Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 11
November 2024

Advancement in the Students' Learning Potency and Machine Learning Insights: Designing the Nexus

[1] Sumant Sekhar Mohanty , [2] Samyo Jagdev, [3]Subhasish Swain
[1]Computer Science and Engineering**Gandhi Institute of Excellent TechnocratsBhubaneswar, India**
[2]Computer Science and Engineering Gandhi Institute of Excellent TechnocratsBhubaneswar, India
[3]Computer Science and Engineering Gandhi Institute of Excellent TechnocratsBhubaneswar, India

Abstract:
The advent of artificial intelligence is revolutionizing the world with cutting-edge technologies similar to machine learning, deep learning, computer vision, and robotics process automation. These innovations become enabling humans to minimize tangible exertion, optimize time, and achieve unparalleled precision in tasks that were previously challenging or impossible for individuals to perform. As research in this field continues to advance, a wide range of applications has emerged, including self-driving cars, facial recognition, emotion detection, speech recognition, and others. Here in this study, it has been developed an intelligent classroom system utilizing computer vision and emotion recognition to analyze multiple students' facial expressions, detect various emotions during lectures, and predict their performance in upcoming exams. Today, machines are increasingly assuming roles similar to humans in decision-making, recommendations, patient diagnosis, and recognition tasks. The future may soon see machines commanding humans, as ongoing research focuses on enabling machines to visualize, detect, and recognize individuals and objects through computer vision. Leveraging this technology, our research has led to the development of a deep learning-based system for recognizing students, interpreting their facial expressions, and gauging their attention levels.
Keywords: - Expert system, Intelligent agent, Deep learning, Eigen-faces, ANN, CNN

Introduction

There are several ways which create disturbance inside a classroom or interrupt the tutor's lecture. Students inside class are sometimes distracted from lectures or starts gossips within friends this is very common happening This research has two different module kinds of work first is to identify the students present in classroom and then capture their presence in every lecture by face recognition system and second model of this research is to find the particular students facial emotion during lectures. This can help to identify the interest level of students about that subject topic during a class lecture, and we can predict their performance by this data. Any other school can use this application, college, universities, or coaching classes to find the attention level and interest of the student for a better result in the upcoming exam. The designand the development of a system that can learn to interact socially with humans like whenever a student is surprised or confused about any topic a display monitor will show some massage after recognition his/her face and emotion. Technically, the research consists of training a deep neural network with faces of students those who are already studying and will read their emotion as the parameter of it labels it. It is a multidisciplinary project involving affective computing, Machine Learning, Deep Learning and computer vision. All these differences are fields interrelated, and to understand how they can provide solutions to complex problems is another project's goal.Although Facial Recognition Systems (FRS) have shown impressive results in controlled settings, they face significant difficulties when dealing with faces in real-world, uncontrolled environments, where variations such as sunglasses, facial hair, hairstyles, and aging can greatly impact accuracy. To address these challenges, the exclusive target is to cultivate a unit convenient in recognizing faces in open atmospheres. Ultimately a tool namely Convolutional Neural Networks (CNNs) has been implemented for the sake of powerful parameter withdrawal technique. Intelligence system is that machine which can learn quickly, assume that we are a tutor and we have 50 students' in class then how we will recognise which student is intelligent? Which can be done merely by assuming learning ability of particular student like the ability to gain knowledge, calculate task, perceive and quick learner from experience, retrieve information & explore knowledge, solve problems and complex ideas, use and adapt language fluently and new situations. Like a human being a device, machine or a system can be called intelligence if it has these similarities:

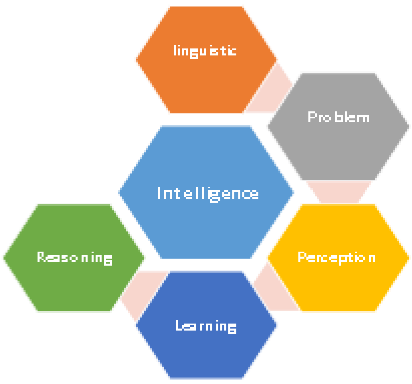

Figure 1: Intelligence System

Reasoning: If it can enable itself for judgment, decision making (E.g. fuzzy rule),and prediction.
Learning: If it is learning from the environment and act of gaining knowledge byprevious data or experience.
Perceptual Learning: If it can recognise things or objects, people and surroundings which have seen before. E.g. identifying and classifying objects.
Problem Solving: Able to solve problems like human
Linguistic Intelligence: Able to understand human language and responseaccordingly.

Expert System

An expert system is a system which mimics like a human being, who can think, decide like a human. Which was introduced by the researchers at Stanford University Expert System is itself an Artificial Intelligence where the machine can determine human-being. As an example, if rainfall occurs, Rahul will not go to school. If rains do not occur, Rahul will surely go to school. The above rainfall condition shows the ability to decide humans and producing accurate output for the small knowledge base.

Knowledge Base

Knowledge is the store of data or facts which can be pulled out when needed in meaningful forms. The data is a collection of different types of events in any format, i.e. word file, mp3, JPEG, etc. Data, information, and experience combined labelled as knowledge.

Intelligent Agent

Being an entity inside the domain of Artificial intelligence, intelligent agent (IA) takes input from surroundings and observe as it is via sensors executes upon the ambience taking actuators like Google self-drive car, which receives a response from camera and sensor and acts on the actuator and achieves goals.

Artificial Neural Network

Inspired from biological human brain system equivalents, Artificial Neural Networks (ANN) are become units, organised manipulative nodes, include any shapes. ANN is something like we make fuzzy rules with multiple inputs to get one output or more input with various inputs, the input we give, i.e. neurons it calculates with some condition or weights and then itgenerates some decision or output like an expert system or a human ideology or better than human perception. It has been justified that the ability of various issue like Computer Vision speech recognition, etc. become cumbersome to sort out via traditionally withdrawing parameters which can be quickly done by Deep Learning which is nothing but a combination of Neural Networks.

The power of Artificial Neural Networks (ANNs) lies in their composition of simple computing nodes, known as neurons, which are organized into layers. These layers are interconnected, similar to the connections between biological neurons via axons. The layers arecategorized into three primary types: input layer, hidden layer, and output layer. The input layer receives the data fed into the network and is connected to one or multiple hidden layers, which are not directly observable, hence their name. Alternatively, the term "hidden" can be interpreted as their functionality being opaque, even after training, making ANNs sometimes referred to as "black boxes" due to their complexity. The hidden layers can be single or multiple, and each is connected to the previous layer or compared to each other. Each neuron inside the concealed and yield layers is strongly linked to each and every neurons in the preceding layer, with each connection assigned a weight that represents the strength of the relationship between connected neurons. Similarly, in the same way how biological neurons are connected, which we have studied during our school days. Finally, the last layer and third most layers are called the output layer, whose task is to produce and execute the outcome to build ANN, considering single yield per class.

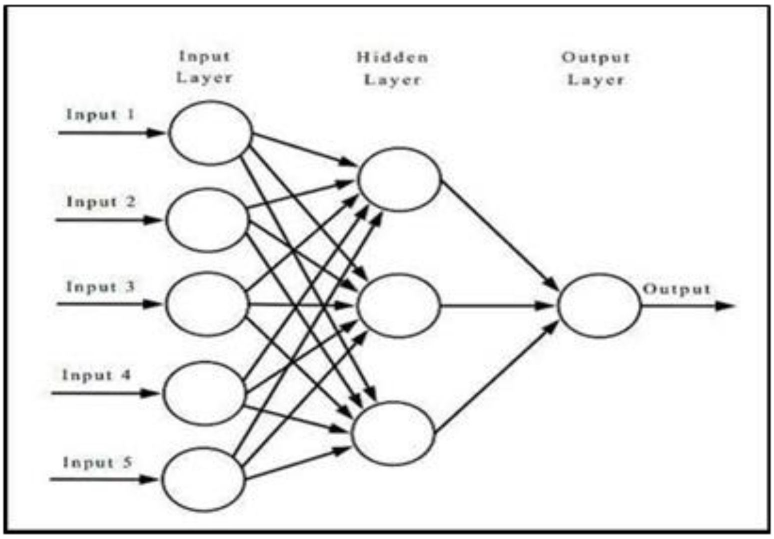

Figure 2: ANN with Multiple Inputs

This algorithm commences by initializing the network's weights, which can be done using various strategies, such as drawing from a probability distribution or random assignment with low values recommended. The process then enters a recurrent loop consisting of three phases. At the opening phase, the very network witnessed for the sake of input communication with and yield numbers become manipulated. Next, the yield is assessed by means of a defeat utility, comparing the predicted output (o) to the actual output, to determine the network's error margin. The final phase involves informing separately mass to reduce the fault, achieved by computing the gradient of individual neuron, representing a step towards the optimal value. The algorithm can be work towards multiple iterations to achieve the best possible outcome.

The gradient descent process unfolds as follows. After calculating the outputs, we determine the error (e) for each output neuron (o), denoted as (o). This error value enables us to compute the gradient of each (o). To do this, we need to find the partial derivative (PD) of the activationfunction with respect to the input (Xo) of the output neuron (o). For the logistic regression case, this becomes:

$$\frac{\partial o}{\partial Xo} = \frac{\partial}{\partial Xo} \varnothing(Xo) = \varnothing(Xo)(1 - \varnothing(Xo))$$

This highlights the importance of using continuously differentiable activation functions, as they allow us to obtain this partial derivative, which is crucial for training the network.

Next, we calculate the gradient on adding the partial derivative including the error monitored. This gradient facilitates the adjustment of the weights of all output neurons, guiding them towards their optimal values. Similarly, we need to do for all segments by adjusting in their weights, except for the input one. To do so, each layer needs to be calculated before reaching to next layer, because of this property it is called back-propagation.

Figure 3: Example of Eigen-faces

Convolutional Neural Network (CNN)

Each layer receives the output from the previous one as input and learns superior standard parameters, as illustrated in the above diagram. Remarkably, the communication at some instances itself utilizes these features to produce an output, while in other cases, it generates them for use by other methods. Deep Learning, also known as advanced machine learning, is essentially a type of Neural Network comprising multiple layers. These networks can automatically discover a regular greatly sensitive parameters, during the application against computer vision. Experimental evidence has shown that these features often outperform manually designed ones, with the added benefit of eliminating the need for manual feature engineering, as the network automatically learns them.

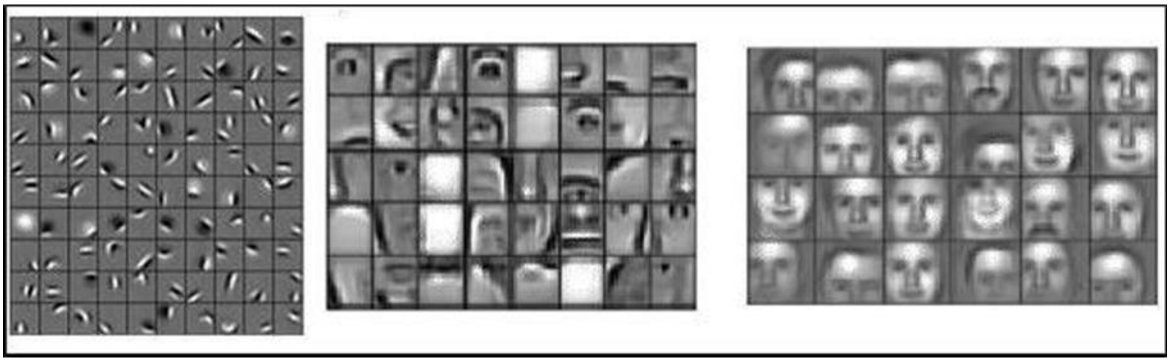

Figure 4: Facial features extracted by a CNN into 3 levels

Furthermore, the learned features can be highly abstract, as seen in Figure 4. Notably, the functionality of CNNs bears a striking resemblance to the biological visual system, although the reason for their success may be beyond the scope of this document. Nevertheless, their impressive results make them a compelling choice for addressing Computer Vision (CV) challenges. In recent years, many of the most successful CV applications have employed CNNs, and this trend is expected to continue. Therefore, this thesis leverages CNNs in its work.

In the field of Face Recognition (FR), two standout applications of CNN are Deep-Face and a similar method. These two have consistently delivered state-of-the-art performance in recent years, with the second approach achieving the highest accuracy. Although alternative methods, including Joint Bayesian techniques, have shown promising results, we opted to focus on CNN. The reason weren't solely result-driven, however additionally interest-driven, as we tendto be in person curious about operating with them.

Result of Emotion Recognition

A significant portion of this project's work was conducted using Jupyter Notebook in an unstructured approach, facilitating effortless data exploration and experimentation with various data pre-processing techniques, model training, finetuning, and evaluation. Common functions were extracted and grouped into library files according to their purpose. preprocessing.py contains functions used during pre-processing steps, such as face angle estimation, image resizing, and cropping to the face region. ui.py houses the implementation of all classifiers, while the wrapper framework is implemented as abstract base classes. The specific classifier is defined by simply extending the relevant base class and providing the learned model and implementation of the feature extraction function. This project file also contains classes Video Processor and Sequence Processor, which handle loading and managing the data source. These Processor objects call a callback function (provided during initialization) for each frame. Additionally, classes Emotion Display and Emotion Plot are defined for displaying the classification results. This file also includes functions specific to landmark features, such as calculating distance features or performing COG-centric normalization. The primary purpose of the main application is to showcase the performance of the implemented models and the deployability of the wrapper framework. The application accepts two command-line arguments: the model name and the sequence source, which can be a webcam, video stream, video file, or a directory containing images. Upon launching, the application opens three windows: one displaying frames with face-aligned landmarks, another showing a bar chart of the current frame's emotion probability classification, and a third window containing a line plot of the classification development over time. Non-essential functionality for the main application was consolidated into supporting scripts, which perform tasks such as data transformation for each source dataset, dataset combination, data pre-processing, deep model training, and evaluation of the trained models.

Conclusion

In this study, a deep learning-based system was developed to classify facial emotions in real- time video images, tackling a challenging problem that has been explored using different methods. Building on the success of feature engineering-based approaches, this researchfocused on feature learning, a key advantage of computer vision and deep learning. With the successful creation of a facial recognition and emotion detection system, future directions include deciphering facial expressions and measuring attention levels. Additionally, the integration of machine learning and student learning potential has the potential to revolutionize the education sector.

Future Scope

As educators and researchers, our ongoing quest should be to discover innovative methods for combining these fields, ultimately empowering students to become proficient machine learning practitioners and effective learners. By doing so, we can tap into new avenues for academic accomplishment and prepare the next generation of leaders to succeed in a complex, technology-dominated world.

References

[1] Sun, Y., Chen, Y., Wang, X., & Tang, X. (2014). Deep learning face representation by joint identification-verification. In Advances in neural information processing systems (pp. 1988-1996).

[2] Tan, X., & Triggs, W. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. IEEE transactions on image processing, 19(6), 1635- 1650.

[3] Lee, K. C., Ho, J., & Kriegman, D. J. (2015). Acquiring linear subspaces for face recognition under variable lighting. IEEE Transactions on Pattern Analysis & Machine Intelligence, (5), 684-698.

[4] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436.

[5] Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015, September). Deep face recognition. In bmvc (Vol. 1, No. 3, p. 6).

[6] Sun, Y., Wang, X., & Tang, X. (2013). Deep convolutional network cascade for facial point detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3476-3483).

[7] Sun, Y., Chen, Y., Wang, X., & Tang, X. (2014). Deep learning face representation by joint identification verification. In Proceedings of

the IEEE conference on computer vision and pattern recognition (pp. 76-83).

[8] Farfade, S. S., Saberian, M. J., & Li, L. J. (2015, June). Multi-view face detection using deep convolutional neural networks. In Proceedings of the 5th ACM on International Conference on Multimedia Retrieval (pp. 643-650). ACM.

[9] Shan, C. (2012). Learning local binary patterns for gender classification on realworld face images. Pattern recognition letters, 33(4), 431-437.

[10] Shrivastava, A., Pfister, T., Tuzel, O., Susskind, J., Wang, W., & Webb, R. (2017). Learning from simulated and unsupervised images through adversarial training. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2107-2116).

[11] Ranjan, R., Patel, V. M., & Chellappa, R. (2017). Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(1), 121-135.

[12] Shaogang, G., & Alexandra, P. (2000). Dynamic vision: from images to face recognition.World Scientific.

[13] Ajay Kumar Singh, Venus Bansal (2018). SVM Based Approach for Multiface Detection and Recognition in Static Images. In Journal of Image Processing and Artificial Intelligence (pp. 1-11).

[14] Castillo, R. V. C., Catanaoan, M. J. T., Obliopas, J. F. E., & Linsangan, N. B. (2018, September). Class Attendance Generation through Multiple Facial Detection and Recognition Using Artificial Neural Network. In Proceedings of the 2018 International Conference on Machine Learning and Machine Intelligence (pp. 38-42). ACM.