

## **ARTIFICIAL INTELLIGENCE FACIAL EXPRESSION RECOGNITION FOR EMOTION DETECTION: PERFORMANCE AND ACCEPTANCE**

*Annu George, Georgia Southern University, [ag006540@georgiasouthern.edu](mailto:ag006540@georgiasouthern.edu)  
Hayden Wimmer, Georgia Southern University, [hwimmer@georgiasouthern.edu](mailto:hwimmer@georgiasouthern.edu)  
Carl M. Rebman Jr., University of San Diego, [carlr@sandiego.edu](mailto:carlr@sandiego.edu)*

### **ABSTRACT**

The understanding of emotions has been an important research topic for many years dating back to almost 150 years to the work of Charles Darwin. Human emotion detection and recognition play an important role in everyday life as a fundamental part in interpersonal communication. Naturally, the relationship between detection and recognition is of high importance in the human computer interaction. Facial expression recognition (FER) systems are a computer-based technology that detects faces, codes expressions, and determines emotional states. This paper discusses the creation and reports the testing of a FER system to determine if machines could more accurately predict emotion than humans. In addition, this study seeks to determine perception of artificial intelligence facial expression recognition. The methods in this study included using a Haar-cascade classifier implemented by OpenCV 2.4.2. Results demonstrated that the AI FER system recognized emotions more accurately and that the human subjects considered the Artificial Intelligence Facial Expression Recognition systems trustworthy and beneficial.

Keywords: Facial Expression Analysis, Emotion expression recognition, Classifiers, Deep Learning,

### **INTRODUCTION**

Emotion plays an important role in interpersonal communication and they represent the psychological state of the human mind and thought processes. Facial expressions and movements provide nonverbal knowledge that contributes to human communication. Expression recognition involves a series of facial characteristics to be extracted from a given subject's facial expression. Recognition of emotion involves classifying facial characteristics into one of many categories of emotion. The function of identifying the emotion, therefore, is a question of pattern recognition despite the expression of an emotion.

The main objective about this research is to use facial recognition methods to detect emotions on faces. A number of applications in the area of human computer interaction can be supported by identifying facial expressions from facial images. For example, facial recognition can assist health providers detect pain or assist with mental health analyses. It includes both measurement of facial detection and recognition of expression. Facial recognition is the first step of the facial expression recognition system to find a face region in the input frame images. Upon determining the face location, various facial feature extraction approaches can be used.

We utilize the feature appearance-based method, that consist of Viola-Jones cascade object detectors. Facial expression-based emotion recognition requires extraction of a set of facial features from the facial expression of a given subject. They rely on parameters of distinctive facial features such as eyes, mouth, and nose. The last stage of the facial expression analysis system is facial expression recognition using Fisher face recognizer. Recognition of emotion here refers to classifying facial characteristics into one of many categories of emotion. Typically, to assess the category of an unknown emotional type, a supervised classifier pre-trained with emotional features is used as input and emotion class as output. The remainder of this study is organized as follow. First is discussion of relevant literature along with research questions and methodology. This discussion is followed with the results of the survey and system testing. The last section contains limitations and future research possibilities.

### **LITERATURE REVIEW**

Face detection enables the identification and validation of human faces from images, videos, and other forms of graphics. It traces facial features, contours, and texture to analyze the unique biometric and demographic details of

individuals. M.-H. Yang, Kriegman, and Ahuja (2002) proposed some methods for facial detection that included: knowledge-based methods which captures relationship between facial features, Feature-based methods which aims to find the structural features, Template-matching methods use to show standard patterns of a face which are stored to describe the face as a whole or the facial features separately, and Appearance-based models which are learned from a set of training images capturing the representative variability of the facial appearance. Each of the methods applied have their own relative performance outcomes and due to the lack of uniformity in how methods are evaluated it's hard to explicitly declare which methods indeed have the lowest error rates. As an ongoing study research area, more complex algorithms can be introduced for advance accuracy (M.-H. Yang et al., 2002).

Humans are capable to express thousands of facial expressions that vary in intensity, complexity and meaning. Kartali, Roglić, Barjaktarović, Đurić-Jovičić, and Janković (2018) proposed algorithms that performs detection, extraction, and evaluation of facial expressions and will allow for real-time emotion recognition. Its aim is to recognize the facial expression stored in a database, then recognize the human emotions in terms of happy, sad, surprise, neutral, disgust, etc. The methods proposed by the authors utilizes three deep learning approaches based on convolutional neural networks (CNN) which are (AlexNet CNN, commercial Affdex CNN solution, and custom-made FER-CNN) , and two conventional approaches for classification of Histogram of Oriented Gradients (HOG) features which are (Support Vector Machine (SVM) of HOG features, and Multilayer Perceptron (MLP) artificial neural network of HOG features).

The concept of face detection has become one of the interesting areas in research that aims to use application of pattern recognition and computer vision. Sharifara, Rahim, and Anisi (2014) proposed an advanced up-to-date method for facial recognition, that aims to increase the accuracy of detecting faces, especially in a complex environment. Some of the face detection methods used were feature-based, appearance-based, knowledge based, and template matching. Also, explained the use of applying Haar-like features and neural networks for facial recognition. Each of the methods applied have increased the efficiency for face detection but as an ongoing study research area, more complex algorithms can be introduced for advance accuracy (Sharifara et al., 2014).

Feature selection is defined as the process of selecting the subset of the most relevant features from the set of features. Ideally, four conditions should be met by a feature selection algorithm that include reliable extraction of relevant feature, identifying non-linear feature interactions, scaling linearly with the number of features and dimensions, and to allow the incorporation of known sparsity structure. In Liu and Yu (2007), the authors proposed a gradient-based feature selection for online boosting that primarily focuses on in person detection and person tracking. This method provides a distinctive scheme of learning discriminative features compared to the common way of searching the feature hypothesis space exhaustively (Liu & Yu, 2007).

Jang and Kim (2008), proposed one of the most efficient stochastic search methods known as evolutionary algorithms (EAs) mainly that focuses on evolutionary pruning method. Evolutionary algorithms have been applied in many classifier training tasks such as facial detection, facial recognition and car detection (Jang & Kim, 2008). These algorithms are known to have robust performance despite domain-specific heuristics. The main objective of this proposed method was being able to construct an efficient cascade structure to minimize the number of classifiers in AdaBoost-based cascade detector without degrading the detection accuracy. The total number of weak classifiers in the proposed structure was reduced to 58.7% of that constructed from the AdaBoost method (Jang & Kim, 2008). This algorithm was able to detect between 90.1% and 94.7% of the faces based on the database under acceptable number of false positives (Jang & Kim, 2008).

Viola and Jones (2001) proposed three main object detection framework which included integral image that allows for very fast feature evaluation, second method included construction of classifier by selecting a small number of important features using AdaBoost, and final method includes combining complex classifiers in a cascade structure which dramatically increases the speed of the detector The presented approach for object detection in this paper minimizes computation time while achieving high detection accuracy

P. Yang, Liu, and Metaxas (2007) proposed a novel approach of facial action units and expression recognition based on coded dynamical features. It utilizes the concept of Haar-like feature for facial representation, then uses the coded dynamic feature to further detect facial expressions. Their proposed method constructs a weak learner for AdaBoost

learning with one feature and uses binary coding based on statistical distribution of training samples, to make it robust to noise. Experiments on the facial expression database showed a promising performance of the proposed method

Shojaeilangari, Yau, Nandakumar, Li, and Teoh (2015) proposed an approach called extreme sparse learning to robustly recognize the facial emotions in real-world natural situations. It has the ability to jointly learn a dictionary set of basis and a nonlinear classification model. Sparse representation is a powerful tool used for reconstruction, representation, and compression of high-dimensional noisy data due to its ability to uncover important information about signals from the base elements or dictionary atoms (Shojaeilangari et al., 2015).

Yuan, Kang, Xu, Yang, and Ji (2018) proposed how deep learning methods is applied to image detection to increase the accuracy of recognition. Some of the deep learning methods used were machine learning techniques that included (Differentiated depth structure, Generative depth structure, Mixed structure), and target detection algorithm based on deep learning (Traditional target detection method, Deep learning target detection algorithm based on Region Proposal). Each of the methods applied have increased the efficiency for face detection but as an ongoing study research area, more complex algorithms can be introduced for advance accuracy such as in-depth neural network integration (Yuan et al., 2018).

Valstar and Pantic (2011), proposed some methods for facial emotion detection that included Facial point detection, and reported use of AU recognition had an accuracy of 95.3% when tested on deliberately displayed facial expressions. Mita, Kaneko, and Hori (2005) proposed a method to utilize a distinctive feature known as joint Haar-like feature for detecting faces in images. A Haar-like feature is signified by taking a rectangular part of an image or object and splitting that rectangle into multiple parts. Branching off from Haar-like feature concept, joint Haar-like feature is proposed to ensure higher classification performance. This method of face detection ensures high detection accuracy, by using a probabilistic outline or finding a discriminant function from a large set of training examples.

A central question of our era is how AI can affect the future workforce, and thus the human condition. The dominant viewpoint is that either humans or computers would win the battle between the AI and humans. In Fleuret et al. (2011), the paper conducts three experimental studies where they Let humans and state-of-the-art AI identify images together and in isolation. The AI is outperforming humans, as predicted. Through delegating to the AI, humans may strengthen but this joint effort is still not outperforming AI itself. The most successful scenario was the inversion, where when it was unclear the AI delegated it to a human. Theoretically, humans could outperform all other configurations if they delegated effectively to AI. In this research even human are bad at delegating if they make an effort to delegate well; the explanation is that with their best intentions, their understanding of task complexity is often not matched with the actual difficulty of the task when the picture is difficult. Human had no idea what they did not know (Fleuret et al., 2011).

## **RESEARCH QUESTIONS**

This study seeks to test the effectiveness of facial expression systems and to further examine issues related towards perception of artificial intelligence and computing ability to detect and determine human emotion. Specifically, the study sought to determine if the subjects perceived computing emotion recognition systems to be of higher benefit, lower risk, and higher trust. The two parts of the research include training and testing a system on images. The second part testing research subjects and asking them to fill out a survey questionnaire. The research questions are listed below. The next section describes the methodology of training, testing, and survey participants.

### **1. Can AI facial expression systems determine human emotional response than a person**

*H1: AI emotion recognition and detection will better predict the human emotional expression than a human person.*

### **2. Does using AI emotion recognition produce more perceived benefit than human emotion detections.?**

*Hpb1: AI emotion recognition and detection is more convenient than human emotion detection and recognition.*

*Hpb2: AI emotion recognition and detection saves more money than human emotion detection and recognition.*

*Hpb3: AI emotion recognition and detection save more time than human emotion detection and recognition.*

### **3. Is there higher perceived risk with trusting AI for emotion detection and recognition.**

*Hr1: AI emotion recognition and detection presents no perceived risks than human emotion detection and recognition.*  
*Hr2: AI emotion recognition and detection presents no perceived financial risks than human emotion detection and recognition.*

*Hr3: Automated AI emotion recognition and detection presents no perceived risks than human emotion detection and recognition.*

**4. Is there higher perceived trust with trusting AI for emotion detection and recognition.**

*Ht1: AI emotion recognition and detection is more trustworthy than human emotion detection and recognition.*

*Ht2: Automated AI emotion recognition and detection has better intent than human emotion detection and recognition.*

*Ht3: Using automated AI emotion recognition and detection is a better solution than human emotion detection and recognition.*

**AI FACIAL EMOTION DETECTION TESTING MODEL**

The detection and recognition implementation tested here is a supervised learning model that will be used to train and predict basic emotions. The methodology used for facial detection and emotion recognition followed several steps of training pipeline that includes training and testing mode for the data set. These methods include using classifiers such as Haarcascade classifier for facial recognition (OpenCV 2.4.2 is used to implement the Haarcascade classifier). OpenCV also has a few face recognizer classes that can be used for emotion recognition and it provides certain pre-trained classifiers for training the datasets. For this project we use test images from the Extended Cohn-Kanade dataset and projected those images onto our Haarcascade classifier and Fisher faces for classification based on our established thresholds.

**Image Data**

The Cohn-Kanade (CK) image set database was released in 2000 for promoting research that focused on facial detection of individual facial expressions. Since then, the CK database has become one of the most widely used test-beds for algorithm development and evaluation (Lucey et al., 2010). The Cohn-Kanade Facial Expression Database is made up of 100 university students aged 18 to 30 years, 65% of whom are female, 15% African-American, and 3% Asian or Latino (Jin & Lin, 2012). Subjects are advised to perform a series of 23 facial presentations, six of which are prototypic emotions that include anger, disgust happy, neutral, sadness, and surprise. The expressions are textually defined in Table 1.

**Table 1.** Basic Facial Expressions (Pandzic & Forchheimer, 2002)

<b>Expression</b>	<b>Textual Description</b>
Anger	Eyebrows are pulled downward and together. The eyes are wide open. The lips are pressed against each other or opened to expose the teeth.
Disgust	Eyebrows and eyelids are relaxed. The upper lip is raised and curled, often asymmetrically.
Happy	Eyebrows are relaxed. The mouth is open, and the mouth corners pulled back toward the ears.
Neutral	Face muscles are relaxed. Eyelids are tangent to the iris. The mouth is closed and lips are in contact.
Sadness	Inner eyebrows are bent upward. The eyes are slightly closed. The mouth is relaxed.
Surprise	Eyebrows are raised. The upper eyelids are wide open, the lower relaxed. The jaw is opened.

For this study, 311 images were selected from the database. The selection criteria was that one of the six basic emotions that could be marked as a series. From the image data set each display of the source image began and ended in a neutral face. Image sequences for frontal views and 30-degree views were digitized into either 640x490 or 640x480 pixel arrays with 8-bit gray-scale color (Lucey et al., 2010). Later these image data set are randomly classified as

training set, and testing set. Each of the images is normalized into 350 x 350 dimension after facial detection, then on the training set, we analyze the distribution of feature units under each expression using the fisher face classifier.

### Facial Detection

The overall face extraction of the image is done first using a Viola-Jones cascade object face detector. The Viola-Jones face detector contains three main ideas that allow an effective face detector to be developed that can run in real time: the integral image, classifier learning with AdaBoost, and the attentional cascade structure. Integral image, also known as a summed area table, is an algorithm to measure the number of values in a grid's rectangle subset. It can be computed from an image using a few operations per pixel to allow for a very fast feature evaluation (Viola & Jones, 2001). Next, for training purpose Viola-Jones uses a variant of Adaboost. For each subsequent weak classifier, the general idea of boosting is to fix the previous classifier's errors.

### Haar-like Features

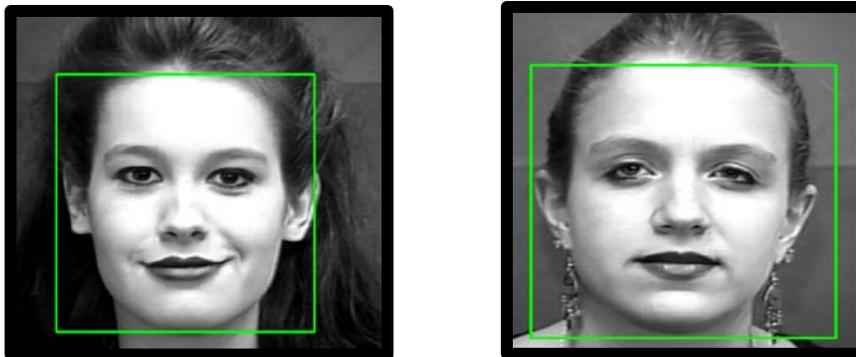
The Haar-like rectangular features are very efficient to compute due to the integral image technique and provide good performance for building frontal face detectors. Haar-like features are used to detect the presence of features in given image. The Viola-Jones face detection procedure classifies images based on the value of simple features. There are many intentions for using features rather than the pixels directly. The Viola Jones detection context seeks to identify faces or features of a face by using simple features known as Haar-like features.

```
import cv2
import glob
faceDet = cv2.CascadeClassifier("haarcascade_frontalface_default.xml")
faceDet_two = cv2.CascadeClassifier("haarcascade_frontalface_alt2.xml")
faceDet_three = cv2.CascadeClassifier("haarcascade_frontalface_alt.xml")
faceDet_four = cv2.CascadeClassifier("haarcascade_frontalface_alt_tree.xml")
```

**Figure 1.** code illustration that implements the Haar-Like feature classifier

The figure above is an example of code illustration that implements the Haar-cascade classifier. It's a classifier which is used to detect the object for which it has been trained for, from the source. It is trained by superimposing the positive image over a set of negative images. Each stage of the classifier labels the region defined as either positive or negative determined by the current location of the sliding window.

The Haar-like rectangular features are very efficient to compute due to the integral image technique and provide good performance for building frontal face detectors. The Haar-like feature method is extremely fast, as the integral value of the object in question can be measured in a single pass and a summed area table can be generated.



**Figure 2.** Facial Detection using Viola Jones Haar-like feature copyright (©Jeffrey Cohn)

These features are defined as the weighted intensity difference between two to four rectangles as shown in figure 2. The process entails passing feature boxes over an image and computing the difference of summed pixel values between

adjacent regions. In order to implement facial detection using Haar-like features, we have to utilize OpenCV's pre-trained classifiers library known as Haarcascade classifiers.

### **Fisher face Classifier**

Fisher face classifier will be used to train and test the dataset. Once the data set is organized, and all faces are detected. The next step is to split data into training and classification set. We use the training set to teach the classifier to recognize the to-be-predicted labels, and use the classification set to estimate the classifier performance. Figure 3 is an example of code illustration that implements the Fisher face classifier. Fisher face classifier is an OpenCV algorithm based on linear discriminant analysis (LDA). Linear discriminant analysis is the statistical method of classifying an observation having component in one of the two groups.

```
from image_commons import load_image

fishface = cv2.face.FisherFaceRecognizer_create()
training_set_size = 0.95
```

**Figure 3.** Code implementation for Fisher face recognizer

This classifier is also a combination of Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA) (Kirana, Wibawanto, & Herwanto, 2018). PCA preserves the details about the distribution but cannot project the optimum matrix. LDA projects the optimal matrix according to the Fisher criterion, but the input space dimension is larger than the number of training objects, so it cannot be applied directly (Kirana et al., 2018). Fisher LDA, also known as the fisher face process, is a supervised classification technique that uses class-specific information to optimize the ratio of scattering data within and between classes.

LDA controls the discriminant dimension in response pattern space, on which the ratio of between-class over within-class variance of the data is maximized. To create the best discrimination between different classes, it looks for the vectors in the underlying space. We used a one versus-all (OVA) approach for each emotion when performing LDA, where all nontarget emotion training samples are grouped. We conducted Fisher LDA for each emotion in which we want to train a predictor, in which the aim is to optimize the objective function that minimizes class variance and maximizes class variance in order to gain strong class distinction between the interest group and the other classes.

### **AI Testing Process**

The completed training implementation used the Haar-like feature of Viola Jones to detect faces, eyes, and mouth. The detected faces are cropped, resized, and mean subtracted. Using the reduced dimensionality learning data set, Fisher LDA is performed to extract Fisher faces on which data can be projected. Here we utilized the concept of Haar-like feature during the training process to detect facial features such as eyes and mouth. Based on the image set database from the Cohn-Kanade we project these images for classification. Based on the classifiers performance on our fisher faces we get an accuracy of 78%. The performance percentage alters every time when run the training data.

Figure 4 shows the performance of the classifier on the training dataset and this accuracy changes every time you run the classifier. The highest observed AI accuracy for training set was 84% (The percentage of accuracy depends on how many sets of classes of emotion you train).

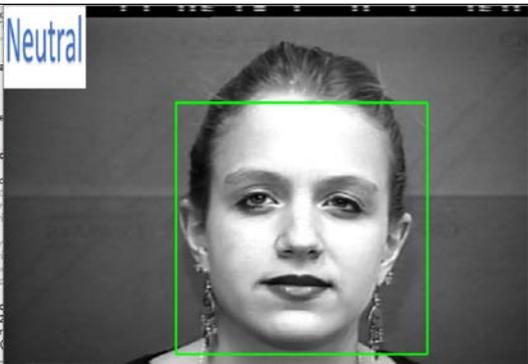
```

IPython console
Console 5/A
Reloaded modules: image_commons
size of training set is: 311 images
predicting classification set
got 78.57142857142857 percent correct!

In [16]:
    
```

**Figure 4.** Fisher face classifier Performance

The overall approach to testing was to move a test image through each of the Fisher face classifiers in order to distinguish the easiest test case. If a classifier only makes one positive prediction, the test image will be assigned as the prediction to the emotion of that Fisher face’s emotion. The data set used was very standardized. All faces are precisely pointing at the camera and in some cases the emotional expressions are actually pretty exaggerated and even comical. Figure 5 is an example of emotion recognizer output, where a test image gave the correct output of the emotion.

Face Detection	Emotion Output
	Face was detected and the correct emotion was printed: Happy
 <p>copyright (©Jeffrey Cohn)</p>	Face was detected and the correct emotion was printed: Neutral

**Figure 5.** Emotion recognizer output

### RESEARCH METHODOLOGY

This study consisted of two parts, the machine testing where a set of images were run through the AI FER system and results reported. The second part consisted of having student participants run through the same set of images. Figure 6

and 7 illustrates how the image questions were presented to the research participants. 73 students participated in the study for extra course credit and demographic data on grade level and race were collected. Of the 73 respondents, 79% were male and 21% were female. Senior level students comprised 78% of respondents, with juniors comprising 12% and the remaining 10% were graduate students. The respondent's answers were recorded and compared to the AI. The results are provided in Table 2. The students were provided information about the study's intent to determine AI Facial Expression Recognition Systems as a security tool and were asked to complete a survey regarding questions of perceived benefit, risk and trust from the Kim et al 2009 study. They are listed in table 3.



*Figure 6 and 7 Sample survey image questions. Picture copyright (@Jeffrey Cohn)*

**Table 2.** T-test on Machine Accuracy versus Human Accuracy

	<i>Human</i>	<i>Machine</i>
Mean	84.85540335	87.06240487
Variance	99.67444614	66.91940545
Observations	73.00000000	73
Pooled Variance	83.29692579	
Hypothesized Mean Difference	0.00000000	
df	144.00000000	
t Stat	-1.46094739	
P(T<=t) one-tail	0.07310422	
t Critical one-tail	1.65550418	

Perceived Benefit (BENEFIT)*	PB1: I think using AI for emotion detection is convenient. PB2: One can save money by using AI technology for emotion detection PB3: AI for Automated Emotion Detection can save time.
Perceived Risk (RISK)*	R1: Utilizing AI for automated emotion detection would possess some risk. R2: Utilizing AI for automated emotion detection would involve financial risk R3: How would you rate your overall perception of risk of AI for automated emotion detection.
Perceived Trust (TRUST)*	T1: AI for automated emotion detection is trustworthy. T2: I believe that AI for automated emotion detection has my best interests in mind. T3: Utilizing AI for automated emotion detection would be a wise idea.

### RESULTS

Our first hypothesis was that AI could better detect and determine emotion expression than humans. The results in provided in table 1 illustrate that AI score higher (87.06%) than the human participants (84.85%). According to the data, the machine's precision outperformed human accuracy. It was proposed that this happens because machines are given more data (images of emotion) to be trained. We also conducted t-test to see whether there was any significance among accuracy between the demographics. There no significance between Gender and Accuracy, Race and Accuracy, and Education level and accuracy. Statistical analyses were also conducted on the constructs from the survey and they are presented in table 4.

**Table 4.** Statistical Analysis of Constructs

			F	Sig.
Perceived Benefit	PB1	CONVENIENT	3.266	<b>0.075*</b>
	PB2	SAVE_MONEY	2.79	<b>0.033**</b>
	PB3	SAVE_TIME	1.387	0.248
Perceived Risk	PR1	OVERALL_RISK	1.044	0.391
	PR2	FIN_RISK	1.236	0.304
	PR3	OVERALL_RATE	0.31	0.870
Perceived Trust	PT1	TRUSTWORTHY	1.64	0.174
	PT2	BEST_INTERESTS	0.892	0.474
	PT3	WISE_USE	2.281	<b>.069*</b>

\*significant at .1 level

\*\* significant at .05 level

Of the other nine hypotheses only 1 were found to be statistically significant and 2 was found to be just short of statistically significant p-.069. Specifically, respondents felt money could be saved by using AI for emotion detection (p=0.033). Two other areas that were quasi significant was that subjects felt that using AI for emotion detection would be convenient (p=0.075) and wise idea (p=0.062). Apparently, the subjects did not feel strongly that AI posed any positive or negative risk and just a small level of trust.

### **LIMITATIONS AND FUTURE RESEARCH**

While this study provides evidence that AI Facial Emotion Recognition systems can be more accurate than human detection, the difference in accuracy could have been stronger. Future testing on different algorithms, training times, and training sets should be conducted to determine if better results can be obtained. The predictor is relatively successful at predicting test data from the same dataset used to train the classifiers. Nevertheless, the predictor is consistently poor at detecting the expression associated with certain emotions such as neutral and surprised. This is possibly occurring due to lack of training and testing images that clearly demonstrate, inadequate labeling for pre-trained data. The classifier also fails to predict test data emotions that have expressions that do not explicitly belong exclusively to one of the six basic expressions, because they have not been trained for other expressions. Future work should focus on enhancing classifier robustness by introducing more training images from various datasets, as well as exploring more accurate methods of detection.

This study also attempted to determine end user perception of AI FER and only produced a few insights. Future research should consider other survey questions that make stronger inquiries towards perception and acceptance. Also, while the sample size was respectable, a more robust demographic representation might help to determine if there are differences in race, gender, or education level. This study also utilized student participants who may have completed the study with less attention or been able to fully appreciate the security context. Attempts were made to remove outliers so that results obtain in this study were valid for conclusions. Lastly, it would be of considerable interest to conduct a study that could account for context in expression. Given that some people emotional appearance does not match their actual meaning (think Tears of a Clown) this would be of great importance and value.

### **CONCLUSION**

In this research we utilized the concept of image processing and classification method to detect faces and recognize emotions. The proposed method uses face images that are used to train the classifier predictor to predict six basic human emotions given a test image. Statistical analysis on machine versus human accuracy determined machine outperformed human accuracy in terms of detecting emotions in this study. Future research is needed to replicate to provide more validation of machine performance. Survey responses indicated that respondents considered AI to be a money saver. Other items of relevance, though not statistically significant was that participants felt AI could be convenient and a wise idea. These are useful results given the increase in need and adoption of computer AI facial expression detection. As the saying goes it is not enough to be doing right, it needs to be perceived as doing right.

### **REFERENCES**

- Fleuret, F., Li, T., Dubout, C., Wampler, E. K., Yantis, S., & Geman, D. (2011). Comparing machines and humans on a visual categorization test. *Proceedings of the National Academy of Sciences of the United States of America*, 108 43, 17621-17625.
- Fügener, A., Grahl, J., & Gupta, A. (2019). Collaboration and Delegation Between Humans and AI: An Experimental Investigation of the Future of Work. ERIM report series research in management Erasmus Research Institute of Management.
- Han, F., Shan, Y., Sawhney, H. S., & Kumar, R. (2008). Discovering class specific composite features through discriminative sampling with swendsen-wang cut. Paper presented at the 2008 IEEE Conference on Computer Vision and Pattern Recognition.
- Jang, J.-S., & Kim, J.-H. (2008). Fast and robust face detection using evolutionary pruning. *IEEE transactions on evolutionary computation*, 12(5), 562-571.
- Jin, D., & Lin, S. (2012). *Advances in Electronic Engineering, Communication and Management Vol. 2*. In: Springer.

- Kartali, A., Roglić, M., Barjaktarović, M., Đurić-Jovičić, M., & Janković, M. M. (2018). Real-time Algorithms for Facial Emotion Recognition: A Comparison of Different Approaches. Paper presented at the 2018 14th Symposium on Neural Networks and Applications (NEUREL).
- Kim, D., Ferrin, D., & Rao, R. (2009). Trust and Satisfaction, Two Stepping Stones for Successful E-Commerce Relationships: A Longitudinal Exploration. *20*, 237-257. doi:10.1287/isre.1080.0188
- Kirana, K. C., Wibawanto, S., & Herwanto, H. W. (2018). Emotion recognition using fisher face-based viola-jones algorithm. *Proceeding of the Electrical Engineering Computer Science and Informatics*, *5*(5), 173-177.
- Liu, X., & Yu, T. (2007). Gradient feature selection for online boosting. Paper presented at the 2007 IEEE 11th International Conference on Computer Vision.
- Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. Paper presented at the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops.
- Mita, T., Kaneko, T., & Hori, O. (2005). Joint haar-like features for face detection. Paper presented at the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1.
- Pandzic, I. S., & Forchheimer, R. (2002). MPEG-4 facial animation. The standard, implementation and applications. Chichester, England: John Wiley&Sons.
- Rankel, P., Conroy, J., Slud, E., & O'leary, D. (2011). Ranking Human and Machine Summarization Systems.
- Sharifara, A., Rahim, M. S. M., & Anisi, Y. (2014). A general review of human face detection including a study of neural networks and Haar feature-based cascade classifier in face detection. Paper presented at the 2014 International Symposium on Biometrics and Security Technologies (ISBAST).
- Shojaeilangari, S., Yau, W.-Y., Nandakumar, K., Li, J., & Teoh, E. K. (2015). Robust representation and recognition of facial emotions using extreme sparse learning. *IEEE Transactions on Image Processing*, *24*(7), 2140-2152.
- Valstar, M. F., & Pantic, M. (2011). Fully automatic recognition of the temporal phases of facial actions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *42*(1), 28-43.
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *CVPR* (1), 1(511-518), 3.
- Yang, M.-H., Kriegman, D. J., & Ahuja, N. (2002). Detecting faces in images: A survey. *IEEE Transactions on pattern analysis and machine intelligence*, *24*(1), 34-58.
- Yang, P., Liu, Q., & Metaxas, D. N. (2007). Boosting coded dynamic features for facial action units and facial expression recognition. Paper presented at the 2007 IEEE Conference on Computer Vision and Pattern Recognition.
- Yuan, N., Kang, B. H., Xu, S., Yang, W., & Ji, R. (2018). Research on Image Target Detection and Recognition Based on Deep Learning. Paper presented at the 2018 International Conference on Information Systems and Computer Aided Education (ICISCAE).