Emotion Recognition Using Convolutional Neural Networks

Shaoyuan Xu [a], Yang Cheng [a], Qian Lin [b], Jan Allebach [a]
[a] School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47906, U.S.A.
[b] HP Labs, Palo Alto, CA 94304, U.S.A.

Abstract

Emotion has an important role in daily life, as it helps people better communicate with and understand each other more efficiently. Facial expressions can be classified into 7 categories: angry, disgust, fear, happy, neutral, sad and surprise. How to detect and recognize these seven emotions has become a popular topic in the past decade. In this paper, we develop an emotion recognition system that can apply emotion recognition on both still images and real-time videos by using deep learning.

We build our own emotion recognition classification and regression system from scratch, which includes dataset collection, data preprocessing , model training and testing. Given a certain image or a real-time video, our system is able to show the classification and regression results for all of the 7 emotions. The proposed system is tested on 2 different datasets, and achieved an accuracy of over 80%. Moreover, the result obtained from real-time testing proves the feasibility of implementing convolutional neural networks in real time to detect emotions accurately and efficiently.

1 Introduction

As one of the most important features of human beings, emotion helps people communicate with and understand each other more efficiently. Therefore, detecting and recognizing various emotions has always been a popular topic. People detect emotions through different methods, such as voice intonation, body language and even electroencephalography [1]. However, the most intuitive and practical way of detecting and recognizing emotions is still through facial expressions. In this paper, we propose a system to detect emotions by examining facial expressions. In our system, we follow the research work proposed by Paul Ekman [2], where the emotions are categorized into 7 classes: angry, disgust, fear, happy,neutral, sad and surprise, except that the category neutral is replaced with contempt.

There has been a lot of research work on emotion recognition, most of which uses traditional computer vision methods, such as LBP [3]; and machine learning classification methods, such as SVM [4]. However, satisfying results could not be achieved due to the limitations of these methods, such as inadaptability to the change of facial muscles. Therefore, we have put much effort in investigating a new approach that take advantages of deep learning[1].

There is also substantial research work done on emotion recognition using deep learning such as traditional model training methods using a specific network [5], or combining deep learning with machine learning such as LBP [6]. Although they obtain comparably high accuracy, there are two aspects that need to be improved. Firstly, most of them use traditional network structures such as VGG Net, Alex Net, or Google Net (including the improved versions of these network structures). This results in a large model size; so that it is extremely difficult to do real time emotion recognition. Secondly, most of the proposed systems only consider the classification scenario where the intensity information is missing in the results. But in practical usage, intensity information is as important as the classification result, because we want to know not only what emotions people have, but also the level of those emotions. In this paper, we solve these two problems by selecting an appropriate network structure for an accurate real-time emotion recognition. At the same time, we extend our classification results to the regression scenario so that the intensity information can be concluded from the results. We trained our emotion recognition model for both the classification scenario and the regression scenario. Figure 1 shows the flowchart of our emotion recognition project.

The paper is organized as follows. In the rest of Section 1, we introduce the seven classes of emotions and an overview of our approaches. Sections 2 and 3 describe how our emotion recognition system is trained using two different models, a classification model and a regression model. The conclusion is provided in Section 4.

2 Seven Classes of Emotions

There are seven universal emotions: angry, disgust, fear, happy, neutral, sad and surprise. Examples of those emotions are shown in Figure 2 [7] [8] and each emotion is described by some characteristics [9].

- Angry: eyebrows are pulled down, upper eyelids are pulled up, lower eyelids are pulled up, margins of lips are rolled in and lips may be tightened.
- Disgust: eyebrows are pulled down, nose is wrinkled and upper lips are pulled up and loose.
- Fear: eyebrows are pulled up, upper eyelids are pulled up and mouth is stretched.
- Happy: muscles around the eyes are tightened, "Crows Feet" wrinkles appear around eyes, cheeks are raised and lip corners are raised diagonally.
- Sad: inner corners of eyebrows are raised, eyelids are loose and lip corners are pulled down.
- Surprise: entire eyebrows are pulled up, eyelids are pulled up and mouth hangs open.
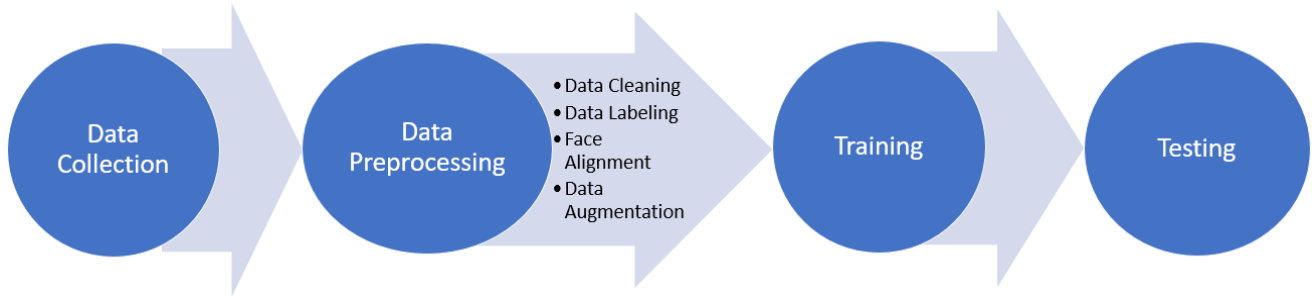

Figure 1: Flowchart of Emotion Recognition Project.

Figure 2: Seven universal emotions: Neutral, Angry, Disgust, Fear, Happy, Sad and Surprise.

3 Emotion Recognition Classification Training

Data Collection

Due to the lack of public datasets for emotion recognition tasks and the low quality of existing datasets, collecting enough datasets and examining them becomes the first challenging task. Firstly, there are 4 publicly available datasets: MUG-FED [10], CK+ [7] [8], Japanese Female Facial Expression (Jaffe) [11] and KDEF [12]. Figure 3 shows some sample images from these four datasets, and Table 1 contains the statistics. Secondly, we have collected our own dataset for testing.

Figure 3: Sample images of MUG-FED, CK+, Jaffe, and KDEF.

Data Preprocessing

Dataset Cleaning

Since most of the public datasets contain raw images, very few of them can be directly used without further examination. Therefore, these datasets need to be cleaned in the first place.

There are 52 subjects in the MUG-FED Dataset. For each subject, it has 5-7 emotions and for each emotion, it has 3-7 attempts. And since all of the images are video frames, each emotion starts from neutral to the emotional expression of the strongest intensity and returns to neutral. Therefore, only the images that contain facial expressions of strong intensity should be selected. Table 2 shows the statistics of the MUG-FED Dataset after it is cleaned. It also provides us with 161 manually labeled images, which is used for validation. Besides these 4 datasets obtained from online sources, an additional 490 images were collected by us and are used to validate the model.

Eventually, there are 3 datasets for training and 3 datasets for validation, as shown in Table 3.

Face Alignment

Face alignment is another key step in dataset pre-processing. The purpose is to remove potential uncertainties when applying our emotion recognition approach to real-time videos. For example, the position and the angle of the subject's head are changing as the video plays, which could affect the accuracy of the classification results if the face is not aligned in advance. With face alignment, the position of the head is aligned and the scale of the head is adjusted to have the same size, which eliminates the influence of any existing distortions on the recognition results.

We propose a novel face alignment algorithm that shows superior results compared to any existing method. Firstly, a face detector is used to detect the face in an image, then the Land Mark (LM) detector [13] is used to detect 68 landmark points of the face. A rotation matrix is then obtained based only the eye center coordinates. The traditional method uses the rotation matrix for face alignment. However, the resultant images can contain comparably useless background of large area and the eyes in different images are not at the same horizontal level, resulting in unsatisfying face classification results.

We improve the traditional face alignment by adding one more step. The aforementioned rotation matrix gets the coordinates of the 68 landmark points in the new coordinate system. Then we use the 1st, 9th and 17th landmark points to get the left, bottom and right boundary of the face. The definition and the location of these facial landmark points are shown in Figure 4 [14] [15], [16]. To get the top boundary, we stipulate that the length from the top boundary to the eye center is one-third of the height of the image. We use the boundary information to crop the original image into the one that contains smaller margins. Finally, the eyes of different images are adjusted to be at the same horizontal level. Figure 5 shows some sample images before and after face alignment. Note that all the images after face alignment are re-scaled to 128×128.

Data Augmentation

To increase the robustness of our model and to prevent it from being over-fitted, we apply data augmentation on the training dataset after face alignment. For each image, 7 images of different brightness and 28 images of different degree of blurring are created, resulting in a final training set with 1,148,812 images. Table 4 shows the statistics of the data augmentation.

Table 1: Statistics of the 4 collected datasets.

Database	Facial Expression	# of Subjects	# of Images	Gray/Color	Size (pixel)	Ground Truth	Type
Extended Cohn-Kanade Dataset (CK+)	Angry, Contempt, Disgust, Fear, Happy, Neutral, Sad and Surprise	123	593 image sequences (327 sequences having discrete emotion labels)	Mostly gray	640×490	Facial expression labels and FACS labels	Posed; spontaneous smiles
Japanese Female Facial Expression (Jaffe)	Angry, Disgust, Fear, Happy, Neutral, Sad and Surprise	10	213 static images	Gray	256×256	Facial expression labels	Posed
Multimedia Understanding Group (MUG-FED)	Angry, Disgust, Fear, Happy, Neutral, Sad and Surprise	86	1462 sequences with more than 100K images	Color	896×896	Facial expression and land mark (LM) labels	Posed
The Karolinska Directed Emotional Faces (KDEF)	Angry, Disgust, Fear, Happy, Neutral, Sad and Surprise	70	4900 images	Color	562×762	Facial expression labels	Posed

Table 2: Statistics of the MUG-FED Dataset after dataset cleaning up.

Emotion Type	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise	Total
# of Images	6220	4856	4605	9329	3719	5562	5623	39914

Table 3: Training and Validation datasets for classification training.

Dataset	Training	Validation
Name	MUG-FED, CK+ and Jaffe	MUG-FED (Manually Labeled by author), KDEF and Images of myself
# of Images	41029	1867

Model Training

In order to apply our emotion recognition system to real-time video, the model needs to be comparably small in size and fast in speed. We have tested several pre-traind models, such as the VGG-S [17], on real-time video with multiple rounds of fine-tuning. The validation accuracies are less than 60% which is far below our requirements. Moreover, the size of the VGG-S model is more than 500 MB which is too large to be implemented efficiently.

To reduce the model size, we modified the original VGG-S [18] model by reducing the kernel size and channel number as shown in Figure 6 [13]. Compared to the original VGG-S model, our model has a size of only 12.1 MB. It takes only 4.5 hours to train on more than 1 million images for 50,000 iterations. Besides a smaller size of the model, the validation accuracy obtained by using the new model reaches 85%, which is significantly higher than our previous results. The reason of getting higher accuracy
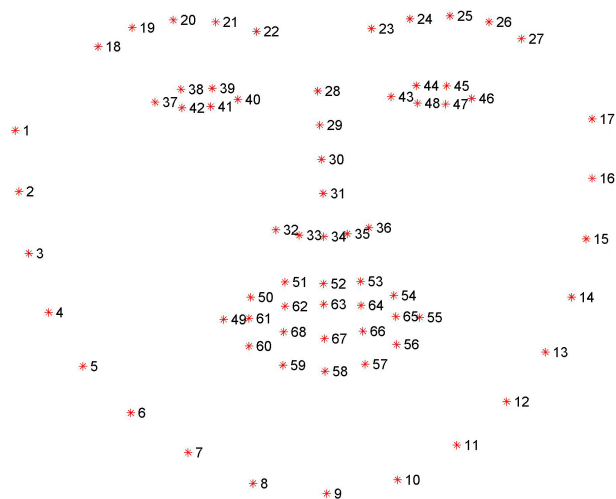

Figure 4: 68 points facial landmark system.

with a model of smaller size is that, if we want to train with the original large-size VGG-S model, it requires millions of raw images and weeks of time to train from scratch which is impossible. Which means, with our training set, smaller model will have higher accuracy. And the modified model can be trained faster and more efficiently with much smaller data set and much less time, in our case, 40k images and 4.5 hours.

Table 4: Data augmentation of classification training dataset.

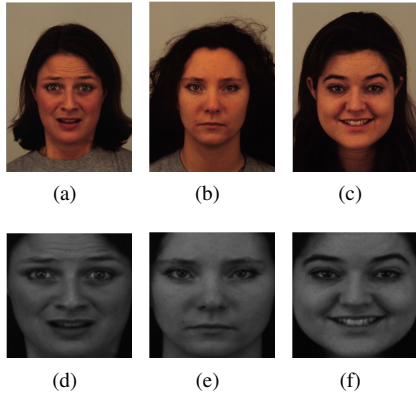| # of Images Before Face Alignment	Brightness Change	Blur (Gaussian, Average, Median)	Total Multiples	# of Images After Face Alignment
41029	7 ×	4 ×	28 ×	1148812

(a)   (b)   (c)

(d)   (e)   (f)

Figure 5: Comparison of images before and after face alignment.

4 Emotion Recognition Regression Training

Introduction

Although our emotion recognition classification model works well, it has its own drawback. And it is especially obvious when the classification model is applied in a real-time demo.

Since our classification training dataset includes a lot of emotions that are not obvious or are of low intensity, this making them similar to one of the emotion categories in particular: neutral. This causes the prediction on the real-time video to be jittery, since in most cases, for example, a person does not need to express his or her happiness by a drop-jaw smile. And also, given an image or a frame of the video, our classification model can only tell if the facial expression is angry, disgust, fear, happy, neutral, sad or surprise; in other words, it is not able to tell the intensity of the emotion.

To solve these two problems, a regression model is used, where the ground truth labels become the intensities of the emotion, such as 20% happy and 80% neutral or 40% sad and 60% neutral. This additional information about the intensity of the emotion can be useful, especially in real-time videos.

Data Collection

Among the four datasets collected from the online sources, only the MUG-FED Dataset is used because of the large number of images the dataset includes. However, the MUG-FED Dataset is more like an "in-the-lab" dataset, where all of the emotions included are standardized and all the images have the same background and consist of a purely frontal face. Since there are very few public in-the-wild datasets, especially for the task of emotion recognition, we collected our own dataset to train the model on a more "in-the-wild" dataset.

Until now, we have collected more than 7000 "in-the-wild" images containing facial expressions and we name this dataset as Emotion Intensity in the Wild Dataset. Table 5 shows the statistics of this dataset. It is worth noting that this dataset includes images containing heads at different angles, people with different races and ages, and backgrounds of different lighting conditions.

To train the regression model, we use both the MUG-FED and Emotion Intensity In the Wild datasets.

Data Preprocessing

Dataset Cleaning

Before an existing dataset obtained from online sources is used, it needs to be examined in the first place. As we introduced in the previous section, the images of the MUG-FED Dataset are consecutive frames obtained from videos. In a video for a specific emotion, the emotion starts from neutral to 100% facial expression and gradually returns to neutral. Therefore, for each attempt of expressing emotion, we select 9 images that contain facial expression intensities from 20% to 100% and back to 20% with an interval of 20%. An example is shown in Figure 7. After dataset cleaning, we have collected 7,451 images for training and 981 images for validation for the MUG-FED Dataset. Each of these 7451 images is labeled with the intensity of the emotion.

And for the Emotion Intensity In the Wild Dataset, after excluding some inappropriate images, we have 6141 images for training and 682 images for validation.

Face Alignment

The procedure of face alignment is the same as the one introduced in the previous section. We utilized the 1st, 9th and 17th landmark points to get the boundary of the faces, cropped them, aligned them and rescaled them to 128×128.

Data Augmentation

We experimented with two strategies of training, one with only the Emotion Intensity In The Wild Dataset, another with the combined dataset of Emotion Intensity In The Wild Dataset and the MUG-FED Dataset. The method of data augmentation remains the same, which includes changing brightness and blurring the images and gives the final training dataset, contains 6,141 images before data augmentation and 171,948 images after, and final validation dataset, containing 682 images before data augmentation and 19,096 images after for Emotion Intensity In The Wild Dataset. For the combined dataset, the final training dataset, containing 13,592 images before data augmentation and 380,576 images after, and the final validation dataset, containing 1,663 images before data augmentation and 46,564 images after, as shown in Table 7.

Model Training

The model framework used for the regression training is the same as for our classification model, except that the softmax loss function is replaced with the sigmoid cross entropy loss function.
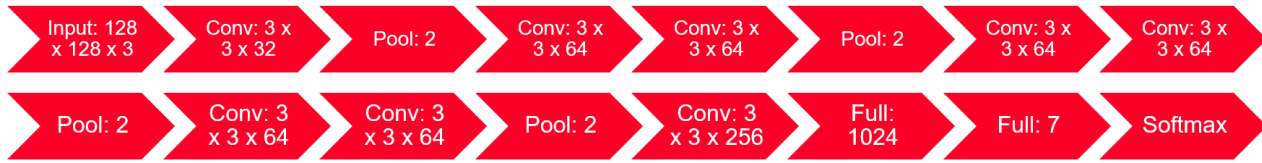
Figure 6: Framework of the classification model.

Table 5: Statistics for Emotion Intensity In the Wild Dataset.

Dataset	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise	Total (Without Neutral)
20%	194	147	93	236	1093	210	91	971
40%	323	218	103	230		164	218	1256
60%	221	243	134	320		137	279	1334
80%	207	198	151	295		81	316	1248
100%	227	178	157	286		84	420	1352
Total	1172	984	638	1367	1093	676	1324	6161 (7254 Total With Neutral)

Table 6: Regression training results.

Dataset	Training Loss	Validation Loss	Regression RMSE	Classification Accuracy
Emotion Intensity In the Wild	0.13	0.3	0.129	77.2%
Combined Dataset	0.122	0.239	0.123	76%

Table 7: Dataset statistics for regression training.

Dataset	Training	Validation
Emotion Intensity In the Wild (After data augmentation)	6141 (171948)	682 (19096)
Combined Dataset (After data augmentation)	13592 (380576)	1663 (46564)

5 Experimental Results

Classification Results

As introduced in Section 3, the classification validation accuracy of the classification model is 85%. Figure 8 shows the validation confusion matrix for the validation dataset.

In order to test our classification model, we collected our own dataset which is called the HP Facial Expression Test Set, which contains 2443 images. The dataset was collected with 5 subjects doing 7 emotions while being video recorded and the images were selected from the video frames. Our model achieves an accuracy of 82% and takes only 13.68 seconds to test on the whole testing dataset (0.0056 s/image). This test was conducted on a workstation with an Nvidia Titan X GPU. Figure 9 shows some sample testing images and Figure 10 shows the testing confusion matrix.

Regression Results

The regression training also gets outstanding results. Figure 11 and Figure 12 shows the classification confusion matrices; and Table 6 shows the regression training results. Noting that, for the training and validation loss values, they are sigmoid cross entropy loss, the smaller the better and for the RMSE values, they represent the standard deviation of the prediction errors and are based on the datum that ranges from 0 to 1.

As indicated by the high accuracy which is around 77% and the small Root Mean Squared Error (RMSE) value which is below 0.13, our regression model performs well on the emotion recognition task.

Real-time Emotion Recognition

Currently, there are not many real-time emotion recognition frameworks, while the existing ones achieve only comparably low accuracy. However, our real-time demo version can detect people's frontal facial expressions accurately. Figure 13 shows some sample results from our real-time emotion recognition demo.

6 Conclusion

In this paper, we apply emotion recognition using a deep learning method. The whole process includes data collection, data preprocessing, model training and model testing. Our contributions include:

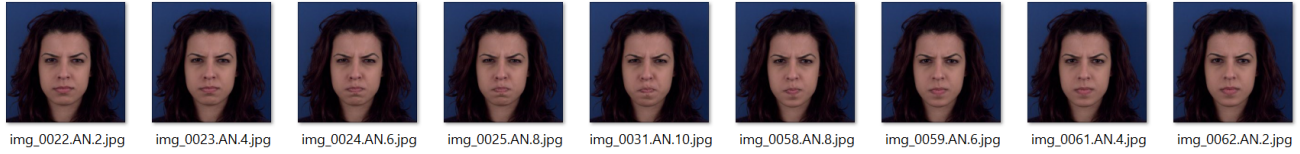| img_0022.AN.2.jpg | img_0023.AN.4.jpg | img_0024.AN.6.jpg | img_0025.AN.8.jpg | img_0031.AN.10.jpg | img_0058.AN.8.jpg | img_0059.AN.6.jpg | img_0061.AN.4.jpg | img_0062.AN.2.jpg |

Figure 7: Sample regression images from MUG-FED Dataset. These images are from one of the attempts that a subject does which have the intensities from 20% (neutral) to 100%, and back to 20% in steps of 20%. Noting that the numbers: 2, 4, ... after the emotion label *AN* are the intensity labels, 2 corresponds to 20% etc.
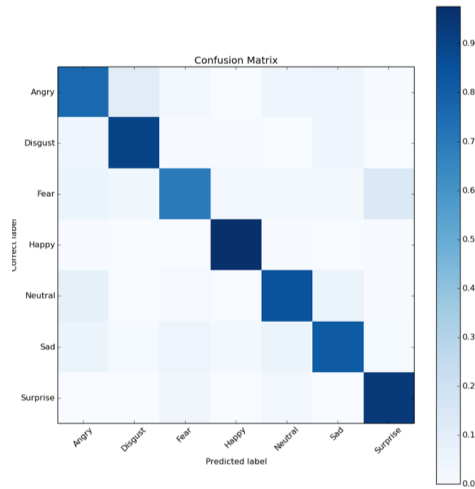

Figure 8: Classification confusion matrix for our classification model validation set. The overall accuracy is 85%.
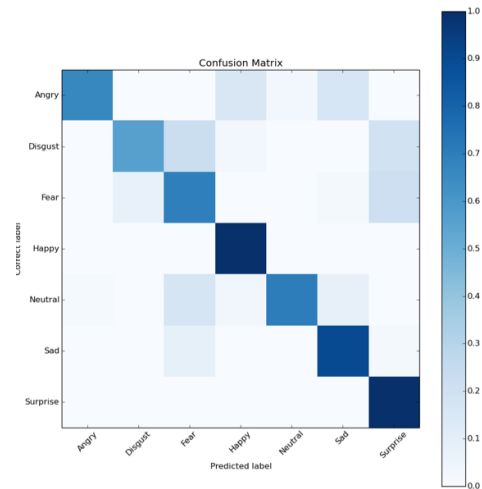
Figure 10: Classification confusion matrix for our self-collected HP Facial Expression Test Set. The overall accuracy is 82%.

- We showed that our model achieved accurate and smooth real-time recognition of both emotion type and intensity by applying it to real-time scenarios.

References

[1] P. Abhang, S. Rao, B. W. Gawali, and P. Rokade, "Article: Emotion recognition using speech and EEG signal a review," *International Journal of Computer Applications*, vol. 15, pp. 37–40, Feb. 2011.

[2] P. Ekman and W. V. Friesen, "Universals and cultural differences in the judgements of facial expressions of emotion," *Journal of Personality and Social Psychology*, vol. 53, 1987.

[3] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, 2009.

[4] S. Xu, C. Lu, M. Shaw, P. Bauer, and J. Allebach, "Page classification for print imaging pipeline," in *Color Imaging XXII: Displaying, Processing, Hardcopy, and Applications, (Part of IS&T Electronic Imaging 2017)*, vol. 2017, pp. 137–142, 01 2017.

[5] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, vol. 00, pp. 1–10, March 2016.

[6] G. Levi and T. Hassner, "Emotion recognition in the wild via convolutional neural networks and mapped binary patterns," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, ICMI '15, pp. 503–510, 2015.

[7] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for fa-

(a)

(b)

(c)

(d)

(e)

Figure 9: Sample images from HP Facial Expression Test Set.

- We built an emotion recognition framework from scratch. We first collected four public datasets and manually cleaned them. After that, we implemented data preprocessing, including labeling data, aligning faces and augmenting data. In the training process, we designed our own model based on a VGG-S model but with a much smaller size, better accuracy and improved efficiency.

- We developed an emotion recognition regression training framework to consider the intensity information of emotions. We collected our own Emotion Intensity In the Wild Dataset and defined a 5-level regression labeling scenario. Our experiment results show that the proposed system can recognize the emotion intensities with promising accuracies.
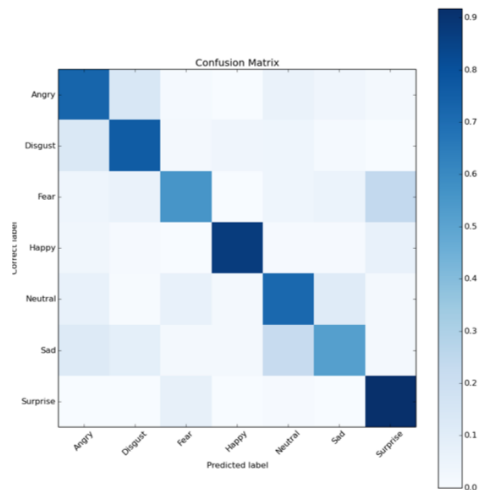
Figure 11: Classification confusion matrix for Emotion Intensity In the Wild Dataset. The overall accuracy is 77.2%.

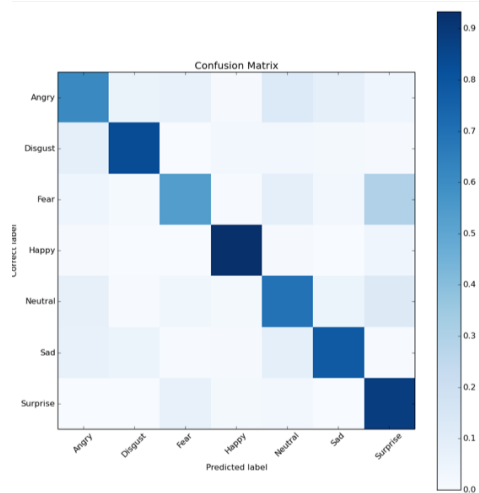Figure 12: Classification confusion matrix for the Combined Dataset. The overall accuracy is 76%.
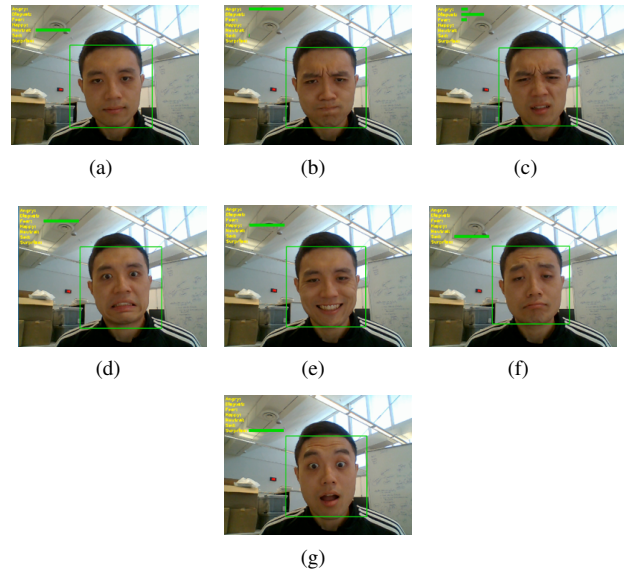

Figure 13: Sample result frames from real-time video demo for all seven emotions. The reader should zoom in to be able to see the labeling of the emotion and intensity provided in the upper left corner of each frame.

cial expression analysis," in *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 46–53, 2000.

[8] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete expression dataset for action unit and emotion-specified expression," in *Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis*, pp. 94–101, 2010.

[9] Humintell, *The Seven Basic Emotions: Do you know them?*, Jun. 2016.

[10] N. Aifanti, C. Papachristou, and A. Delopoulos, "The MUG facial expression database," in *Proc. 11th Int. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, Apr. 2010.

[11] M. J. Lyons, S. A. ad Miyuki Kamachi, and J. Gyoba, "Coding facial expressions with Gabor wavelets," in *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200–205, 1998.

[12] D. Lundqvist, A. Flykt, and A. hman, "The Karolinska directed emotional faces - KDEF," in *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet*, 1998.

[13] R. Mao, Q. Lin, and J. P. Allebach, "CNN based facial landmark detection," in *Imaging and Multimedia Analytics in a Web and Mobile World 2018, (Part of IS&T Electronic Imaging 2018)*, 01 2018.

[14] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: Database and results," in *Image and Vision Computing (IMAVIS), Special Issue on Facial Landmark Localisation "In-The-Wild"*, 2016.

[15] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "A semi-automatic methodology for facial landmark annotation," in *Proceedings of IEEE Intl Conf. Computer Vision and Pattern Recognition (CVPR-W), 5th Workshop on Analysis and Modeling of Faces and Gestures (AMFG 2013)*, 2013.

[16] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: The first facial landmark localization challenge," in *Proceedings of IEEE Intl Conf. on Computer Vision (ICCV-W), 300 Faces in-the-Wild Challenge (300-W)*, 2013.

[17] G. Levi and T. Hassner, "Emotion recognition in the wild via convolutional neural networks and mapped binary patterns," in *Proc. ACM International Conference on Multimodal Interaction (ICMI)*, Nov. 2015.

[18] A. Z. K. Simonyan, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.

Authors' Biographies

Shaoyuan Xu

Shaoyuan Xu received his Bachelor of Science degree in Electrical Engineering from University of Illinois at Urbana-Champaign in 2015. He was admitted by Purdue as a master student. Then he joined Professor Allebach's team in Spring 2016 and transfered to direct Ph.D. in Fall 2016. After he joined the team, he started to work on projects in the area of im-

age processing and image quality. In Summer 2016, he went to HP Boise for a summer internship working on the page classification project. From Summer 2017 till now, he works at HP Labs in Palo Alto on the emotion recognition project. Besides image processing, image quality and deep learning, his research interests also include computer vision and machine learning.

Yang Cheng

Yang Cheng was born in Chongqing, China, in 1995. He received his Bachelor of Science degree in Electrical Engineering from Purdue University in 2018, and decided to continue pursuing a Ph.D. in the same area at Purdue University. Currently, he works with HP Labs in Palo Alto on the facial landmark points detection project, under the supervision of Dr. Jan Allebach and Dr. Qian Lin. His research focus is on image processing, computer vision and deep learning.

Qian Lin

Dr. Qian Lin is a distinguished technologist working on computer vision and deep learning research in HP Labs. She is also an adjunct professor at Purdue University, supervising PhD students in their deep learning research. Dr. Lin joined Hewlett-Packard Company in 1992. She received her BS from Xian Jiaotong University in China, her MSEE from Purdue University, and her Ph.D. in Electrical Engineering from Stanford University. Dr. Lin is inventor/co-inventor for 44 issued patents. She was awarded Fellowship by the Society of Imaging Science and Technology (IS&T) in 2012, and Outstanding Electrical Engineer by the School of Electrical and Computer Engineering of Purdue University in 2013.

Jan P. Allebach

Dr. Jan P. Allebach received his B.S. from the University of Delaware in 1972, his M.S. from Princeton University in 1975 and his PhD from Princeton University in 1976. He is now the Hewlett-Packard Distinguished Professor of Electrical and Computer Engineering at Purdue University. Allebach is a Fellow of the IEEE, the National Academy of Inventors, the Society for Imaging Science and Technology (IST), and SPIE. He was named Electronic Imaging Scientist of the Year by IS&T and SPIE, and was named Honorary Member of IST, the highest award that IST bestows. He has received the IEEE Daniel E. Noble Award, and is a member of the National Academy of Engineering. He recently served as an IEEE Signal Processing Society Distinguished Lecturer (2016-2017). His current research interests include image rendering, image quality, color imaging and color measurement, printer and sensor 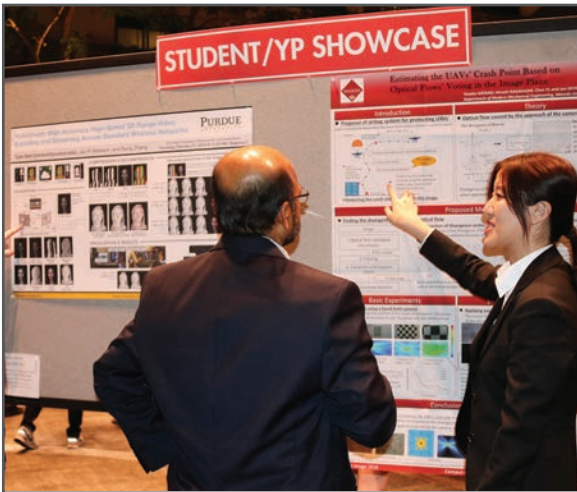forensics, and digital publishing.