

# Conceptual Preview of the AI and Image Analytics Methodologies in Robotics

Sathyanarayanan P.S.V

Topcoder (Data Analytics and AI)

Chennai, India

## ABSTRACT

Due to the need of thinking cum adaptable machines in the variety of applications, the usage of AI-Artificial Intelligence concepts has seen tremendous hike as far as the current researches are concerned. AIs are still a growing technique, which makes it still difficult to evaluate than several technology-specific modifications, particularly in accordance with several decisions that the organization's management takes regarding the adapting and utilizing these innovative concepts, which in common, shaping the impact of that concept are found to be in earlier stages of development. When AI is to be utilized for the machine-human level interactions, the user recognition and the intended interaction becomes a mandatorily important task which needs more attention towards the image analytics. Thus, this paper is aimed to discuss, correlate basic image analytic methodologies along with its corresponding utilization in the robotics industry with reference to the image processing techniques. Majorly, the image processing operations taking up the facial/emotion recognition is discussed in a brief preview in this paper with reference to the utilized methodologies. After discussing many existing/current methodologies used in the facial recognition related applications deployed in the robotics industry, futuristic research direction will be suggested so that the forthcoming researches could benefit with the better application of the AI-specific methodologies.

**Keywords:** Artificial Intelligence, AI, Image Analytics, Image Processing, Robotics.

## 1. INTRODUCTION

Artificial Intelligence is generally nothing but making/initiating the machines to act and think according to the surrounding circumstantial happenings with its intelligence (i.e.) AI enables the machines to have intelligence just similar to the human being have. These machines may be robots, cars, personal electronic gadgets, IoT enabled devices, etc. Below, the figure depicting the AI fundamental will be included.

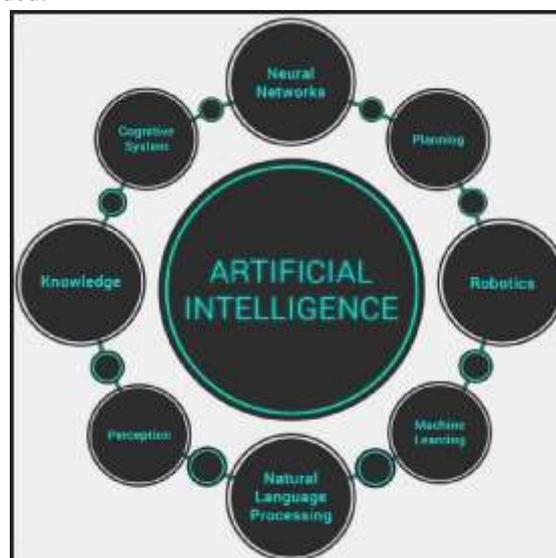


Fig. 1. Elements of AI (Image courtesy: scorio.com)

The above figure 1 indicates the elements constituting AI (simple schema). The robotics part that really belongs to AI includes all intelligent algorithms that perform navigation in global or local, direction/operation scheduling, and knowledge/intelligent dependent maintenance.

Specifically, robots are generally categorized into different kinds based on the applications in which they are utilized to perform certain intended tasks. Most important and widely utilized type of robots include the social robots. Social robots were the robots which was designed with the aim of detecting/recognizing the facial/emotions that the humans express while saying the

control commands. Some of the emotion related discussions will be included in the following sections with reference to the appropriate figure as and when required.



Fig. 2. Six crucial emotions as defined by [1]

The above figure is showing the six crucial emotion as per the words of [1]. In many researches taking up the social robots, indications of the emotions used to consists of 2 base concepts. The widely utilized and popular concept was put forth by [1]. As per this work, each emotions expressed could be categorized into 6 contributing categories such as disgust, happiness, fear, anger , surprise, and sadness along with the neutral expressions as depicted in the figure 2. Additionally, some other emotional expressions like contentment, shame, contempt, fun, embarrassment, pleasure, guilt, pride, excitement, relief, satisfaction were also seen added along with the earlier defined six facial emotions known as the secondary emotions. But, when recognition of emotion was done, only the 6 crucial emotions were taken in accordance with 2 significant reasons. First reason being somewhat related to the work, which was majorly utilized for making an effective and sophisticated multi-class recognition issue solving operation. As far as the next reason was concerned, when we take a good look at the recent robotics applications, they could be associated to the six crucial emotions only, at least when computation application was desired. Ekman along with the other researchers say that these six emotions must be taken as independent phases of the mind. Also, in the ending emotion expressions, attributes like couple of micro facial expressions, associated to every variety of emotions, could obviously intersect.

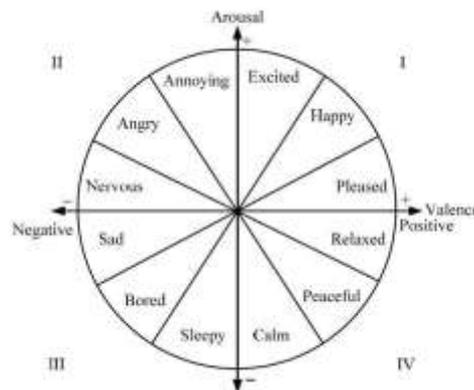


Fig. 3. Valence-arousal Concept by [2]

Other concept concerning the social robot’s emotions was put forth by [2] in the name of Valence-arousal. As per this work, the emotions could be obtained in the empirical parts of a phase of 2 major coordinates. One being the valence part contributing negative or else a positive intention to the emotional phase, and other being the arousal part measuring the feeling level as shown in the above figure 3. This work too defined variety of emotion phases in the human being’s face. But, again for the sake of avoiding unnecessary confusions in handling these several emotions, only the six crucial emotions defined by [1] were taken in accordance with the 2 earlier discussed reasons. This concept was somewhat more preferable than the earlier model by Ekman because of its improved suitability when the classification issue was taken up and its capability to refer implicitly to a hyper-phase migration issue.

When it comes to feature Engineering domain and corresponding learning, it was the base for the robots to operate based on recognizing some kinds of emotions or facial expressions. As a matter of fact, [3] framed the issue of local indication of imaging information as the estimation of optimum needed statistics which were invariant to nuisance variability triggered by the illumination along with the viewpoint. This work made us to realize that the under the influence of adverse scenarios, the issues were associated to the “attribute descriptors” wisely employed in the Vision of the Computer. These scenarios could be compensated if several perspectives of the alike circumstances were accessible. Majorly, work of [3] came up with a sampling

dependent and a point-valuation dependent approximation of some indications, correlated roughly with the picture-to-picture matching (Inter-picture matching). They also made use of benchmarks for different perspective based baseline matching, comprising of the collection of the artificial and real objects with the deployment of motion of the ground truth camera and 3-D geometry which were dense.

An ideal face recognizing methodology was developed by [4] for the lesser resolution RGB-D Kinect information. Several methodologies were devised for preprocessing of the image in accordance with the noises in the considered information. Initially, symmetry of the face was utilized on the basis of 3-D point cloud for acquiring the front viewed canonical picture independent of the earlier pose and afterwards, depth information would be transformed to XYZ ordinary maps. Later on, several channelled Discriminant Conversions were then deployed for projecting the RGB to Discriminant Colour Space and ordinary maps to Discriminant Ordinary Maps. Next, the several channelled ideal Sparse Coding operation was devised which used to code the every channels present in the investigated picture as a sparse incorporation of training samples comprising variety of pixel weightings. Then, weights were determined in a dynamic basis in an iterative operation for achieving the perfectness against pose related deviations, facial expressions, disguise, and illumination. Contradicting to the literature available methodologies, this devised several channelled operation was able to be perfect to the undesired deviations. Flaws in reconstruction of the investigated picture were fused and normalized accordingly to make the decision regarding the identification. The devised methodology was cross verified with 4 common databases. This method was able to attain 98.4% rate of detection when utilizing CurtinFaces, a Kinect database comprising 4784 RGB-D pictures with 52 themes. When, 1<sup>st</sup> against all rules on the Bosphorus (i.e.) standards of FRGC v2 and CASIA were utilized, the devised methodology was able to attain 95.2%, 97.6% and 95.6% rates of detection respectively. At this time, the yielded outcomes of detection rates were the best detection rates recorded by using the considered common databases.

Another instant of recognition in the facial expressions were developed by [5] by devising a innovative methodology which made use of next gen several class classification and kernel variable improvisation methodologies along with the feature extracting methodology concerning the support vectored devices. The outline of this proposed method began with the extraction of the intended features from the source vectors, and afterwards, extracted features would be converted into a Gaussian space with the utilizations of compression based sensing operations. This operation assured dimensionality reduction in the case of feature vectors after which matching of the feature vectors would be carried out with radial dependent function kernel utilized in SVMs-Support Vector Machines for the classifying purposes. Even before performing the operation of classification, an improvised variable for training of the SVMs was spontaneously calculated on the basis of the earlier devised method relying on the receiver processing properties of the classifier of the SVMs. By using an improvised kernel variable, the devised innovative several class classification method was utilized for classifying any vectors in the final phase. In every experimental investigations proceeded on the 2 face expression dependent databases by making use of several evaluation methodologies, the devised methodology beats the other literature available similar works when using the alike configurations and database. Yielded outcomes was able to firmly confirm the soundness and benefits of the devised method against all the considered existing works.

Conventional and simple facial expression recognition operations were carried out by [6] after considering the importance of the intensity image and effects caused by the poses, disguise, and illumination. The proposed paper was objected to devise an ideal face expression dependent recognition operation with the utilizations of several channelled attributes and RGB-D pictures. On the basis of visual saliency and picture entropy, facial textured attributes were initially acquired from the RGB pictures and depth pictures would build the descriptors of the HOG-Histogram of Oriented Gradient. As a next step, they obtained RGB pictures' geometric attributes by deploying the AAM-Active Appearance Model. By joining the AAM geometric attribute with the HOG texture attribute, this work constructed an idle several channelled attribute vector for the recognition of facial expressions. By using this as the base, a better SVM dependent methodology like GSSVM, was utilized for classifying recognizing the facial expressions. This devised GS-SVM methodology implements Grid exploration operation for improvising the suitable variables for the classifier of SVM and the accuracy of every variable combinations was estimated in the particular span. At last, the devised methodologies were cross-verified and assessed on the combined database of RGB-D. As far as the experiment based outcomes were concerned, the devised methodology was not only able to attain the better mean rate of the recognition but also was able to be perfect enough for the uncontrollable scenarios.

Image analytics always needs the aid of Labelling methodology for proceeding with the segmentation or classification based upon the requirements of the end user. By keeping in this mind, [7] came up with the several labelled picture classification since it was one among the crucial issue in the semantic dependent processing of the picture. Conventional partially supervised and several labelled learning operations generally gain the knowledge of the functions in the classification technique in a space that was continuously labelled. Neglecting of non-continuous restriction in the semantic labels hinders the fulfilment of the taken up classification operation. This work had particularly taken the non-continuous restriction to devise the DSML-Discrete Semi-supervised Multi-label knowledge gaining methodology for classifying the picture under consideration. Here, in this devised DSML methodology, they made use of the partially supervised outline with the non-continuous restriction. As a next step, they further include anchor graph learning for yielding the optimized scalability to deliver an ADMM dependent varying improvisation

operation for the sake of redressing the considered outlines. When validated experimentally, the yielded outcomes indicate the better performance of the proposed DSML than the remainder of the considered advanced partially-supervised operations.

## **2. SOCIAL ROBOTS APPLICATIONS**

Social robots are the sort of the town for serving many applications involving the robotics utilizations. Along this path, a parallel cum practical periodic outline for mental phases and emotions extraction followed by the recognition was done by [8] by using the human being motions recorded in the video fragments. When setting up for the experimental investigation, our hands were detected with assessment of the skin-colored entities in motion. Analysing the detection rate indicates that frequency and acceleration properties of the detected entities were suitable for the classifying the emotions recorded while the human in motion. The yielded results of recognition of the mental and emotional phases were evaluated and analysed with the help of 2 independent standards of the category CMAs-Certified Movement Analysts whom utilize the operation based on the LMA-Laban Movement Analysis. They told debated that LMA dependent computer investigation could act as the general language for the sake of expressing and understanding emotional motions intermediary to the humans and robots, in the manner resembling the general coding rule intermediary to the perception and action which the humans have. This work also made us realize that primates were constituted by the non-minor neuron regularity. The result was portion of a bigger design on communication intermediary to the humanoid type of robots and human beings with the objective of training social behavioral talents to affected children with autism by using the robots influencing in a non-artificial conditions.

Since comprehending the mental phase of other pupil was one among the crucial task that the intelligent robots and agents have to get adapted to in certain social scenarios, [9] came up with the methodology to deal with this crucial and significant task. But, the mental operations indulged in ‘mind reading’ were always complex foe one to take up. One supporting methodology was to utilize the concept of Simulation, which was working on the basis of the significant portion of the researches carried out in the domain of neuropsychology. Even then, computing the suitable or appropriate computing theory or concept for utilizing in the simulation based emotion identification was more of an astonishing task, beyond the capabilities of our understanding. These difficulties were tried to be resolved by taking up this research by making deployment of both the neuroscience discovering along with Simulation concepts on the Mirror-Neuron regularities as the base for the devised innovative computing methodology as the possibility of handling the facial expressions that were affective in nature. This methodology was devised on the basis of the observations being mapped probabilistically from the variety of entities onto a one rigid entity (‘inner transcoding of stimuli that were external’), and further, onto a latent space (‘concept dependent reply’). In short, by combination of the earlier literature available arrangements and their innovative operations, they were able to give rise to few astonishing preparatory outcomes.

Smarter human-to-robot communication was always a challenging yet interesting research domain in both the arenas of robot maintenance and smartness in computation power. So, [10] had taken up the work involving innovative incrementing and adaptive picture dependent semantics learning outline on the basis of the particular utilization base of the social robots. This makes the robot of these days to possess the capability for learning to identify newer pictures on the basis of earlier communication made between the robots and the human beings. Contradicting the many smart picture dependent semantics learning processes, concentrating on how to identify large scaled training information, this research was concentrating on how to learn picture semantics from the scratch and empower the learning process in much increments under the aid of communication made between the robots and the human beings. In the devised outline, the consumer initially presented unlabeled pictures to a humanoid type of robots for identifying it. Further, the robot would reply back the consumer regarding what was the picture dependent on by making use of partially supervised incrementing learning outline. Successively, the consumer “teaches” the robot with the correct picture label if the robot comes up with unintended answer. Consumers could give several unlabeled pictures to the robot in the devised outline for facilitating the processes of learning and teaching. Additional correlations and experimentations show that the devised methods could give rise to promising outcomes. Their developed outline had wide variety of utilizations including rehabilitation and education.

Robots of these days were keen to work beyond the predefined restricted commercial surroundings, so [11] had taken up this work. These robots also track unarranged and unsearched arenas for the different kinds of utilizations in the practical world as social and service type of robots. Therefore, along with these newer physical restrictions, these robots would also deal with the human being. The above discussed prospective make us to take the communication made between the robots and humans beings into account right from the starting point when designing. The chances for the robot to identify emotions of its consumers along with the need to appropriately respond and behave in a particular manner. This robot capability might impart a base part in the society’s amalgamation. But, this capacity was still well beyond the reach of ours. In the last few decades, many researches to apply automatic concepts to several utilizations beyond the industries scopes have been taken up by many researchers and philosophers. However, only some utilizations have made efforts to take the emotional phase of the consumers in the robotic behavioral systems, so this gap makes one to arise numerous queries like- how could the human being’s emotions be modelled for an exact indication of their mind phase? Which classification methods and which sensing modalities should be suiting more apt to acquire the desired skills? To add, which utilizations were appropriately suiting the robot to comprise this kind of sensitivity

values? Along all these ways, this work was objected for giving a common preview of the earlier researches to facilitate the robots to identify the emotions of the human beings and communicate accordingly.

### **3. NEURAL NETWORK BASED ROBOTIC RECOGNITION**

Human beings could observe and absorb newer facts from the new and earlier unknown surroundings could give rise to the new levels of exposures that might be very provoking for the artificially developed robotic regularities. Here, a cognitive type of arrangement was proposed by [12] which deploys emotional smartness to comprehend innovative theories from the earlier unknown type of exposures. Fundamental concept was that the appraisal of the emotion related exposure expressed internally as numerous of MoNADs would aid the devised arrangement for identifying novel concepts and easy the production and gaining of newer theories. Aiming the measurement of emotion dependent cognition effects on learning, the arrangement was applied in the robot and investigated with numerous of standards consisting variety of colour related configurations. Many crucial discoveries were made in this work. First to start with, the modal neural networks' dynamic phase doesn't converge with few of the attracting mediums while receiving the input of unknown type. Subsequently, in addition to the above finding, the devised methodology converges in a faster manner to an attracting medium in reply to input that was known familiarly. By using the parameter of period, the devised methodology developed the capacity for learning the earlier unknown characters and theories as the newer MoNADs. Here, in this methodology, it was said that the model triggered a subliminal reply of the human brain to the surroundings which were unknown. Discoveries of this work too contributed much towards the wide utilizations for the futuristic emotion dependent artificial smartness concepts.

The usage of DNNs-Deep neural networks in the form of a learning model for modelling the human being's brain hierarchical arrangement were taken up by many of the researchers and academicians. But, DNN experiences some issues of computation related complexities and learning efficiencies. For redressing the above said issues, DSAN-Deep Sparse Auto encoder Network was utilized by this work of [13] to learn facial attributes, which takes the hidden units' sparsity to learn high-level arrangements. In the same time, SR- Softmax Regression was also deployed for classifying expression attribute. This work made use of SRDSANs-SR dependent DSANs for identifying the face associated emotions by the communication made between the robots and human beings. This methodology was objected for dealing with the large information in the deep learnings' output with the utilizations of SRs for the sake of out beating the issues of gradient diffusion and local extrema occurring in the training operation, for finely tuning the total network related weights to attain the global optimum, thereby making the whole neural networks' depth ideal to make it perfect, thereby enhancing the fulfilment of the facial emotion recognition. Yielded outcomes indicates that the mean identification accuracy of SRDSAN was better than the utilized SRs along with the earlier literature available neural networks. The preparatory utilization investigations were performed in the devised ESRS-Emotional Social Robot System by using 2 robots which were mobile, wherein emotion based social robot was capable enough for identifying emotions like angry and happiness.

Another kind of neural network cum deep learning methodology of CNNs-Convolutional Neural Networks were rapidly becoming famous when the large sized picture identification, detection, classification, and localization were to be performed. Earlier CNNs methodologies make use of only one system for extracting the attributes and identification accuracy of the utilized systems was not enough for the practical utilizations. For the sake of augmenting the identification accuracy, an ECNN-Ensemble of Convolutional Neural Networks dependent facial identification was devised by [14]. This devised methodology provides the resolutions for the issues of low resolution, facial expression difficulties, aging, and pose deviations. This devised ECNN method performs better with the earlier available state of the art systems such as Xception, Inception-v3, VGG16, ResNet50, and VGG19 CNN systems comprising the category-5 accuracy of 97.12 percent on the Face dataset of Web and 100 percent on face dataset of the YouTube.

Study of robotic facial expression had a range of computer-human interactions purposes. Traditional methods investigate as a classification function primarily prototypical facial expressions comprising no more than eight unique emotions. But, in practice in a naturalistic environment spontaneous facial expressions can portray not only a mixture of emotions, but also different intensities within a cluster of emotions. Such strategies are not efficient or sufficient in these instances. Hence, this work [15], they devised to train deep CNNs-Convolutional Neural Networks to examine explainable facial expressions in a paradigm of dimensional emotions. This proposed approach encompasses either a set of basic expressions of emotion, or a full range of certain emotions and dramatic frequencies of emotion that we both feel in ourselves and perceive in others in our daily routines. Interestingly, they initially projected facial expressions into dimensional metrics because they converted the examination of facial expression from an issue of classification in to one of regression. They further evaluated our facial expression regression approaches on the basis of CNNs and these approaches demonstrated impressive outcomes. In addition, a bilinear pooling encoding 2nd-order attribute statistics optimized our method. They had indicated that such bilinear-CNN frameworks exceeded their respective benchmarks considerably.

Auto interpretation of face expressions with variations in scale, pose occlusions and alterations which would greatly enhance the normal interaction of humanoid robots. In order to tackle such difficulties, this research work of [16] endorses

unattended instant facial point identification incorporated with regression-based frequency assessment for facial AUs-Action Units and emotion grouping. Utilizing Gabor sorting, BRISK-Binary Robust Invariant Scalable Key Points, an ICP-Iterative Closest Point process and fuzzy c-means grouping, the devised face point predictor could identify fifty four facial points in pictures of faces with occlusions, pose differences and scaling discrepancies. In particular, ICP was first extended to produce neutral landmarks for the occluded facial components in conjunction effectively with pictures comprising occlusions. After which FCM was utilized by taking into consideration the firsthand knowledge of the un-occluded facial components to perhaps predict the structure of the occluded facial zone. Post-mark comparison optimization must then be implemented to extract the optimal geometry for the occluded face element to even further alter the neutral ICP-generated landmarks and recreate the occluded facial zone. They just utilize neural networks and support vector regression for 18 defined AUs together to execute AU strength computation. Only essential feelings, as well as neutral gestures, were subsequently identified by FCM. This also demonstrated great capability to address compound and recently arrived category identification of newer emotions. The overall regularity was combined with a robot, empowering it to accomplish difficult tasks of identification of facial emotions in everyday life.

#### **4. IOT DEPENDENCY ON AI**

The IoT- Internet of Things [17] was a modern Internet movement that was rapidly being incorporated with tremendous success in business manufacturing, education, financial, as well as other modern data society fields. In general, artificial intelligence-supported IoT greatly improves performance in a wide range of everyday applications with applications of the dominant business, transport, robotics, manufacturing, and automation systems. Aim of this paper was to provide a global overview of key issues related to IoT and AI synergy, including current and prospective opportunities of immense worth to society. Beginning with an overview of the domains of IoT and AI, the paper described what the idea of ' IoT-AI synergy ' implies, demonstrates the factors driving the growth of ' IoT allowed by AI ' and discusses the principles of 'IIoT- Industrial IoT ' , ' IoRT-Internet of Robotic Things ', and 'IAIoT- Industrial Automation IoT '. A variety of case studies were then illustrated and few IoT / AI-assisted robotics and industrial automation technologies are eventually addressed.

#### **5. MISCELLANEOUS ROBOTIC AND INTERDISCIPLINARY APPLICATIONS**

In this work they present Human-like Sensing or ' ' 5 sense computing " as a possible futuristic Internet-of-Things extension. We explore how 3D optical vision, thermal perception, acoustic profiling, smell and tactile sensing could aid with solutions for remote evaluation and inspection. As a potential integrated system for collecting these data, they proposed a robot mounted opto-thermal and acoustic sensing device. We present outcomes of field studies with the proposed system and illustrate how such systems could provide appropriate strategies for remote evaluation and inspection.

Accidents and attacks involving materials that were biological, chemical, radiological / nuclear or explosive were uncommon, but may have a major impact. Because investigating these incidents is not the habitual work of anyone, a variety of AI approaches could reduce the cognitive load of researchers and help decision-making, such as: Scene assessment planning; ongoing risk assessment and revising; autonomous car monitoring for object and sensor data gathering; photo / video review of points of interest; detection of defects; and information collection. Practical simulations may support the creation and testing of AI dependent technologies due to the unusual and elevated-risk complexity of these incidents. We constructed practical CBRNE situation structures and put in place a preliminary range of tools.

[20] Face recognition had received considerable attention in recent years, one of the basic subjects of computer vision and pattern identification. Furthermore, the facial pictures were reshaped to a lengthy vector for these traditional identification methodologies, thus losing some of the earlier spatial restrictions of each pixel. A fresh tensor dependent extraction feature algorithm called tensor reservoir preserving discriminant analysis was suggested in this paper for facial picture identification; the proposed method includes two phases: first, a low-dimensional tensor subspace was obtained from the initial input tensor samples; second, a discriminative locality alignment was used to obtain the ultimate representation of the vector function for subsequent facial recognition. On the one hand, the suggested TRPDA methodology makes full use of the input samples ' natural architecture and introduces an optimizing parameter that can handle the problem of tensor spectral analysis directly, thus reducing the cost of computing relative to those conventional tensor dependent selection techniques. At the other side, the suggested TRPDA methodology extracted feature by searching a subspace for the tensor that maintains many of the internal-class source survey rank order data. In order to determine the efficacy of the proposed TRPDA algorithm, tests on the three facial databases are carried out here.

[21] Human beings change their facial expressions to reflect their inner state and quite often confuse people about their real emotional state. Proofs in experimental psychology indicated that there were brief and subtle discriminative facial reactions. It means that it would be easier to determine such behaviour if filmed at a higher frame rate in higher resolution. We suggest SASE-FE, the first database of congruent and incongruent facial expressions with underlying emotional states. They demonstrate that

ultimately, studying spatio-temporal representations of the information can address the problem of identifying whether facial gestures are expressions of genuine emotions or not. In addition, the proposed approach enhances state-of-the-art performance for video emotion detection on CK+ and OULU-CASIA datasets and achieved competitive outcomes when classifying BP4D facial action units.

[22] It was important for a robot to direct human attention and actions in order to interact smoothly with people. They constructed a model of human visual attention in this research to direct human attention dependent on a study of a fulfilment of magic tricks. They calculated human gaze points of people watching videos of a presentation of a magic act and correlated them to the position in which the magician wanted to draw the attention of a spectator. The analysis showed that the relationship between the eyes, hands, and gaze of the magician plays a major role in guiding the attention of the spectator. The analysis using the observed gaze points revealed that, when people watch a video of a magic trick success, the proposed model could better explain human visual attention than the saliency graph.

[23] Recognition of FER-Facial expression was extended to HRI-Human-Robot Contact. An assistant robot with near human interaction should be able to recognize the expression of the human face. FER was a non-trivial issue because each individual has his or her own way of revealing his or her emotion, and the facial expressions of two different individuals may not be completely identical. The expression of the face can be divided into four phases: positive, original, peak, offset, and then negative. They utilize Gabor filter in the first stage to obtain the facial features and implement SVMs to classify AUs. They utilize random forest classifiers in the 2nd stage, based on the identified AUs, to recognize the facial expressions. Finally, they showed the outcomes of the experiment and compare our method to the other methods.

[24] Past approaches to the study of facial expression modelling and analysis use three different techniques: units of facial movement, geometric features and modelling based on graphs. These techniques have an interrelationship. By using these mappings between major geometric features involved in facial expressions and the subset of facial action units whose presence or absence is specific to facial expression, the study of facial expression is significantly improved. This paper blends dimensional reduction and object identification techniques with search space pruning achieved through this special sub-set of facial action units to substantially prune the search space. The performance results indicate a 70 percent improvement in performance over time while preserving the accuracy of emotion recognition on the public facial expression database.

[25] This research note concentrates on emotions symbolized in a moving humanoid NAO robot being recognized by humans. Attendees whose facial expressions were registered and analysed with a Noldus FaceReader were provided with emotional gestures emulated on NAO intended to cause joy or sadness. Our preliminary results suggest accurate emotion recognition in modelling the affective movements of the robot using the Laban choreographic method.

[26] What is the correct representation of objects for manipulation? We want robots to perceive scenes visually and learn about the objects in them that (i) objects in them that (i) is task-agnostic and can be used as a building block for a variety of manipulation tasks, (ii) is generally applicable to both rigid and non-rigid objects, (iii) take advantage of good 3D vision priors and (iv) benefit entirely through self-surveillance. Regarding previous methods, this was hard to achieve: a lot of recent grasping work does not apply to grasping specific objects or other tasks, while task-specific training can take several tests to categorize well across object settings or other activities. In this study we propose Dense Object Nets as a coherent representation of objects for visual comprehension and manipulation, based on the recent advances in self-supervised dense descriptor training. We show that for a wide variety of previously unseen and potentially non-rigid objects, they can be trained rapidly (about 20 minutes). We also present new contributions to allow multi-object descriptor learning and demonstrate that we can either acquire descriptors which generalize across object classes or descriptors that are distinct for each object instance by changing our training method. At last, they show the novel approach of robotic manipulation of trained complex descriptors. They prove grasping specific points on an object across potentially deformed object setups, and showing how to transfer specific grasps through entities in a class utilizing class particular descriptors.

## 6. CONCLUSION

This paper discussed and correlated fundamental image analytic methodologies along with its corresponding utilizations in the robotics industry with reference to the image processing techniques. Moreover, the image processing operations taking up the facial/emotion recognition is discussed in a brief preview in this paper with reference to the utilized methodologies. After discussing many existing/current methodologies used in the facial recognition related applications deployed in the robotics industry, futuristic research direction is also given below.

Another highly valuable subject in the implementation and application of IoT / IIoT-AI is the study of ethics / morality which sets out the principles and rules to be implemented in the field to ensure the ethical / moral use of IoT / AI in everyday applications

## REFERENCES

- [1] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *Journal of personality and social psychology*, vol. 17, p. 124, 1971.
- [2] J. L. Salmeron, "Fuzzy cognitive maps for artificial emotions forecasting," *Applied Soft Computing*, vol. 12, pp. 3704-3710, 2012.
- [3] J. Dong, N. Karianakis, D. Davis, J. Hernandez, J. Balzer, and S. Soatto, "Multi-view feature engineering and learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3251-3260.
- [4] B. Y. Li, M. Xue, A. Mian, W. Liu, and A. Krishna, "Robust RGB-D face recognition using Kinect sensor," *Neurocomputing*, vol. 214, pp. 93-108, 2016.
- [5] A. M. Ashir, A. Eleyan, and B. Akdemir, "Facial expression recognition with dynamic cascaded classifier," *Neural Computing and Applications*, pp. 1-15, 2019.
- [6] L. Cai, H. Xu, Y. Yang, and J. Yu, "Robust facial expression recognition using RGB-D images and multichannel features," *Multimedia Tools and Applications*, vol. 78, pp. 28591-28607, 2019.
- [7] L. Xie, L. He, H. Shu, and S. Hu, "Discrete Semi-supervised Multi-label Learning for Image Classification," in *Pacific Rim Conference on Multimedia*, 2018, pp. 808-818.
- [8] T. Lourens, R. Van Berkel, and E. Barakova, "Communicating emotions and mental states to robots in a real time parallel framework using Laban movement analysis," *Robotics and Autonomous Systems*, vol. 58, pp. 1256-1265, 2010.
- [9] J. Vitale, M.-A. Williams, B. Johnston, and G. Boccignone, "Affective facial expression processing via simulation: A probabilistic model," *Biologically Inspired Cognitive Architectures*, vol. 10, pp. 30-41, 2014.
- [10] H. Zhang, P. Wu, A. Beck, Z. Zhang, and X. Gao, "Adaptive incremental learning of image semantics with application to social robot," *Neurocomputing*, vol. 173, pp. 93-101, 2016.
- [11] F. Cavallo, F. Semeraro, L. Fiorini, G. Magyar, P. Sinčák, and P. Dario, "Emotion modelling for social robotics applications: a review," *Journal of Bionic Engineering*, vol. 15, pp. 185-203, 2018.
- [12] K. Kushiro, Y. Harada, and J. Takeno, "Robot uses emotions to detect and learn the unknown," *Biologically inspired cognitive architectures*, vol. 4, pp. 69-78, 2013.
- [13] L. Chen, M. Zhou, W. Su, M. Wu, J. She, and K. Hirota, "Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction," *Information Sciences*, vol. 428, pp. 49-61, 2018.
- [14] V. Mohanraj, S. S. Chakkaravarthy, and V. Vaidehi, "Ensemble of Convolutional Neural Networks for Face Recognition," in *Recent Developments in Machine Learning and Data Analytics*, ed: Springer, 2019, pp. 467-477.
- [15] F. Zhou, S. Kong, C. Fowlkes, T. Chen, and B. Lei, "Fine-grained facial expression analysis using dimensional emotion model," *arXiv preprint arXiv:1805.01024*, 2018.
- [16] L. Zhang, K. Mistry, M. Jiang, S. C. Neoh, and M. A. Hossain, "Adaptive facial point detection and emotion recognition for a humanoid robot," *Computer Vision and Image Understanding*, vol. 140, pp. 93-114, 2015.
- [17] S. G. Tzafestas, "Synergy of IoT and AI in modern society: The robotics and automation case," *Robot. Autom. Eng. J*, vol. 31, pp. 1-15, 2018.
- [18] A. Pal, R. Dasgupta, A. Saha, and B. Nandi, "Human-like sensing for robotic remote inspection and analytics," *Wireless Personal Communications*, vol. 88, pp. 23-38, 2016.
- [19] D. L. Smyth, J. Fennell, S. Abinesh, N. B. Karimi, F. G. Glavin, I. Ullah, *et al.*, "A virtual environment with multi-robot navigation, analytics, and decision support for critical incident investigation," *arXiv preprint arXiv:1806.04497*, 2018.
- [20] D. Tao, Y. Guo, Y. Li, and X. Gao, "Tensor rank preserving discriminant analysis for facial recognition," *IEEE transactions on image processing*, vol. 27, pp. 325-334, 2017.

- [21] K. Kulkarni, C. Corneanu, I. Ofodile, S. Escalera, X. Baro, S. Hyniewska, *et al.*, "Automatic recognition of facial displays of unfeigned emotions," *IEEE transactions on affective computing*, 2018.
- [22] Y. Tamura, T. Akashi, S. Yano, and H. Osumi, "Human Visual Attention Model Based on Analysis of Magic for Smooth Human–Robot Interaction," *International Journal of Social Robotics*, vol. 8, pp. 685-694, 2016.
- [23] S.-C. Hsu, H.-H. Huang, and C.-L. Huang, "Facial expression recognition for human-robot interaction," in *2017 First IEEE International Conference on Robotic Computing (IRC)*, 2017, pp. 1-7.
- [24] M. Ghayoumi and A. K. Bansal, "Unifying geometric features and facial action units for improved performance of facial expression analysis," *arXiv preprint arXiv:1606.00822*, 2016.
- [25] K. Arent, M. Gakis, J. Sobecki, and R. Szczepanowski, "Toward emotion recognition embodied in social robots: implementation of laban movement analysis into NAO robot," *Journal of Automation Mobile Robotics and Intelligent Systems*, vol. 11, 2017.
- [26] P. R. Florence, L. Manuelli, and R. Tedrake, "Dense object nets: Learning dense visual object descriptors by and for robotic manipulation," *arXiv preprint arXiv:1806.08756*, 2018.

#### About the Author



SathyaNarayanan is working with Topcoder and comes with rich background in Data, Analytics and AI space. Sathya already has various research journals published on AI, Cloud and IoT He also sits in the advisory council of Computer Society of India and looks forward to collaborate with young and efficient researchers in Data, Analytics and AI space