International Journal of Control and Automation
Vol. 13, No. 2, (2020), pp. 309 - 317

A Wide-Ranging View of Face Emotion Recognition System

Preet Navdeep[1] , Dr. Neeraj Sharma[2] , Dr. Manish Arora[3]
[1.] *Research scholar, IKGPTU, Jalandhar*

[2.] *Dean and Professor, GJIMT, Mohali*

[3.] *Additional Director, NIELIT, Ropar*

[1] *preet.sra@gmail.com*, [2] *nrjsharma@yahoo.com*, [3] *manisharora@nielit.gov.in*

Abstract

Facial expressions contribute in revealing the emotional state of an individual. For an effective communication, facial expressions act as vital means of unveiling intentions and emotions to others. The system has an immense popularity in the area of fraud detection, health care and gaming. Although, many techniques have been developed to recognize facial emotion, but still there exist challenges in applicability of facial expression detection system. In this paper, we present panoramic view of the existing approaches and methods for correctly recognizing the facial emotion of humans. Main Phases for carrying out facial emotion recognition process includes feature extraction and classification. Various types of emotions are studied in this paper.

Keywords: Facial Emotion Recognition, facial expressions, RNN, CNN,

1. Introduction

Unlike other animals, Humans or Homo sapiens have the power or the capability to express themselves be it physically or emotionally. The central notion of being human is the capability of expressing one's emotions [1]. Humans can express mainly through speech and gestures, these can be hand gestures, facial expressions or other body gestures. The best way of communication is undoubtedly speech but, the primary way to express a person's emotional or mental state is through facial expressions. In some cases when we have some image data or some videos that need to be analysed, facial expressions are the main factor that needs to be focused upon because different facial expressions denote an individual's inner intentions or emotional states [2]. Facial appearances provide evidence on the intentions and characters of people. The first noted investigation on examining facial expressions was reported early in 1862 [1].

The recognition of facial expressions has become significant due to its potential applicability to deal with fraud of passports, cases of missing individuals, and to support law enforcement [22]. Some applications and methods use this kind of system to analyze an individual's feelings, opinions, and nature. But, there are some issues related to illumination, pose, and quality of an image. The basic system is categorized into four phases: a) Face detection aids in detecting the face region, b) Normalization to reduce the effects of lighting, c) Feature extraction to extract the necessary features, and d) Classification to categorize features as emotional states [21]. This paper provides an insight into the existing techniques of recognizing the emotions, their comparison, benefits and issues faced by the current FER systems.

ISSN: 2005-4297 IJCA

2. Motivation

Emotions are a major component of being human [1]. Faces are one of the most relevant social inducements as they transfer essential information while interacting with others. Particularly, the facial expressions convey about what an emotional state a person is going through currently. Some movements are observed during facial expressions which are basically due to the contraction and expansion of muscles beneath the skin. These movements basically express a non-verbal form of communication to the spectator expressing the emotional state of the person. [18]. The Facial Emotion Recognition (FER) System deals in recognizing or classifying the emotional state of an individual correctly and efficiently. Recognition of emotions has large applicability in the field of human-computer interaction (HCI) such as affective computing, online interactive gaming, healthcare, etc. [1]. The main aim of FER is to transform the emotions of an individual into a piece of valuable information which may further help in understanding the behaviour of that particular individual [15]. Understanding the emotional state of a person is very significant for providing appropriate services to the customers or treatment in the case of medical science. Also, it is essential for getting relevant feedback. Consequently, more advanced FER systems should be developed.

3. Background Study

Abdulsalam et al. in [1] suggested a Deep CNN network of three convolutional layers, two fully connected layers (FCs), BN and a Rectified Linear unit for the identification of emotions. The system was conducted on ten emotions including the emotions of neutral, pride, contempt and embarrassment in addition to the six basic emotions. It was performed in two phases: i) Training, and ii) Testing. The pre-processing carried histogram equalization, resizing and converting the image to grayscale. Chen et al. [2] proposed a facial expression identification system based on kernel sparse representation for efficient classification of facial expression. The proposed method was evaluated on JAFFE dataset containing 213 images which involved 7 basic emotions. The system achieved overall accuracy of 85.71%. It was observed that the proposed system outperform other methods such as SRC, SVM etc. Hossain et al. [3] proposed an audio video emotion recognition system using Deep learning method which uses SVM for the final classification of the emotion. The process involves pre-processing and frame selection of speech and video signals respectively which are then fed to the corresponding 2D and 3D CNN networks. Then these are passed through extreme learning machines (ELM) that calculates the probability distribution. Finally, SVM helps in classifying the emotions. Zadeh et al. [5] proposed a system that uses Gabor filters as a feature extraction technique and Convolutional Neural Networks for classification. The process is as follows: a) Resizing, b) First Gabor filter, c) Second Gabor filter, d) CNN and e) Emotions are recognized. Using Gabor filters improves the learning speed as it extracts sub features supplying them to CNN. The seven basic expressions of anger, happiness, disgust, fear, neutral, sadness and surprise are considered. Li et al. [6] proposed a technique for facial expression classification using key points that helps in extracting features on the face. The key points are the landmarks for locating facial properties which subsequently enhanced the classification of emotion. The geometric and textual features were then integrated with SVM classifier and the results were evaluated on Extended Cohn-Kanade Dataset (CK+). Reddy et al. [8] proposed a Face Emotion Recognition system that employs Stacked Auto-Multiplexer (SCAM) to recognize the emotions and Convolutional Neural Networks besides Batch Normalization. The deep CNN is pre-trained as a C Auto-Multiplexer (CAM). The system was implemented on seven fundamental emotions specified as sad, surprised, happy, neutral, fear, disgust, and angry. The CNN included convolutional layers, BN, ReLU, and MAM pooling layers. Tzirakis et al. in [9] proposed a multimodal system that operates on raw signals to perform spontaneous emotion prediction from

ISSN: 2005-4297 IJCA

the speech and visual data. Scale Invariant Feature Transform (SIFT) and Histogram of Oriented Gradients (HOG) is used for the extraction of features. Long-short term memory networks are applied followed by CNN. Jain et al. [11] presented a system for the recognition of emotions that uses a Hybrid Convolutional-Recurrent Neural Network. Initially, normalization of nose and mouth is done followed by mean subtraction. It cites relations within the facial images and temporal dependencies can be analyzed during classification. The extraction of features is done by CNN which consists of a class of networks that has 6 convolutional layers and the last two layers are fully connected. The RNN is responsible for classification. The blended model considerably decreases false detection. Jain et al.[12] suggested a Multi-angle Optimal Pattern-based Deep learning (MAOP-DL) which consists of a) Extended Boundary Background Subtraction(EBBS) for performing background subtraction to overcome illumination, b) Multi-angle Texture Pattern + STM for extracting texture patterns, c) Densely extracted SURF + Local Occupancy patterns (LOP) for relevant key features of facial points, d) Priority Particle Cukoo Search Optimization (PPCSO) for selection of features, and e) Long Short-term CNN memory for classification. In [13], Tang et al. presented two models for facial emotion recognition: i) Bimodal Deep Denoising Auto-encoder model and ii) Bimodal-LSTM model. These models can use both frequency domain and temporal information of the features. The models used are tested on two datasets SEED and DEAP which involve EEG features, eye movements, and peripheral physiological signals. It is found that the Bimodal-LSTM model outperforms the other. Xie et al. [14] presented a Feature redundancy-reduced Convolutional network for the facial emotion recognition system. Emotions of sadness, surprise, happiness, disgust, fear, and anger were considered. The system achieved an accuracy of 92.06, it performs well in classifying expressions of surprise, happiness, and disgust but seems confusing to classify that of anger and sadness. Liliana et al. [15] proposed a system for emotion recognition from facial expression using the combination of Support Vector Machine and conditional Random Fields. The methodology was tested on six emotions mainly happy, sad, fear, disgust, angry, and surprise. The system evaluated the performance by creating own dataset by integrating 12 emotions as humans exhibit many different emotions. Patwardhan et al. [16] proposed a method for face emotion recognition by combining SVM classifier with Rule based learning in order to expedite the detection efficiency. The proposed method was conducted using Static images that contributed in giving body measurements and different features, Temporal data that was used for providing content about facial expressions of actor, hand gestures and body movement and 3D data frames taken from sensors. Garcia et al. [17] proposed two structures, the first one reduces the number of layers then observes the fact and the second approach divides the input images horizontally into two streams based on positions of eyes and mouth. The method is tested on seven primary emotions. Datta et al. [19] presented a system for detecting face emotion using geometrical and textual features of face. SVM classifier was used for efficient detection of facial emotion. The method was evaluated on CK+ dataset having 6 basic emotions in images. The proposed method observed 91.85% overall system accuracy in detection of facial emotions. Fan et al. [20] presented a video-based face emotion recognition system by integrated the two methods: 1) CNN-RNN which was used for feature extraction and later the emotions were encoded, 2) C3D Hybrid Networks which modelled appearance and the motion of video. The proposed method was able to achieve an accuracy of 59.02%. Aggarwal et al. [21] presented a method for face emotion identification and detection based on johns and viola algorithm in combination with principle component analysis. A brief introduction was given about the techniques used in face emotion detections 1) Hidden Markov Model 2) Geometrical Feature matching 3) Neural network 3) Template Matching. The proposed method achieved an accuracy of 99.84% with an error rate of 0.023. Levi et al. [23] presented a system that converts images to LBP (Linear Binary Codes) and then performs classification through CNN. The system is conducted on two datasets: i) SFEW, and

ii) EmotiW that consisted of six primary emotions. Ng et al. [24] proposed an Emotion Recognition System using Deep Learning technique (CNN). However, the proposed method was evaluated on small dataset by using Transfer learning and was able to achieve 55.6% accuracy in recognition of seven emotions exhibit by primarily humans. The proposed work concluded that the cascading fine tuning method is more efficient method than single tuning approach.

3.1. Face Emotion Recognition Framework

Face emotion recognition involves several stages as discussed below:

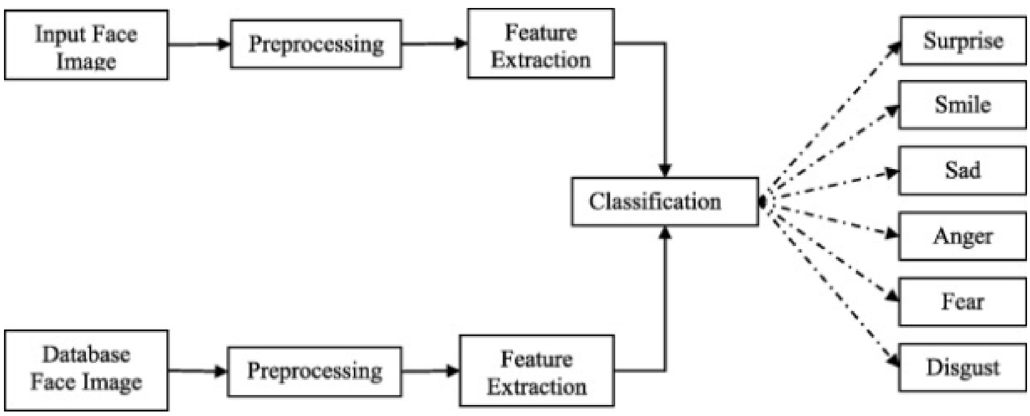

Figure 1. Face Emotion Recognition system (FER) [10]

i. Face acquisition: image acquisition is the process of acquiring facial images from the available input sources (Camera, video, sensors) and other information databases. Since the human image may contain other body parts which may not be required for emotion detection thus, the facial images are first acquired using image acquisition techniques.

ii. Preprocessing:
Since the images acquired from different data sources are somewhat unclear and distorted which may include hindrance in detecting the features properly. Therefore, this is the stage where facial images are processed using techniques such as filters, gray scaling and normalization.
   i) Normalization: it is the process of bringing an evenly effect to the image which are obscure and are low in illumination. It basically the changes the pixel values to bring a change in the contrast of an image and stretches the histogram range.

iii. Feature Extraction:
It is the significant process in the field of computer vision. It is the process which is widely followed in facial emotion recognition. It involves extracting of facial features which aids the machine learning algorithm in classifying different emotions. Feature extraction techniques are mainly Edge based, Texture based, Region based, Shape based.

iv. Classification of emotions:
This is the final stage in face emotion recognition. Classification in face emotion recognition is the process of identifying different emotions in separate classes such as anger, sad, happy, disgust, surprise and neutral. It involves several techniques such as Support Vector Machine, Convolutional Neural Network, and Recurrent Neural Network etc. In the research done so far, CNN is popularly used for efficient classification of emotion from facial images.

3.2. Role of Machine Learning and Deep Learning in Facial Emotion Recognition

Machine learning is the process of making the machines more intelligent and thereby reducing the manual and arduous work. With the advancement in technology, many machine learning algorithms are extensively used for Face Emotion Recognition. These algorithms enhance the quality and detection rate of emotions. Classifiers such as SVM, Decisions rule requires preprocessing of facial images for efficient recognition.

Deep learning is subdivision of machine learning algorithms which consists of multiple hidden layers. Each layer helps in processing the image and producing the better-quality image and thus avoids using any other preprocessing method. Algorithms such as CNN, RNN are widely implemented in face emotion recognition.

3.3.Benefits of Face Emotion Recognition

There are many advantages of a Face emotion recognition system. Some of them are listed below:

3.3.1 The FER system aids in detecting mental health of an individual. A good mental health is very important for living a health life, thus FER system analyzes the emotions to understand the emotional health and then helping the psychologists and doctors to improve the mental health of the patients.

3.3.2 It can help in easily giving a feedback to an organization or an interactive TV program [7].

3.3.3 The most popular benefit is in online gaming. The games are tested by recording the facial expressions of the players and understanding their response on various features of the game [1].

3.3.4 It can help in understanding the behaviour of an individual [7].

3.4.Issues

3.4.1 Technical challenges

FER face a lot of difficulty when recognizing the expression from dynamic objects in a video stream. Since the video move at a particular rate and sometimes the action remains incomplete in detecting the expression which may lead to other problems such as, misclassification. Also, video contains diverse range of people which may include people of different race, gender etc who may exhibit different kind of emotions. Machine learning classification algorithms, if not well trained with wide ranging groups of people may cause problem in classifying the expression of people.

3.4.2 Illumination and Occlusion issue

There is problem when working with untrained data as the change in posture might cause occlusion problem. This can be prevented by using frontal side of images to increase the efficacy. In addition to that, illumination and contrast plays an important role in detecting emotions. This is where normalization is used for better recognition.

3.4.3 Misinterpretation of features

Gestures play a significant role in emotion recognition. But there exist wide range of factors that affect the classification accuracy. Factors such as distance between eyes, thin lips, pale lips, skin color hinders the detection rate. Therefore, algorithm must be robust enough to handle all the diverse data about the humans.

3.5. Application Areas of Face Emotion system

3.5.1 Online gaming

Various Multiplayer online games (MOGs) are very popular due to several features that they provide. In MOGs, players can communicate, interact, and collaborate with other players. But, the major drawback lies in the facial communication. Although facial animations have been provided in many MOGs, players have to use text commands which are inefficient in controlling the expressions of their respective avatars naturally. Therefore, there is a need for the players to convey their emotions in real-time effortlessly and naturally which can be done through a facial emotion recognition system. The system can also be used to test any new game launched by analyzing the user's expressions.

3.5.2 Customer's feedback assessment

A potential area that uses a Facial expression recognition system is customer satisfaction measurement. It helps the organizations in collecting the feedback of the users just by understanding their facial expressions instead of urging them to make an input to the feedback form. The expression of customers being served at the counter is taken which is then analyzed to examine the response of the customer while the services are being provided. The customer feedback is a vital quality measurement component of many organizations. This helps in improving the quality of the products and services provided.

3.5.3 Healthcare

The most significant aspect of being healthy is emotional and mental health. To improve the quality of life, it is necessary to have good emotional health while bad mental health can lead to social or mental health problems. To fight with bad mental health it first needs to be recognized, therefore, many healthcare systems are designed to recognize the emotions of an individual by analyzing the facial expressions so that appropriate services and medications can be provided for the treatment of mental and emotional health.

4. Comparative Analysis of Face Emotion Recognition system using different learning techniques:

Authors	Dataset	No. of Samples	Types of Emotions	Feature Extraction	Feature Classification	Outcomes
Abdulsalam et al.[1] (2019)	ADFES-BIV	-	6 Basic + Neutral +pride+ contempt+ embarrassment	-	Deep CNN	Accuracy=95.12%
Chen et al. [2] (2019)	JAFFE	213	Angry, Sad, Happy, Fear, Neutral, Disgust, Surprise	LBP operator	Kernel sparse representation classifier	Accuracy=85.71%
Jain et al. [4](2019)	CK+, JAFFE	-	Angry, Sad, Happy, Fear, Neutral, Disgust, Surprise	-	CNN, Deep Residual Block	Accuracy=94.23%
Zadeh et al. [5] (2019)	JAFFE	Female=213	Angry, Sad, Happy, Fear, Neutral, Disgust, Surprise	Gabor Filter	CNN	Accuracy=97.16%

ISSN: 2005-4297 IJCA

Reddy et al. [8] (2018)	KDEF	Male=35 Female =35	Angry, Sad, Happy, Fear, Neutral, Disgust, Surprise	CNN with batch Normalizati on	Stacked Auto Multiplexer (SCAM)	Accuracy=93.52%
Tzirakis et al. [9] (2017)	RECOLA Database of AVEC(201 6)	43 participan ts	-	Scale Invariant Feature transform, HOG	CNN	-
Jain et al. [11] (2018)	TFD & JAFFE	Greater than 3000 images	Angry, Sad, Happy, Fear, Neutral, Disgust, Surprise	CNN	RNN	Accuracy=95.46% Mean Class accuracy=93.46%
Jain et al. [12] (2017)	CK+, MMI	130	Angry, Sad, Happy, Fear, Disgust, Surprise	MAP-STM, DESURF-LOP	STM-CNN	Accuracy=98.72%
Xie et al. [14] (2017)	CK+, JAFFE		Angry, Sad, Happy, Fear, Disgust, Surprise		FRR-CNN	Accuracy(CK+)=92 .06%, Accuracy(Mixed)= 83.96%
Liliana et al. [15] (2017)	CK+, Customized Dataset	-	Angry, Sad, Happy, Fear, Disgust, Surprise	Gabor Filter	SVM & CRF	Avg. Accuracy=86.90%
Garcia et al. [17] (2016)	KDEF	70 individual s	Angry, Sad, Happy, Fear, Neutral, Disgust, Surprise	-	CNN	Accuracy=96.93%
Datta et al. [19] (2016)	CK+	-	Angry, Sad, Happy, Fear, Disgust, Surprise	-	Hierarchal Multiclass SVM	Accuracy=91.85%

5. Conclusion

Emotion recognition allows users to examine the sentiments of a human face. Many organizations have been deploying several computer vision techniques for efficient face emotion recognition by integrating advanced algorithms and formulas with the existing techniques and practices. Due to its complexity, detection of facial expressions has become a challenging task in fields like biometric face identification, law enforcement, health care, psychology, gaming and identification of missing individual. For this, Facial Emotion Recognition (FER) systems have been developed to extract facial features like eye movements and lip movements further, classifying these features into different categories of emotions. On the basis of analysis done so far, it can be inferred that Deep Learning techniques such as CNN, RNN are the most efficient and widely implemented approaches for Facial Emotion Recognition (FER).

References

[1] Abdulsalam, W. H., Alhamdani, R. S., & Abdullah, M. N. (2019). Facial Emotion Recognition from Videos Using Deep Convolutional Neural Networks. *International Journal of Machine Learning and Computing*, *9*(1), 14–19.

[2] Chen, S.-P., Wang, X.-H., Peng, Z.-Q., Chen, D.-C., & Hong, J. (2019). Facial Expression Recognition via Kernel Sparse Representation. *DEStech Transactions on Computer Science and Engineering*, (ammso).

[3] Hossain, M. S., & Muhammad, G. (2019). Emotion recognition using deep learning approach from audio–visual emotional big data. *Information Fusion*, *49*, 69–78.

[4] Jain, D. K., Shamsolmoali, P., & Sehdev, P. (2019). Extended deep neural network for facial emotion recognition. *Pattern Recognition Letters*, *120*, 69–74.

[5] Zadeh, M. M. T., Imani, M., & Majidi, B. (2019). Fast Facial emotion recognition Using Convolutional Neural Networks and Gabor Filters. *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI).*

[6] Li, R., Tian, J., & Chua, M. C. H. (2018). Facial expression classification using salient pattern driven integrated geometric and textual features. *Multimedia Tools and Applications*, *78*(20)

[7] Panchal, S., Hiremath, A., & Toravi, N. R. (2018). Automatic Face Emotion Recognition. *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*

[8] Reddy, T. M., & Singh, D. R. P. (2018). Real Time Human Emotions Recognition using Auto-Multiplexers based Deep Convolutional Neural Networks. *SSRG International Journal of Electronics and Communication Engineering (SSRG - IJECE) – Special Issue ICETST Nov 2018*.

[9] Tzirakis, P., Zhang, J., & Schuller, B. W. (2018). End-to-End Speech Emotion Recognition Using Deep Neural Networks. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.

[10] Revina, I., & Emmanuel, W. S. (2018). A Survey on Human Face Expression Recognition Techniques. *Journal of King Saud University - Computer and Information Sciences*.

[11] Jain, N., Kumar, S., Kumar, A., Shamsolmoali, P., & Zareapoor, M. (2018). Hybrid deep neural networks for face emotion recognition. *Pattern Recognition Letters*, *115*, 101–106

[12] Jain, D. K., Zhang, Z., & Huang, K. (2017). Multi angle optimal pattern-based deep learning for automatic facial expression recognition. *Pattern Recognition Letters*.

[13] Tang, H., Liu, W., Zheng, W.-L., & Lu, B.-L. (2017). Multimodal Emotion Recognition Using Deep Neural Networks. *Neural Information Processing Lecture Notes in Computer Science*, 811–819.

[14] Xie, S., & Hu, H. (2017). Facial expression recognition with FRR-CNN. *Electronics Letters*, *53*(4), 235–237.

[15] Liliana, D. Y., Basaruddin, C., & Widyanto, M. R. (2017). Mix Emotion Recognition from Facial Expression using SVM-CRF Sequence Classifier. *Proceedings of the International Conference on Algorithms, Computing and Systems - ICACS 17*.

[16] Patwardhan, A. S., & Knapp, G. M. (2016). Augmenting Supervised Emotion Recognition with Rule-Based Decision Model.

[17] Ruiz-Garcia, A., Elshaw, M., Altahhan, A., & Palade, V. (2016). Deep Learning for Emotion Recognition in Faces. *Artificial Neural Networks and Machine Learning – ICANN 2016 Lecture Notes in Computer Science*, 38–46.

[18] Zhang, Y.-D., Yang, Z.-J., Lu, H.-M., Zhou, X.-X., Phillips, P., Liu, Q.-M., & Wang, S.-H. (2016). Facial Emotion Recognition Based on Biorthogonal Wavelet Entropy, Fuzzy Support Vector Machine, and Stratified Cross Validation. *IEEE Access*, *4*, 8375–8385.

[19] Datta, S., Sen, D., & Balasubramanian, R. (2016). Integrating Geometric and Textural Features for Facial Emotion Classification Using SVM Frameworks. *Advances in Intelligent Systems and Computing Proceedings of International Conference on Computer Vision and Image Processing*, 619–628.

[20] Fan, Y., Lu, X., Li, D., & Liu, Y. (2016). Video-based emotion recognition using CNN-RNN and C3D hybrid networks. *Proceedings of the 18th ACM International Conference on Multimodal Interaction - ICMI 2016*.

ISSN: 2005-4297 IJCA

[21] Agrawal, S., & Khatri, P. (2015). Facial Expression Detection Techniques: Based on Viola and Jones Algorithm and Principal Component Analysis. *2015 Fifth International Conference on Advanced Computing & Communication Technologies*.

[22] Chelali, F. Z., & Djeradi, A. (2015). Face Recognition Using MLP and RBF Neural Network with Gabor and Discrete Wavelet Transform Characterization: A Comparative Study. *Mathematical Problems in Engineering*, *2015*, 1–16.

[23] Levi, G., & Hassner, T. (2015). Emotion Recognition in the Wild via Convolutional Neural Networks and Mapped Binary Patterns. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction - ICMI 15*.

[24] Ng, H.-W., Nguyen, V. D., Vonikakis, V., & Winkler, S. (2015). Deep Learning for Emotion Recognition on Small Datasets using Transfer Learning. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction - ICMI 15*.

[25] Ouellet, S. (2014). Real-time emotion recognition for gaming using deep convolutional network features.