Kotropoulos, Fast and accurate sequential floating forward feature selection with the bayes classifier applied to speech emotion recognition, Signal Processing 88 (2008) 2956 – 2970.



	FERET	GENKI	JAFFE
Full frame	88	84	70
Best rectangular window	89	87	82
Our mask	91	92	82

Figure 12: Experimental results on different databases.

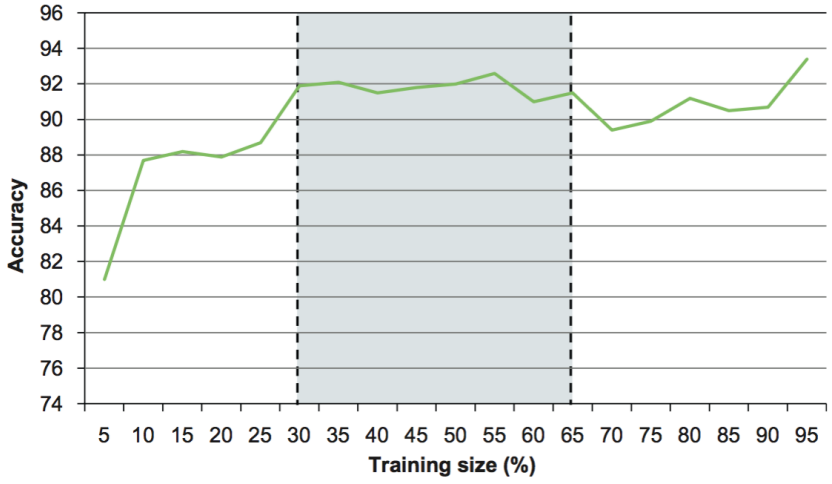


Figure 13: Training size vs. accuracy for GENKI dataset.

[4] K.-M. Lam, H. Yan, An analytic-to-holistic approach for face recognition based on a single frontal view, Pattern Analysis and Machine

Intelligence, IEEE Transactions on 20 (1998) 673 –686.

[5] T. Ojala, M. Pietikinen, D. Harwood, A comparative study of texture measures with classification based on featured distributions, Pattern Recognition 29 (1996) 51 – 59.

[6] H. Tang, B. Yin, Y. Sun, Y. Hu, 3d face recognition using local binary patterns, Signal Processing (2012) –.

[7] M. Turk, A. Pentland, Eigenfaces for recognition, Journal of Cognitive Neuroscience 3 (1991) 71–86.

[8] K. Etemad, R. Chellappa, Discriminant analysis for recognition of human face images, in: J. Bign, G. Chollet, G. Borgefors (Eds.), Audio- and Video-based Biometric Person Authentication, volume 1206 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 1997, pp. 125–142. 10.1007/BFb0015988.

[9] A. Hyvrinen, E. Oja, Independent component analysis: algorithms and applications, Neural Networks 13 (2000) 411 – 430.

[10] N. Guan, D. Tao, Z. Luo, B. Yuan, Online nonnegative matrix factorization with robust stochastic approximation, Neural Networks and Learning Systems, IEEE Transactions on PP (2012) 1.

[11] N. Guan, D. Tao, Z. Luo, B. Yuan, Nenmf: An optimal gradient method for nonnegative matrix factorization, Signal Processing, IEEE Transactions on 60 (2012) 2882 –2898.

[12] N. Guan, D. Tao, Z. Luo, B. Yuan, Manifold regularized discriminative nonnegative matrix factorization with fast gradient descent, Image Processing, IEEE Transactions on 20 (2011) 2030 –2048.

[13] S.-I. Choi, J. Oh, C.-H. Choi, C. Kim, Input variable selection for feature extraction in classification problems, Signal Processing 92 (2012) 636 – 648.

[14] S. Gong, P. W. McOwan, C. Shan, Conditional mutual information based boosting for facial expression recognition, in: British Machine Vision Conference, volume 1, Guide Share Europe, Berlin, 2005, pp. 399–408.

[15] G.-M. Jeong, H.-S. Ahn, S.-I. Choi, N. Kwak, C. Moon, Pattern recognition using feature feedback: Application to face recognition, International Journal of Control, Automation and Systems 8 (2010) 141–148. 10.1007/s12555-010-0118-7.

[16] V. Praseeda Lekshmi, S. Kumar, D. Vidyadharan, S. Naveen, Analysis of facial expressions using pca on half and full faces, in: Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on, pp. 1379 –1383.

[17] M. Song, D. Tao, Z. Liu, X. Li, M. Zhou, Image ratio features for facial expression recognition application, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 40 (2010) 779 –788.

[18] Z. Zhang, M. Lyons, M. Schuster, S. Akamatsu, Comparison between geometry-based and gabor-wavelets-based facial expression recognition using multi-layer perceptron, in: Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on, pp. 454 –459.

[19] M. E. Aroussi, M. E. Hassouni, S. Ghouzali, M. Rziza, D. Aboutajdine, Local appearance based face recognition method using block based steerable pyramid transform, Signal Processing 91 (2011) 38 – 50.

[20] M.-Y. Chen, C.-C. Chen, The contribution of the upper and lower face in happy and sad facial expression classification, Vision Research 50 (2010) 1814 – 1823.

[21] M. Hall, Feature subset selection: A correlation based filter approach, 1997.

[22] H. Liu, R. Setiono, A probabilistic approach to feature selection - a filter solution, in: ICML, pp. 319–327.

[23] R. C. Holte, Very simple classification rules perform well on most commonly used datasets, Machine Learning 11 (1993) 63–90. 10.1023/A:1022631118932.

[24] M. Pantic, L. Rothkrantz, Toward an affect-sensitive multimodal human-computer interaction, Proceedings of the IEEE 91 (2003) 1370 – 1390.

[25] M. Song, D. Tao, Z. Sun, X. Li, Visual-context boosting for eye detection, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 40 (2010) 1460 –1467.

[26] T. Danisman, I. M. Bilasco, N. Ihaddadene, C. Djeraba, Automatic facial feature detection for facial expression recognition, in: VISAPP (2), pp. 407–412.

[27] http://mplab.ucsd.edu, The MPLab GENKI Database, GENKI-4K Subset, 2011.

[28] H. Rowley, S. Baluja, T. Kanade, Neural network-based face detection, Pattern Analysis and Machine Intelligence, IEEE Transactions on 20 (1998) 23 –38.

[29] S. Milborrow, F. Nicolls, Locating facial features with an extended active shape model, in: D. Forsyth, P. Torr, A. Zisserman (Eds.), Computer Vision ECCV 2008, volume 5305 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 2008, pp. 504–513.

[30] M. Riedmiller, H. Braun, A direct adaptive method for faster backpropagation learning: the rprop algorithm, in: Neural Networks, 1993., IEEE International Conference on, pp. 586 –591 vol.1.

[31] M. Lyons, J. Budynek, S. Akamatsu, Automatic classification of single facial images, Pattern Analysis and Machine Intelligence, IEEE Transactions on 21 (1999) 1357 –1362.

[32] P. Phillips, H. Moon, S. Rizvi, P. Rauss, The feret evaluation methodology for face-recognition algorithms, Pattern Analysis and Machine Intelligence, IEEE Transactions on 22 (2000) 1090 – 1104.

[33] R. Zhi, M. Flierl, Q. Ruan, W. Kleijn, Graph-preserving sparse nonnegative matrix factorization with application to facial expression recognition, Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 41 (2011) 38 –52.

[34] Y. Shinohara, N. Otsu, Facial expression recognition using fisher weight maps, 2004.

[35] I. Kotsia, I. Buciu, I. Pitas, An analysis of facial expression recognition under partial facial image occlusion, Image and Vision Computing 26 (2008) 1052 – 1067.

[36] Z. Liang, Y. Li, T. Zhao, Projected gradient method for kernel discriminant nonnegative matrix factorization and the applications, Signal Processing 90 (2010) 2150 – 2163.