INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

IJARIIT

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 7, Issue 5 - V7I5-1370)
Available online at: https://www.ijariit.com

Facial emotion recognition using Convolutional Neural Networks for Autism and Dementia-related diseases

Hanshul Bahl
hansshul@yahoo.com
Birla Vidya Niketan, New Delhi, Delhi

ABSTRACT

Transfer research findings are important for mouth-based emotional recognition since there are limited data sets and most contain virtual actor emotional gestures rather than real-world categorization. Through way of transfer learning, we can use less training data than training an entire network from scratch, thereby improving the network efficiency with emotional data and improving the overall output accuracy of the neuronal network within the desired area. The suggested solution seeks to dynamically enhance the understanding of feelings, taking into consideration not just new situations, but changed contexts, because even though the face in whole is apparent in an unfavourable perspective, the picture of the mouth can be accessible. Typical applications include automatic monitoring of vital bedridden patients in a hospital management system and wearable applications assisting people with conditions with facial expressions challenging to see or to understand. This achievement takes advantage of previous work on mouth-based deep-learning emotional recognition and has also been validated and compared to a variety of other networks utilizing an extensive dataset for facial emotional recognition, well-known in publication. The exact recognition of the mouth was also contrasted to that of a total emotional knowledge; we find that the lack of precision is mostly offset by continuous success in the area of visual awareness of emotions. Therefore, we may conclude that in the dynamic phase of emotional recognition our system shows the value of mouth detection.

Keywords— *Visual, Emotion, Facial, Convolutional, Neural, Network, Autism, Dementia, Disease, Recognition, Detection, Perception*

1. INTRODUCTION

Visual Emotion Recognition (ER) has been researched extensively as one of the first affective computational techniques that incorporates eye, mouth and different factial elements dependent on visual features of the face. Various methods to visual recognition have received multiple classifications of various methods for visual recognition [1–3]. In recent experiments with the mouth alone to identify face feelings, positive results have been made, though still being not adequately recognised by the most experienced. These works use neural Convolution Neural Networks (CNNs) to monitor fundamental human emotions or to provide digital resources, especially for health care systems, to provide advanced and omnipresence equipment including smartphones or computer cameras 4]. The goal is to provide textual, audio or visual input. When trained on a single person, a neural network can achieve an excellent outcome with a relatively limited dataset of photos, for example, to recognize individual conditions that require immediate medical attention or to evaluate changes over time, which imply an underlying degenerative element. After we show that the mouth itself may promise a successful emotional recognition, we concentrate in this work on the mouth as a special feature for recognition on the facial expression by advanced techniques of deep learning. The purpose was to increase the emotional recognition mouth-based method to generalize the previous experiment on a multi-user dataset. For this reason, advanced deep learning methods, i.e. information transfers with continuous learning components, have been introduced. These methods may be found in the clinical system in particular. Connecting this architecture with the necessary resources, for example, will allow people to easily communicate feelings automated, for instance, by delivering more cognitive stimulation to autism-affected users or other social-relationship disorders, where a person has trouble identifying emotions conveyed by others. A further illustration is a procedure that identifies extreme problems and requests for assistance by a human assistant, for example in hospitalized people that have significant discomfort or psychiatric help. These applications may provide healthcare workers with input to take advantage of continued learning. In the course of previous training on massive datasets of photographs we testet the strongest neural network[1] for facial recognition and acknowledgement of the mouth e.g. lip reading[6]. The transfer of information enabled our neural networks to be pre-trained with low quality features such as edges, corners and colors. We performed ad-hoc training on the identification of the face in emotional groups in the last layers of the CNN. The testing finished with a review of a variety of CNNs on a commonly used dataset [8] of human faces markened by Ekman's emotions.

The biggest issue with the detection of facial expressions is the absence of adequate evidence for training videos. Many of the current datasets take just Ekman[9] or its sub-sets[5] into account, discarding more nuanced frameworks, such as Plutchik[10], or emotionally-supporting models[11]. Other than the naturally normal types of facial expression[12] Image data sets often include non genuine expressions (e.g., simulated by skilled actors). Since in-depth learning will remove characteristics that people do not even know, non-real aspects of emotion render it impossible to efficiently train a neural network for emotional recognition. Therefore, our study has based our work on a subset of the Ekman model which, by investing a reasonable dedication to selecting the photos from the available datasets, both from internet sources and self-generated content, has been increased by the neutral speech as a control state for emotional recognition results (4) Emotional photographs may be collected for emotions which are readily stimulated, e.g., excitement and pain, but other emotions including ethical concerns such as rage and anxiety are very difficult. We used transfer learning to improve results for every emotion with a relatively small set of images.

Our research uses the high specificity of CNNs by handling mouth pictures in order to identify emotional conditions via the new progresses. Affective computation, first described in 2003 by Rosalind Picard[16], will soon be one of the trendiest multidisciplinary research studies in artificial intelligence. Affective computation with a variety of fields, for example in psychiatry, neuroscience, neurology, liberal arts, information science and robotics, has recently highlighted the value of extracting and identifying affective mental disorders, e.g. emails, moods, thoughts and personality characteristics. In terms of the interest in such modern studies, much of the research currently available is still centered on superficial sensation analyzes or pure theoretical modeling (e.g. in psychology and social science) and product-based applications [17], particularly for market-ing purposes. In addition to automatic sensors and computer assistants with emotional intelligence skills, we want to concentrate on self-help[18], medical management and connectivity for human comprehension and help with real-life challenges with every difficult job (e.g. injury [19], psychological issues in crises and specific environments). Every day, citizens still use, and generate an immense number of heterogeneous data in a manageable and customized subset of classified goods in our socially integrated environment. Data can be used for emotional identification (ER) explicitly, for example pictures and text exchanged on social networks, or include inferable emotional intelligence factors such as wearable physiological data, including slept patterns, continuous pulse and monitoring of breathing, art-specific expression of emotions [21], and consumer data from thrilling games [22]. If the study of sentiment involves just the identification of the optimistic, negative or neutral character of sentiments, moods and emotions, so the mechanism of emotional recognition remains little studied. Since the scientists do not agree on all models accessible, research on ERs in any field begins with Ekman's well established model, which we have chosen for our analysis. Recent investigations also shown that Ekman's key emotional conditions, including satisfaction, sorrow, rage, displeasure and neutrality[24], are understood in a manner that can easily be contrasted with a multidimensional approach[11] on the basis of texts[24] and physiological clues such as pulse rate, skin behavior, movements and facial expressions. Facial recognition is already widespread across all ER methods. From this viewpoint, we chose to concentrate on the mouth as a special facial feature, which can be seen almost often in every facial speech, even though some societies are more underestimating that the mouth is expressive than others. In South America, for example, many people rely more on the top of their face, as Ekman himself emphasized in his first experiments on the marking of face emotions[9]. The new problem is that the neutral emotion is, in general, the most confused of human and artificial agents when contrasting those original trials with our findings in AI, and is tending to be represented as indignation or sorrow. Their automated mouth-dependent recognizers in same error groups often confirm that the auto-matic agent will correctly identify pictures, based on human marks. The fact that the same emotion is often mistaken.

In this analysis, we checked the multi-user data set methodology to find answers to the following research questions:
- How correctly would face feelings just from the mouth be recognized?
- Is the methodology suggested capable of identifying emotions while focused on a wide spectrum of facial images?

Our mouth-based ER development takes advantage of transfer learning (i.e. infor-mation transmission), which uses neural networks of general intent pretrained by comprehensive and later fine-tuned data sets of various types (i.e., mouth-based emotion recognition). Given an area where this approach is useful and relevant data are accessible as input in order to monitor the weights of your neural network and continue training, our implementation can easily be modified for continuous education. Such an approach could boost emotion detection performance in real-time circumstances, where a prior testing will use the final weights in a corresponding time. In order to pause and roll back in cases of conspicuous mistakes and over-fitting in a particular setting, for example, a certain camera resolution or a certain light situation in one location or at one particular moment, a secondary result of such equipment needs further remarkable computing power and a semi-supervised device. Continuous learning will consistently change and increase the accuracy of the network, contributing to a more stable framework over a longer period.

2. LITERATURE REVIEW

Face detection is one of the biometric techniques used to identify human identifi-cation utilizing automated methods focused on physiological attributes. Human facial features are an important aspect of identity when it comes to face recog-nition. This is because the face is the primary interface for presence, contact, speech, and collective recognition of the human sensory organ. Goldstein and Chance[15] performed the first comprehensive research on future growth of facial recognition capabilities for schoolchildren of various ages. Children from 5, 8 and 13 years old would conduct an acknowledgment assignment, by displaying first a selection of unknown faces and then choosing the previously found faces from a broader collection of unknown faces at a later point. The efficiency of this role in the age community has been consistently improved[16]. Nonetheless, Kanade [17] performed automatic facial detection, which used easy methods of processing the picture with a 16-facial vector. The parameters are size, region and angle ratios. By using Euclidean distance (ED) calculation, the maximum production of 20 subjects with two images per individual was achieved by 75 percent[18]. With its techniques, algorithms and processes, facial detection has since greatly enhanced. The face reconnaissance technique is likely classified into three approaches:(1) a practical approach where analyzer analyzes local characteristics such as the nose, eyes and the mouth and the geometric interplay (i.e. elastic bunch graph maping (EBGM)), the Secret Markov model (HMM). Either an identifying or confirming face may be recognised. Identification is a 1:M which compares an image query to a reference image in the database to find an identity of the query, and a 1:1 match which compares the image query to the image prototype for identity argument [19]. However, Grother[20] experimented with another method where the test face may not actually be within a database but contrasts the query face to all the faces in the database with the score measurement for each face, which is numberically ranked to first generate the maximum score. With improvements in facial recognition technologies, certain recognition systems are able to reach over 90% precision, but a real-time solution is a major challenge owing to the broad variety of variations. While generic face recognition typically follows a pattern of symmetric structure indicating the replication in a separate location or direction of the face structure, facial asymmetry may occur when the human face is deformed or damaged by illness or genetic condition. The right and left symmetry in the area of biology defines inherited consistency and wellbeing in health[21]. In order to measure the dynamics of symmetric variance and asymmetry [22], morphometric facial analysis with the analyses of landmark configurations is also used. The most visible facial characteristics are added to the advent of face recognition. It was a smart method to imitate the potential of the human face to see through certain intuitive functions (eyes, mouth and cheeks)[23] and geometric proportions (distance points, height-to-width ratio)[24]. These phenomena are also important today as the discarding of certain facial features or portions of the face will boost recognition [25]. Therefore, it is important to consider which facial characteristics lead to optimal efficiency and which add noise. However, an abstract mathematical approaches including Eigen faces [26,27] have appeared, which slowly leaves behind the anthropometric methodology. Another approach has also been implemented. However, there are also some characteristics that are important to facial recognition, such as skin colour [28]. It is nevertheless important, in computational or mathematical terms, to present abstractions and to tackle challenges.

3. CONVOLUTIONAL NEURAL NETWORKS

Convolutionary neural networks form a class of deep neural networks which prove especially efficient for various data tasks, which are organized on the topology of the grid, such as time series and visual inputs. Typical implementations include picture identification, segmentation, identification and retrieval [28–30], where CNNs have been up-to-date. Three important variables, such as sparse associations, parameter-sharing and equivalent depiction, are due to this major breakthrough. [31«] Large sample sizes of testing may be processed productive and decrease training times dramatically through utilizing more deep networks of millions of parameters to acquire more complicated and distinctive picture features. Another benefit in the form of conventional, completely linked networks requiring defined inputs is the probability that variable input size is possible. Some typical completely connected layers are replaced by convolution lays in convolutionary neural networks which search data in subsets by organized, organized grid-like data and multiply each sub-set by a kernel matrix (i.e. filter) to generate a feature map in the output. The mechanism similarly responds to the sensory feedback of individual neurons: each neuron has a tiny component, named its receptive field, and avoids more detail. The receptive field. For a picture of size 1000 * 1000, training of 1000,000 weight for one layer neuron, training of the entirely linked neural network of the same feather recognition and thus classification capability will take far more effort. CNNs will instead have certain trainable parameters, but even less contingent on the scale of the implemented kernel. CNNs are typically interlaced sequences of three distinct layer types: convolution layers mentioned above, traditional completely linked layers and pooling layers, used to combine various functional characteristics in one layer, i.e. to sample the maps according to a number of strategies. Although there are various numbers of layers, filter sizes and amounts, the CNNs used in the literature to execute comparison task tasks such as the ImageNet classification are often made up of building blocks.

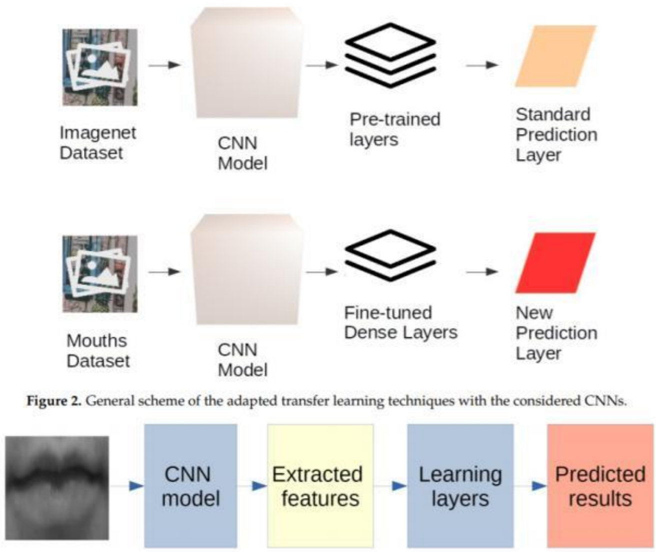

Figure 1. General scheme of the adapted transfer learning techniques with the considered CNNs.

Testing of four convolutionary neural networks was conducted: VGG16 [32], InceptionResNetV2 [33], InzeptionV3 [34] and Xception [35]. Among CNN's, the networks outperform AlexNet[28], one of the largest image datasets for the transfer-learning pre-workout level, on the commonly used ImageNet [7]. Both neural networks whose general behaviour is defined in Figure 1 were evaluated by means of transfer training. We used the pre-trained neural networks that implemented transfer learning, and the weights were determined on the ImageNet dataset [7]. For the training process the first layers of the networks were frozen; i.e. the weights of the convolutionary layers were fixed and the mouth emotion data set was not altered during the finely-set step. This was achieved because the CNNs were able to distinguish low-level characteristics such as peaks, arcs, corners, coins and colour distribution. The final stages have been substituted by two thick layers, each with 64 neurons and one softmax, to optimize the network to mouth emotion recognition. This layer classifies the possibility of the most desirable class, thereby restoring the hierarchy of emotions. Only the last layers, the completely linked SoftMax level and the final level of the CNN networks therefore adjust and can be retrained throughout the fine tuning process. The model was configured only to save weights if it increases the emotional classification precision of the earlier period, contributing to the final best configuration of the neural network training. Two optimizers have been tested: Adam and SGD. InceptionV3 and VGG16, Adam did higher, having a learning score comparable to 0.001 and 0.01 in Xception. SGD did higher with a Learning Rate similar to 0.001, momentum = 0.9 and nesterov equal to true with InceptionResNetV2. The other conditions were as follows. Batch size of 25 with a cumulative epoch amount of 100; we used the early stop technique: the workout is stopped if the outcomes have not changed.

4. FACE RECOGNITION TIMELINE

Computer vision typically includes facial identification using a three-step pro-cess[30], which follows a logical pipeline: face and labeling detection, extraction and classification (or recognition). These are discussed independently under the sphere of visual deficiency and autism facial perception in the following pages.

For face recognition and various other applications such as the capacity for the device to identify and localize facial features in a picture, face and landmark identification is a key first phase in the face analysis. The application is various, including face recording, compression or pose estimation. There are several subproblems in the definition of face recognition. Some simultaneously identify and find faces; some first run a detection routine and then locate faces if positive; while others can require tracking algorithms. Different facial detection methods are seen in [29], but only a few are actually applied to face detection systems, such as feature invariant and appearance dependent methods.

Role extraction is the next stage after facial recognition. This includes the obtaining of specific visual properties, which could be variations, angles or measurements for some face areas. This is a method of collecting knowledge from a facial picture that is capable of classification, and with an appropriate error rate. In terms of memory use and machine time, this method should be successful. It's not complicated for us as human beings to remember individuals whom we meet, even if they occlude, because of their common faces, but because of their devices.

The extraction method may consist of reduced dimensionality, extraction of functions and selection of features, but often overlaps. Basically, the usable data is based on an initial data mixture or transition. The selection function selects the appropriate and better component of the input function by discarding the unnecessary features resulting in the smallest classification error. Reduced dimensionality may be done or integrated into any of these measures before removing or choosing features. In the end, the number of characteristics should be carefully selected, especially since there are numerous measurements that which impact the algorithm of recognition such as tempo, dimensionality curse, number of attributes, sample quantity and classification difficulty. Redundant characteristics will adversely impact device accuracy. However, role extraction can be split into four categories: a local, holistic approach to functionality, model statistical designs, and approach to neural network [45,46]. In this segment, only 3 practical extractions directly relevant to DS extraction are addressed.

The local characteristic approach is the most primitive type of facial characteris-tics extraction and concentrates mostly on local texture characteristics such as eyes, lips, eyebrows and nose by locally examining tiny image patches and aggre-gating details from local patches in a total picture representation[46]. Examples of this include transformed Scale Invariant (SIFT) [47] and locally applied binary patterns (LBP) [31,33,48]. Local features may also be look-based, utilizing color and form-based picture features. The key information in the face picture is retained while the redundant information is discarded. Examples include the ICA, the principal PCA, which was used in[50—52], and the Gabor Wavelet Transform (GW T), which were used by [34,36,53]. Examples include: the ICA, the key component analysis (PCA). Local properties can also be geometrically oriented, integrating the distances and angles between these and other geometric features into a single function vector that can be used for subsequent analyses.

5. NEUTRAL NETWORK BASED

Neural networks are ideal not only for a classification model, but also for the extraction of the characteristics as set out in [37]. The analysis in [39] found that CNN is better than other approaches, such as local feature-based and predictive models for the extraction of facial features. Few neural networks were suggested by numerous authors in relation to DS face recognition function extraction algorithms. Bruna and Mallat were tolerant of analyzing mathematical profound neural networks (DCNNs) for feature extraction[61]. In [47], the study considered that the deformation stability of the corresponding extractor was based on a wavelet transformation and the non-linearity module in each Network Layer. More experiments have since been carried out promising better results in medical prediction predivision diagnosis. In Thai children with de-identified computer-aided facial examination 30 frontal pictures of children with DS and 140 non-DS were present with the probabilities for identification of Down's disease[62]. Face2Gene technology has been used to compare the face photos.

The methodology began with face and bottom identification, accompanied by the calculation of several distances, angles and proportions of 130 points. The input face was sliced to face and fed into the DCNN. This lead to a vector, showing its correspondence in the Face2Gene database with each syndrome. With clinical accessibility or anthropometric calculation the authors proposed that the rating accuracy might improve. A novel framework has been proposed in[37] to diagnose facial developmental disorders. For recognition, six categories of disturbances including DS using DCNN were considered. The vision and the profound learning algorithms were used as a facial recognition system called the DeepGestalt framework. The similarities of these two-dimensional images is tested by hundreds of genetic syndromes focused on (unconstrained). The algorithm has trained more than 17,000 patients from a shared database of genotypes.

Although the method of local appearance extraction retains the main details in terms of shape and colour, it does not include particular facial regions, such as in local texture-based environments, preserving specific facial zones such as the eyes, lips, eyebrows and nose. This will help us find internal facial characteristics that are clinically required for the proper diagnosis. While the only known high precision rates for DS detection are geometrical attributes, as seen in [13], combining the technique with other local extraction processes will contribute to a better outcome. The value of geometric local feature extraction is the opportunity to blend geometric features with appearance or texture. As shown in [33,35], this increases classification efficiency. The statistically driven method reveals that it can capture the traits that are most used in the prediction of genetic diseases of dysmorphic syndromes such as PCA, LDA and ICA. PCA employs expressive functions which approximates data using the medium square error criterion in a linear subspace.

The SVM definition focuses on decision-making planners, who establish limits of decision[64]. The concept is the SVM. It is a valuable controlled data classification learning technique involving the creation and evaluating of data consisting of certain data cases[65], with the goal value and many attributes of each training package. The features were tracted by crossvalidating the technique of leave-on-one-out with a precision of 94.3 percent using an SVM classifier in [33] on DS and stable patients. But it is the intrinsic weakness of any kind of research to catch people from diverse ethnicities and tribes across the globe. While it includes several players and nations, the research only represents a tiny portion of the world's population. The authors used PCA as the extraction function of [34] to categorize photos with 94.4 percent precision, using SVM.. Although the approach is not meant to assess the diagnosis of genetic diseases, it will promote an unprejudiced reduction in the field of diagnostic search[34]. In [35] for the extraction of function, a statistical shape model, local forms were defined with ICA. The method has been tested on 130 pictures and with SVM accuracy 96.7 percent. The geometric elements of [32] which are the primary clues for diagnosis of syndrome were extracted nineteen paired landmarks. SVM classification was used to separate regular cases with a remarkably correctly validity of 97.92 percent between normal and irregular cases. An additional methodology has been outsourced by the CENTRIST descriptor, with a precision of 98.39 percent as classified by SVM in [13]; verified using a 10-times cross-validation procedure. Although the classification scheme has predicted high precision, 16 points and 14 features might not be adequate for predicting genetic disease correctly. SVM-RBF proved to have the best accuracy of 95.6 percent among the four classifiers that were employed to compare effects of recognition in[54]. The amount of landmarks can therefore not be adequate to characterize hereditary disease facial variance. The classification method stated that dimorphic faces were recognised by 97.34% and 96% of the SVM[36]. Nevertheless, there could be very little amount of DS patient subjects with genetic diseases to characterize heterogeneity.

6. CONCLUSION

Geometer representation has been shown to achieve improved autism efficiency and visual disability identification between the state-of-the-art extraction meth-ods described in the feature extraction phase. Emotional/double recognition in healthcare administration and automatic monitoring of vital patients such as individuals in hospitals are clear implementations of our methodology. Where direct human assistance is inaccessible, for example during the night or in wards where support is not tolerated, the system may support patients with the aid of the ER machine to advanced detections of the original unpleasant or awkward situation, to lift a signal, to make sanitation workers reply immediately, thereby preventing distress for the patient. Emotional recognition can be used to help victims following traffic collisions, in emergency or postoperatively in order to consider their discomfort thresholds, for example, to prepare supportive systems. When a critical condition is identified by a machine, the event should be recorded to nurses or an intervention/check-out order can be made to a patient room. With the right datasets of paine photos, our method may easily be applied to the previously mentioned scenario. The early detection of depressed States or the support of individuals with mental problems, including blind users or people with autism-related spectrum disorders, could offer an alternative approach. In future the framework of our emotional processing can provide the consumer with input (e.g. email, emoticon, sound) or a medium for transmitting knowledge to apps.

While the features are taken from photographs in different regions for various testing purposes, it is important to notice that for its particular identification reason the extraction of face-lifting characteristics for other images should not be generalized. Besides the study suggested in [34,37,39], the majority of the other papers have used limited data of less than 200 in the field of deep learning, which is known as an insignified number. Furthermore, several field works accept a comparatively limited number (30, 15 or less) of DS samples, while efficiency may be validated and improved through cross validation. To standardize the phototype in an effort, the path to the new area of precision medicine has been raised. This has since opened the door for patients with genetically disabilities to easily obtain an exact genetic diagnosis[39]. The patients and families require early and thorough analysis of a genetic condition and special powerful attention. In any event, it can be stressed that the correct identification of dysmorphic characteristics depends on professional practitioners and wellness workers respectively [13]. The literature analyzed found that SVM or SVM classificators showed greater accuracy of identification than neural network-based classifiers. That is more related with classification features, but most neural network-based recognition systems have not given the information extracted, so it is hard to assess. Given this, there are several benefits to the neural network methodology: a coherent approach for extraction and classification of features with scalable approaches for the quest for mild not-linear solutions[69]. The efficiency of a topic classifier not only depends on the sophistication of the classifier, but also on the amount of features and sample images used. However, the classification efficiency can decrease if the training sample counts small compared with the number of features. This can be minimized by utilizing the amount of features as a minimum of 10 times the training samples [69]. This is the "dimensional curse," however when they are redundant, a large number of features could generate false positives.

REFERENCES

[1] Gervasi, O.; Franzoni, V.; Riganelli, M.; Tasso, S. Automating facial emotion recognition. Web Intell. 2019, 17, 17–27.
[2] Sagonas, C.; Tzimiropoulos, G.; Zafeiriou, S.; Pantic, M. A Semi-automatic Methodology for Facial Landmark Annotation. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Portland, OR, USA, 23–28 June 2013; pp. 896–903.
[3] Kazemi, V.; Sullivan, J. One millisecond face alignment with an ensemble of regression trees. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1867–1874.
[4] Riganelli, M.; Franzoni, V.; Gervasi, O.; Tasso, S. EmEx, a Tool for Auto-mated Emotive Face Recognition Using Convolutional Neural Networks. In Lecture Notes in Computer Science, Proceedings of the International Conference on Computational Science and Its Applications, Trieste, Italy, 3–6 July 2017; Springer International Publishing: Cham, Switzerland, 2017; Volume 10406, pp. 692–704.
[5] Biondi, G.; Franzoni, V.; Gervasi, O.; Perri, D. An Approach for Improving Automatic Mouth Emotion Recognition. In Lecture Notes in Computer Science, Proceedings of the Computational Science and Its Applications— ICCSA 2019, Saint Petersburg, Russia, 1–4 July 2019; Misra, S., Gervasi, O., Murgante, B., Stankova, E., Korkhov, V., Torre, C., Rocha, A.M.A.C., Taniar, D., Apduhan, B.O., Tarantino, E., Eds.; Springer International Publishing: Cham, Switzerland, 2019; Volume 11619, pp. 649–664.
[6] Gervasi, O.; Magni, R.; Ferri, M. A Method for Predicting Words by Interpreting Labial Movements. In Lecture Notes in Computer Science, Proceedings of the Computational Science and Its Applications—ICCSA 2016, Beijing, China, 4–7 July 2016; Gervasi, O., Murgante, B., Misra, S., Rocha, A.M.A., Torre, C.M., Taniar, D., Apduhan, B.O., Stankova, E., Wang, S., Eds.; Springer International Publishing: Cham, Switzerland, 2016; Volume 9787, pp. 450–464.
[7] Deng, J.; Dong, W.; Socher, R.; Li, L.; Li, K.; Li, F.-F. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
[8] Sagonas, C.; Antonakos, E.; Tzimiropoulos, G.; Zafeiriou, S.; Pantic, M. 300 Faces In-The-Wild Challenge: Database and results. Image Vis. Comput. 2016, 47, 3–18.
[9] Ekman, P. An Argument for Basic Emotions. Cogn. Emot. 1992.
[10] Plutchik, R. A psychoevolutionary theory of emotions. Soc. Sci. Inf. 1982, 21, 529–553.
[11] Franzoni, V.; Milani, A.; Vallverdú, J. Emotional Affordances in Human-Machine Interactive Planning and Negotiation. In Proceedings of the International Conference on Web Intelligence, Leipzig, Germany, 23–26 August 2017; pp. 924–930.
[12] Franzoni, V.; Milani, A.; Vallverdú, J. Errors, Biases, and Overconfidence in Artificial Emotional Modeling. In Proceedings of the International Conference on Web Intelligence, Thessaloniki, Greece, 14–17 October 2019.
[13] Franzoni, V.; Milani, A.; Nardi, D.; Vallverdú, J. Emotional machines: The next revolution. Web Intell. 2019, 17, 1–7.
[14] Gervasi, O.; Magni, R.; Macellari, S. A Brain Computer Interface for Enhancing the Communication of People with Severe Impairment. In Lecture Notes in Computer Science, Proceedings of the Computational Science and Its Applications—ICCSA 2014, Guimarães, Portugal, 30 June– 3 July 2014; Murgante, B., Misra, S., Rocha, A.M.A.C., Torre, C., Rocha, J.G., Falcão, M.I., Taniar, D., Apduhan, B.O., Gervasi, O., Eds.; Springer International Publishing: Cham, Switzerland, 2014; Volume 8584, pp. 709– 721.
[15] Bastianelli, E.; Nardi, D.; Aiello, L.C.; Giacomelli, F.; Manes, N. Speaky for robots: The development of vocal interfaces for robotic applications. Appl. Intell. 2016, 44, 43–66.
[16] Picard, R.W. Affective Computing: Challenges. Int. J. Hum. Comput. Stud. 2003, 59, 55–64.
[17] Cieliebak, M.; Dürr, O.; Uzdilli, F. Potential and limitations of commer-cial sentiment detection tools. In Proceedings of the CEUR Workshop Proceedings, Valencia, Spain, 17–18 June 2013; pp. 47–58.
[18] Franzoni, V.; Milani, A. Emotion Recognition for Self-aid in Addiction Treatment, Psychotherapy, and Nonviolent Communication. In Lecture Notes in Computer Science, Proceedings of the Computational Science and Its Applications—ICCSA 2019, Saint Petersburg, Russia, 1–4 July 2019; Misra, S., Gervasi, O., Murgante, B., Stankova, E., Korkhov, V., Torre, C., Rocha, A.M.A.C., Taniar, D., Apduhan, B.O., Tarantino, E., Eds.; Springer International Publishing: Cham, Switzerland, 2019; Volume 11620, pp. 391–404.
[19] Hayes, G.R.; Hirano, S.; Marcu, G.; Monibi, M.; Nguyen, D.H.; Yeganyan, M. Interactive visual supports for children with autism. Pers. Ubiquitous Comput. 2010, 14, 663–680. [CrossRef]
[20] Picard, R.W.; Vyzas, E.; Healey, J. Toward machine emotional intelligence: Analysis of affective physiological state. IEEE Trans. Pattern Anal. Mach. Intell. 2001, 23, 1175–1191.
[21] Bertola, F.; Patti, V. Emotional responses to artworks in online collections. In Proceedings of the CEUR Workshop Proceedings, Valencia, Spain, 17–18 June 2013; Volume 997.
[22] Canossa, A.; Badler, J.B.; El-Nasr, M.S.; Anderson, E. Eliciting Emotions in Design of Games—A Theory Driven Approach. In Proceedings of the 4th Workshop on Emotions and Personality in Personalized Systems (EMPIRE), Boston, MA, USA, 16 September 2016.
[23] Angelov, P.; Gu, X.; Iglesias, J.A.; Ledezma, A.; Sanchis, A.; Sipele, O.; Ramezani, R. Cybernetics of the Mind: Learning Individual's Perceptions Autonomously. IEEE Syst. Man Cybern. Mag. 2017, 3, 6–17.
[24] Biondi, G.; Franzoni, V.; Li, Y.; Milani, A. Web-Based Similarity for Emotion Recognition in Web Objects. In Proceedings of the 9th Interna-tional Conference on Utility and Cloud Computing, Shanghai, China, 6–9 December 2016; pp. 327–332.
[25] Chollet, F. Keras. 2015. Available online: https://github.com/fchollet/ke ras
[26] Antoniou, A.; Storkey, A.; Edwards, H. Data Augmentation Generative Adversarial Networks. arXiv 2017, arXiv:1711.04340.
[27] Sagonas, C.; Tzimiropoulos, G.; Zafeiriou, S.; Pantic, M. 300 Faces in-the-Wild Challenge: The First Facial Landmark Localization Challenge. In Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops, Sydney, Australia, 2–8 December 2013; pp. 397–403.

[28] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. Commun. ACM 2017, 60, 84–90.
[29] Perri, D.; Sylos Labini, P.; Gervasi, O.; Tasso, S.; Vella, F. Towards a Learning-Based Performance Modeling for Accelerating Deep Neural Networks. In Lecture Notes in Computer Science Book, Proceedings of the Computational Science and Its Applications—ICCSA 2019, Saint Petersburg, Russia, 1–4 July 2019; Misra, S., Gervasi, O., Murgante, B., Stankova, E., Korkhov, V., Torre, C., Rocha, A.M.A., Taniar, D., Apduhan, B.O., Tarantino, E., Eds.; Springer International Publishing: Cham, Switzerland, 2019; Volume 11619, pp. 665–676.
[30] Farabet, C.; Couprie, C.; Najman, L.; LeCun, Y. Learning Hierarchical Features for Scene Labeling. IEEE Trans. Pattern Anal. Mach. Intell. 2013, 35, 1915–1929.
[31] Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning; MIT Press: Cambridge, MA, USA, 2016.
[32] Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 2014, arXiv:1409.1556.
[33] Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. arXiv 2016, arXiv:1602.07261.
[34] Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. arXiv 2015, arXiv:1512.00567.
[35] Zhao, Q.; Okada, K.; Rosenbaum, K.; Kehoe, L.; Zand, D.J.; Sze, R.; Summar, M.; Linguraru, M.G. Digital facial dysmorphology for genetic screening: Hierarchical constrained local model using ICA. Med. Image Anal. 2014, 18, 699–710.
[36] Saraydemir, ¸S.; Ta¸spınar, N.; Ero €gul, O.; Kayserili, H.; Dinçkan, N. Down syndrome diagnosis based on gabor wavelet transform. J. Med. Syst. 2012, 36, 3205–3213.
[37] Shukla, P.; Gupta, T.; Saini, A.; Singh, P.; Balasubramanian, R. A deep learning frame-work for recognizing developmental disorders. In Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, 27–29 March 2017; pp. 705–714.
[38] Özdemir, M.E.; Telatar, Z.; Ero €gul, O.; Tunca, Y. Classifying dysmorphic syndromes by using artificial neural network based hierarchical decision tree. Australas. Phys. Eng. Sci. Med. 2018, 41, 451–461.
[39] Gurovich, Y.; Hanani, Y.; Bar, O.; Nadav, G.; Fleischer, N.; Gelbman, D.; Basel-Salmon, L.; Krawitz, P.M.; Kamphausen, S.B.; Zenker, M. Identifying facial phenotypes of genetic disorders using deep learning. Nat. Med. 2019, 25, 60.
[40] Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 8–14 December 2001; p. 1.
[41] Bradski, G.; Kaehler, A. The OpenCV library. Dr Dobb's J. Softw. Tools 2000, 3, 122–125.
[42] Kruszka, P.; Addissie, Y.A.; McGinn, D.E.; Porras, A.R.; Biggs, E.; Share, M.; Crowley, T.B.; Chung, B.H.; Mok, G.T.; Mak, C.C. 22q11. 2 deletion syndrome in diverse populations. Am. J. Med Genet. Part A 2017, 173, 879–888.
[43] Zhu, X.; Ramanan, D. Face detection, pose estimation, and landmark localization in the wild. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2879–2886.
[44] Li, H.; Lin, Z.; Shen, X.; Brandt, J.; Hua, G. A convolutional neural network cascade for face detection. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 5325–5334.
[45] Rai, M.C.E.; Werghi, N.; Al Muhairi, H.; Alsafar, H. Using facial images for the diagnosis of genetic syndromes: A survey. In Proceedings of the 2015 International Conference on Communications, Signal Processing, and Their Applications (ICCSPA'15), Sharjah, UAE, 17–19 February 2015; pp. 1–6.
[46] urovich, Y.; Hanani, Y.; Bar, O.; Fleischer, N.; Gelbman, D.; Basel-Salmon, L.; Krawitz, P.; Kamphausen, S.B.; Zenker, M.; Bird, L.M. DeepGestalt-Identifying rare genetic syndromes using deep learning. arXiv 2018, arXiv:1801.07637.
[47] Lowe, D.G. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 2004, 60, 91–110.
[48] Ahonen, T.; Hadid, A.; Pietikainen, M. Face description with local binary patterns: Application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2006, 28, 2037–2041.
[49] Yang, M.-H. Face recognition using kernel methods. In Proceedings of the 2015 Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 1457–1464.
[50] Kuru, K.; Niranjan, M.; Tunca, Y.; Osvank, E.; Azim, T. Biomedical visual data analysis to build an intelligent diagnostic decision support system in medical genetics. Artif. Intell. Med. 2014, 62, 105–118.
[51] Boehringer, S.; Guenther, M.; Sinigerova, S.; Wurtz, R.P.; Horsthemke, B.; Wieczorek, D. Automated syndrome detection in a set of clinical facial photographs. Am. J. Med. Genet. Part A 2011, 155, 2161–2169.
[52] Boehringer, S.; Vollmar, T.; Tasse, C.; Wurtz, R.P.; Gillessen-Kaesbach, G.; Horsthemke, B.; Wieczorek, D. Syndrome identification based on 2D analysis software. Eur. J. Hum. Genet. 2006, 14, 1082. [
[53] Loos, H.S.; Wieczorek, D.; Würtz, R.P.; von der Malsburg, C.; Horsthemke, Computer-Based recognition of dysmorphic faces. Eur. J. Hum. Genet. 2003, 11, 555.
[54] Zhao, Q.; Okada, K.; Rosenbaum, K.; Zand, D.J.; Sze, R.; Summar, M.; Linguraru, M.G. Hierarchical constrained local model using ICA and its application to Down syndrome detection. In Proceedings of the 2013 International Conference on Medical Image Computing and Computer-Assisted Intervention, Nagoya, Japan, 22–26 September 2013; pp. 222–229.
[55] Cerrolaza, J.J.; Porras, A.R.; Mansoor, A.; Zhao, Q.; Summar, M.; Lingu-raru, M.G. Identification of dysmorphic syndromes using landmark-specific local texture descriptors. In Proceedings of the 2016 IEEE 13th Interna-tional Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 13–16 April 2016; pp. 1080–1083.
[56] Wilfred, O.O.; Lee, G.-B.; Park, J.-J.; Cho, B.-J. Facial component features for facial expression identification using Active Appearance Models. In Proceedings of the AFRICON 2009, Nairobi, Kenya, 23–25 September 2009; pp. 1–5.
[57] Mahoor, M.H.; Abdel-Mottaleb, M.; Ansari, A.-N. Improved active shape model for facial feature extraction in color images. J. Multimed. 2006, 1, 21–28.

[58] Lee, Y.-H.; Kim, C.G.; Kim, Y.; Whangbo, T.K. Facial landmarks detection using improved active shape model on android platform. Multimed. Tools Appl. 2015, 74, 8821–8830.
[59] Erogul, O.; Sipahi, M.E.; Tunca, Y.; Vurucu, S. Recognition of down syndromes using image analysis. In Proceedings of the 2009 14th National Biomedical Engineering Meeting, Izmir, Turkey, 20–22 May 2009; pp. 1–4.
[60] Gabor, D. Theory of communication. Part 1: The analysis of information.Inst. Electr. Eng. Part III Radio Commun. Eng. 1946, 93, 429–441.
[61] Bruna, J.; Mallat, S. Invariant scattering convolution networks. IEEE Trans. Pattern Anal. Mach. Intell. 2013, 35, 1872–1886.
[62] Vorravanpreecha, N.; Lertboonnum, T.; Rodjanadit, R.; Sriplienchan, P.; Rojnueangnit, K. Studying Down syndrome recognition probabilities in Thai children with de-identified computer-aided facial analysis. Am. J. Med. Genet. Part A 2018, 176, 1935–1940.
[63] Intrator, N.; Reisfeld, D.; Yeshurun, Y. Face recognition using a hybrid supervised/unsupervised neural network. Pattern Recognit. Lett. 1996, 17, 67–76.
[64] Vapnik, V. The Nature of Statistical Learning Theory; Springer Science & Business Media: Berlin, Germany, 2013.
[65] Rejani, Y.; Selvi, S.T. Early detection of breast cancer using SVM classifier technique. arXiv 2009, arXiv:0912.2314.
[66] Kohonen, T. Self-Organization and Associative Memory; Springer Science & Business Media: Berlin, Germany, 2012; Volume 8.
[67] Islam, M.J.; Wu, Q.J.; Ahmadi, M.; Sid-Ahmed, M.A. Investigating the performance of naive-bayes classifiers and k-nearest neighbor classifiers. In Proceedings of the 2007 International Conference on Convergence Informa-tion Technology (ICCIT 2007), Gyeongju, Korea, 21–23 November 2007; pp. 1541–1546.
[68] Amadasun, M.; King, R. Improving the accuracy of the Euclidean distance classifier. Can. J. Electr. Comput. Eng. 1990, 15, 16–17.
[69] Jain, A.K.; Duin, R.P.W.; Mao, J. Statistical pattern recognition: A review. IEEE Trans. Pattern Anal. Mach. Intell. 2000, 22, 4–37.

