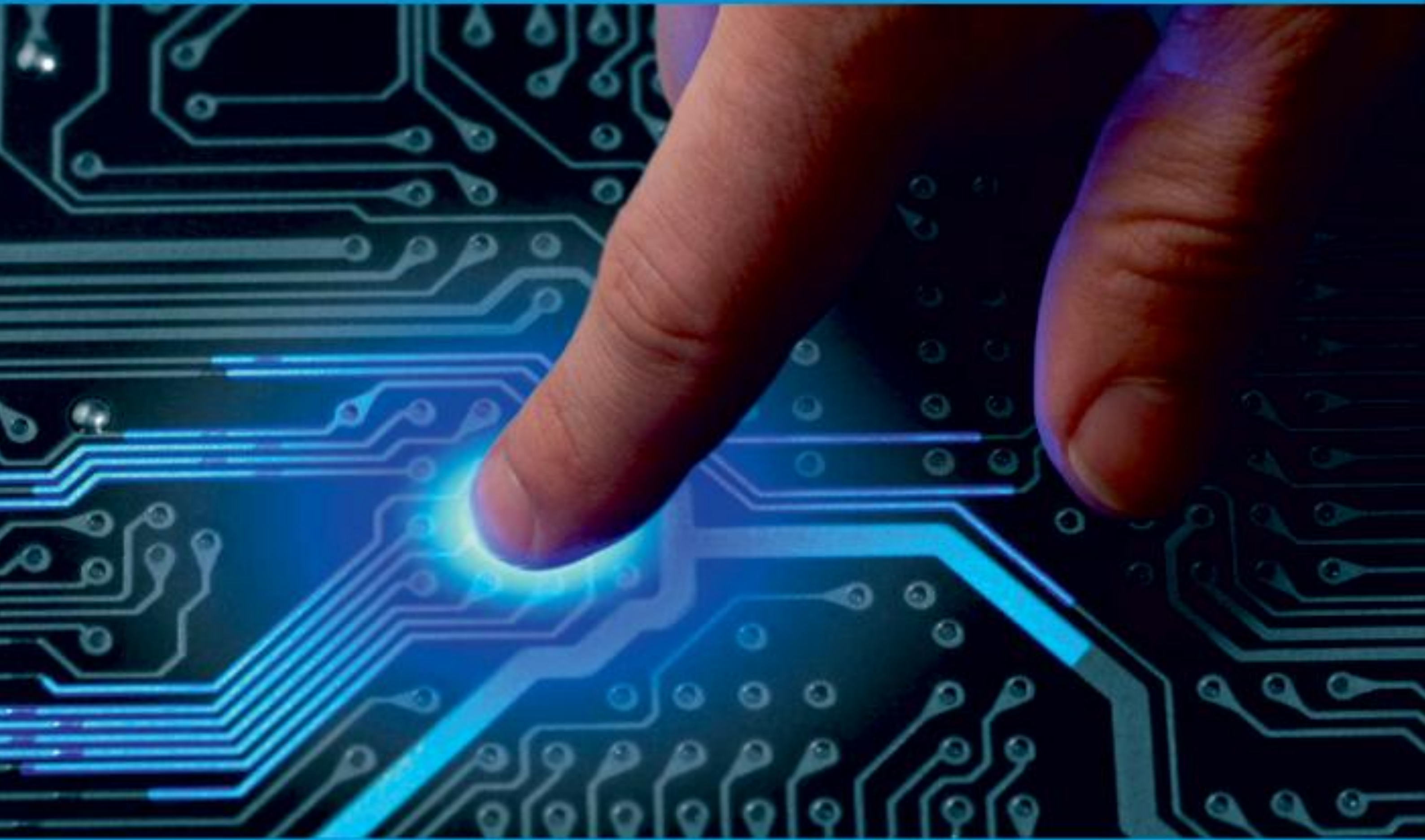
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005233 |

Hand Gesture Recognition Using CNN

Bala Srinivasa Reddy[1], Bejjam Harsha Teja[2], Dokku Nandini[3], Chalapaka Jahnavi[4], V Ramachandran[5]

U.G. Student, Department of Computer Science and Engineering, Vasireddy Venkatadri Institute of Technology, Nambur, Andhra Pradesh, India [1,2,3,4]

Professor, Head of Department of Computer Science and Engineering, Vasireddy Venkatadri Institute of Technology, Nambur, Andhra Pradesh, India [5]

ABSTRACT: Our daily lives have become increasingly dependent on computers, which are employed in a variety of industries. Human-computer interaction can be facilitated by the use of hand gestures. The position and shape of gestures varies from one person to the next. As a result, this problem exhibits non-linearity. Convolutional neural networks (CNNs) have recently been shown to be superior for picture representation and classification in recent study. A static hand gesture detection approach utilising CNN was developed in this study because CNN can learn complex and non-linear correlations between images. More than 20,000 samples will be used to train the model, and more than 7,000 samples will be used to validate it. The Hand Gesture Recogniser can identify hand gestures and display text indicating what those gestures signify. The project introduces a hand gesture detection programme that uses image processing and deep learning. Convolutional neural networks, in particular, are employed in deep learning. At least once, the system is trained on several types of hand motions. The system then tries to recognise the model's hand gestures and displays the corresponding text on the screen. When we train the model, we utilise 2D static hand movements, which we anticipate it to recognise in test data. In terms of hand gesture recognition, CNN is the best.

KEYWORDS: Convolutional Neural Networks, Static Hand Gesture Recognition, Human Computer Interaction.

I. INTRODUCTION

People are connecting to computers in more ways than ever before because to the rapid advancement of technology. Standard modes of human-computer interface such as remote controls, keyboards, and mouse are no longer the most up-to-date means of communication. Because they necessitate a lengthier period of time spent in front of the screen. It's the most adaptable, up-to-the-minute, and current method of communication. Many of the most recent electronic products now provide these features.

With specific movements, we're offering a new method to engage with the system. A gesture is defined as a human body movement that conveys meaning and information to others [1]. A The use of hand gestures is a sensible strategy for creating a user-device interface that is both convenient and flexible. HCI systems can make use of applications like virtual object manipulation, gaming, and gesture recognition. Three key aspects of computer vision are involved in hand tracking: hand segmentation, component detection, and tracking. With hand gestures, people can express themselves more clearly than with any other method of communication. One of the following methods can identify hand gestures: The hand shape ratio in a posture is static, but the hand motion in a gesture might be dynamic. Gesture has been defined by Bobick and Wilson. Gesture, in their view, is any movement of the body meant to communicate with another agent . For a gesture to be effective, both the sender and the recipient must have the same kind of information. There are two types of gestures: dynamic and static. In contrast to static gestures, dynamic gestures are intended to alter over time. Static motions are the focus of this research. Design, robotics, virtual reality, and, most crucially, sign language could all benefit from the automatic recognition of hand motions. The most pressing issue is how to make a computer interpret the gestures made by the user's hands. The position of fingers and the form of hands can vary greatly in hand motions. The non-linearity of hand movements is one of the challenges that must be overcome. It can be done by analysing the photographs' metadata and content. Hand gesture photographs have meta information that can be utilised to identify the gestures. Feature extraction and classification are two separate activities that are combined in the process. An image's features must be retrieved before it can be used to recognise gestures. Then, whichever classification method is most appropriate can be used to sort the data. So, the most difficult part is figuring out how to get at and use these features so that they can be classified. Classification and identification rely on large features. Natural data in its raw form cannot be processed by conventional pattern recognition methods [2]. As a result, extracting characteristics from raw data necessitates a great deal of work and is not automated. Deep learning neural networks, such as CNNs, are able to extract characteristics on the fly and use fully linked layers to classify data [3]. With these two steps together, CNN is able to reduce memory usage and computational complexity while also

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005233 |

improving performance. Moreover, it is capable of deciphering the images' complicated and non-linear associations. Because of this, a CNN-based solution will be employed.
It is hence the focus of this study to construct an algorithm that can evaluate vast volumes of picture data and recognise static hand motions using a CNN.

II. LITERATURE SURVEY

A literature review focused on the current tasks related with this subject matter. A method described by Shweta S. Shinde, Rajesh M. Autee, and Vitthal K. Bhosale [4] uses MATLAB to extract the elements of hand gestures, and then MATLAB built-in commands are utilised to turn the recognised gesture into speech. When it comes to gestures, the Indian hand signal language uses two hands instead of only one, but the American language uses just one hand, so Sangeetha, Valliammai, and Padmavathi [2] came up with a device that uses both hands to generate gestures. Instead of using any external hardware, they use MATLAB to implement their system, which captures the runtime snapshot and then extracts the photo frames and does photo processing with the help of HIS mannequin before extracting the characteristic data via a distance-radically-changing approach. Using this mannequin, the results are found to be satisfactory for the majority of hand signs.
Signal awareness machines based on Harris algorithm for extraction of function extraction have been developed by Anchal Sood and Anju Mishra [3]. The Nx2 matrix matrix is used to store the extracted functions. In a similar fashion, this matrix is applied to the database photo. There are a few roadblocks in the way. When looking at the range pricing for skin segmentation, very light to exceptionally dark-brown ancestry presents an error. But the results are effective.
We, Prashant G. Ahire and Kshitija B. Tilekar, and Tejaswini Pramod and A. Jawake B. Warale's A real-time images was utilised as an input to the [4] device, which then used MATLAB [5] to recognise hand gestures. After analysing the image, they applied a correlation-based technique to map causes, in accordance with the machine's specifications, the model provides an friendly outcome for the gestures by hand. Fortunately, we have dataset that is ready made of it. Future dataset generation can be helpful with this literature.

III.PROPOSED SYSTEM

HGR (Hand Gesture Recognition), Hand movements that are fed into the model during the learning phase are an important part of our approach. HGR is turning hand gestures into words. The goal of this study is to identify meaningful words from hand gestures without verbal communication.

A. *Algorithm Details:*

Here, we will explore the concept of the model, approaches applied in this project and also the design of the system.
There has been tremendous progress in artificial intelligence in closing the gap between human and machine capabilities. Researchers and hobbyists [6] alike are working on a wide range of projects in order to achieve incredible results. One of many such disciplines is the realm of Computer Vision [7].
One of the goals of this field is to give machines a human-like ability to see the world and use that perception for various activities like image and video recognition, image analysis and classification, media recreation and recommendation systems. Over the years, a single method — the Convolutional Neural Network [8] — has been used to develop and improve computer vision with deep learning.

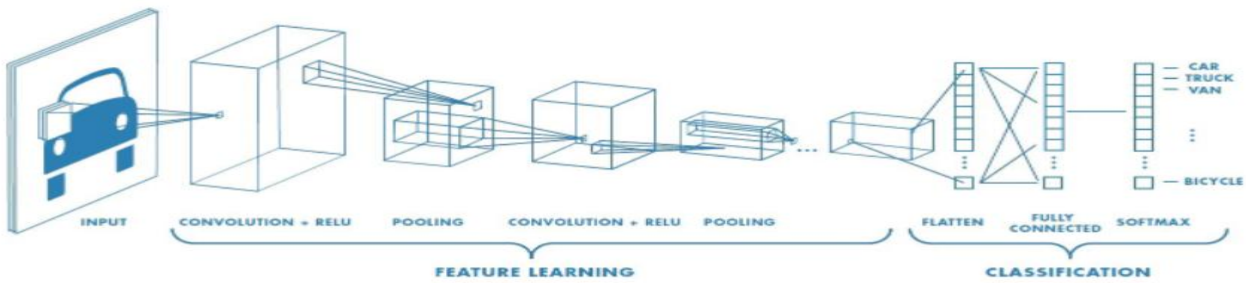


Figure 1:Example of Convolutional Neural Network classification on images

This Deep Learning technique, known as a Convolutional Neural Network (ConvNet/CNN), can take in an image, assign importance (learnable weights and biases) to distinct aspects/objects in the image, and distinguish one from the other. Compared to other classification algorithms, the amount of pre-processing required in a ConvNet is significantly

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005233 |

less. Filters are created by hand in rudimentary approaches, but with enough training, ConvNets can learn to create these filters and attributes themselves.

Architecture of ConvNet [9] is similar to that of human brain connectivity patterns and was inspired by visual cortex arrangement. The Receptive Field is a portion of the visual field in which individual neurons are responsive to stimuli. The entire visual field is covered by a collection of these fields.

Through the use of relevant filters as demonstrated in Figure 1, a ConvNet is able to correctly capture spatial and temporal correlations in an image. Because of the reduced number of parameters and the reusability of weights, the architecture does a better job of fitting the image dataset. To put it another way, the network can be honed to a higher level of sophistication through training

B. *Module Implementation:*

To implement the system, python was used as the programming language and a Ipython notebook was used to write and run code. The library Keras was used for building the CNN classifier. The library PIL was used for image preprocessing. Sklearn was used to calculate the confusion matrix. Matplotlib was used to visualize model accuracy and loss values and confusion matrix. NumPy was used for array operations. The training process on dataset is composed of two phases.

1) Training with Base Dataset: In this phase, the model was trained using the base dataset achieved after pre-processing.

2) Training with Expanded Dataset: In this phase, the dataset was augmented. Data augmentation is a technique to increase the number of data by applying zoom, shear, rotation, flip etc [10].

This process not only increases the data but also brings variation in dataset which is essential for CNN to learn sophisticated differences of images. A random image was selected to provide the demonstration.

We can also view the Hand Gesture Recognition System with simple approach as below Figure 3:

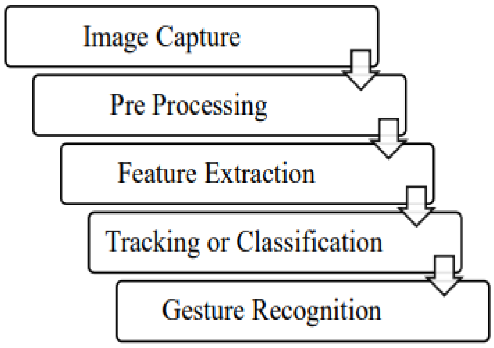


Figure 2: Typical Hand gesture recognition system approach

Image Capture

Image can be the static 2d images of hand gestures. They can also be the frames of the video, as video is a collection of frames one after the another in a sequence.

Pre-processing

This step we have already covered in the before, exactly the same is going to happen in this current flow.

Feature Extraction

This is the crucial part for gaining the knowledge about all the gestures through training. These vary from one person to another as it is not the same for everyone.

Tracking or Classification

In this step, we classify the input into one of the class of the available labels. In this step, it will give you the probability that the input matches with the trained images.

Gesture Recognition

This step going to predict most probable closet gesture of the hand with the input image.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |
|| Volume 10, Issue 5, May 2022 ||
| DOI: 10.15680/IJIRCCE.2022.1005233 |

The complete Hand Gesture System architecture is given below Figure 3 with view to imagine how the model is going to be in virtual.

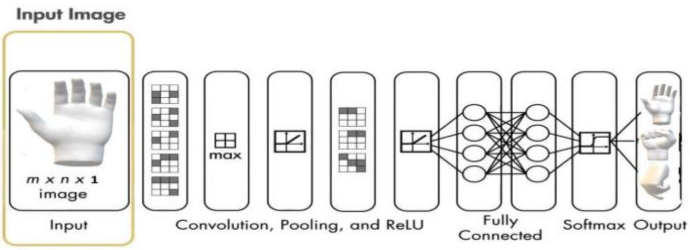


Figure 3: Complete architecture of Hand gesture Recognition System

IV. RESULTS AND DISCUSSION

In this section, we'll examine the model's performance in a variety of scenarios. The results of the tests help to form conclusions and analyse the modifications that need to be made in order to improve good accuracy. Thus, the project's long-term goals may be better defined.

Tests are carried out on a model by feeding it input photos and predicting how the gesture will be displayed on the screen. Python's 'random' predefined module is used to randomly select these test photos from the test dataset. Figures 4 and 5 show the outcomes of the tests. The input photographs are shown in the first row, while the predicted output images are shown in the second row.

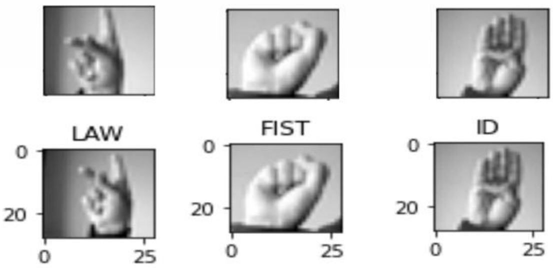


Figure 4:Test case results (1)

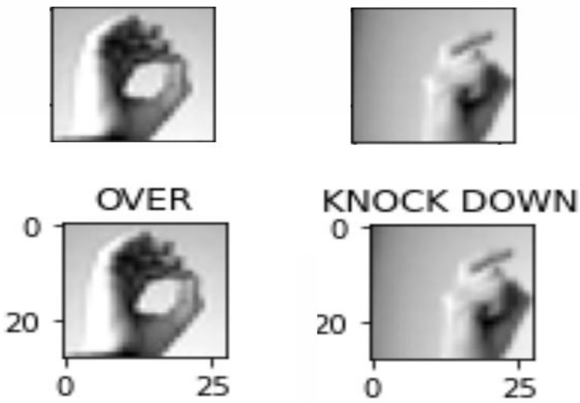


Figure 5: Test case results (2)

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005233 |

The following describes the results obtained from the experiment using the CNN configuration. The experimental result shows that the model has achieved 87% accuracy. The accuracy comparison graphs among the training and testing phases are shown on Figure 6.

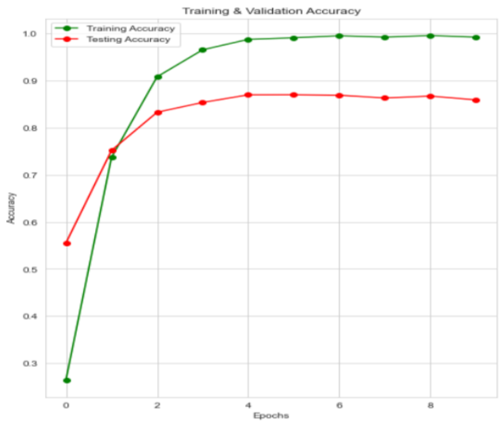


Figure 6: Training and validation accuracy

The graph shows that the model was trained with 10 epochs. The accuracy of the model was getting higher with each epoch up to 10. The loss of data for model was also plotted to corresponding train and test data sets in Figure 7.

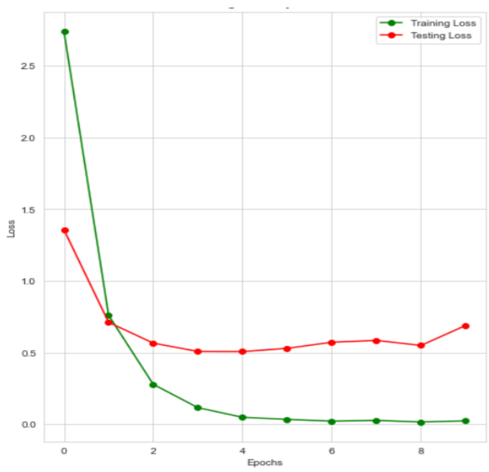


Figure 7: Training and Testing loss

A. *Performance Analysis*

In this section, we'll go over the findings in more detail. Hand gestures can be accurately predicted by a model, but there are some circumstances where the model fails to predict hand gestures accurately. That's because our model's accuracy is 87 percent, which means it's only a little more than 13 percent off. But it does not imply a complete lack of accuracy in the model's depiction. Overfitting occurs if a model achieves 100% accuracy in its predictions. Accuracy will suffer when testing is included. We must keep in mind that the model's output is the likelihood that an input image matches one of the previously learned hand gestures. The model is able to anticipate the input image like a human because of the knowledge we shared at the end of the training step. We must train the model as often as possible with different types of data in order to improve its accuracy and put it to use in real time.

V. CONCLUSION

Hand gesture detection is essential for a genuine Human-Computer Interaction (HCI). Detection, segmentation, and tracking are the most important parts of gesture recognition. In this project, a system for recognising hand motions utilising characteristics extracted and classified using CNN techniques has been developed. We recorded twenty-four

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005233 |

2D hand motions at short distances, each with a different hand position and shape. Experiments were conducted in order to compare the effectiveness of the CNN method's training and testing phases. Training is more accurate than testing, according to the data. The number of gestures will be increased in the future to include more of the more prevalent ones.
This research examines the potential and limitations of recognising hand motions. As a result of this investigation, we can state with confidence that CNN is a data-driven methodology with significant implications for deep learning. It also examines how data affects the model over time. In order to accomplish so, we've done the proper data analysis and pre-processing.
This project demonstrated the process of building a CNN model. We've gone over all of the major CNN types in great depth. Examples of their architectures can also be shown. The model is put to the test by feeding input photos to it, which the model interprets to print out the gesture's intended meaning. This project's premise is straight forward;however, it does necessitate some theoretical research. The architecture model aids you in forming a clear picture of the project. The implementation of our project gives you a clear picture of its progress. We also demonstrated how our project works by providing you with a set of sample test cases and their corresponding test results.

This Hand gesture recognition experiment has been a success, with an accuracy rate of 87 percent. We also strive to improve accuracy by conducting more data training with varied hand motions. There is a lot of potential for this project.

REFERENCES

[1] A. D. Wilson and A. F. Bobick, "Learning visual behavior for gesture analysis," in Proceedings of International Symposium on Computer Vision-ISCV. IEEE, 1995, pp. 229–234.
[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. 521, no. 7553, p. 436, 2015.
[3] Y. LeCun, P. Haffner, L. Bottou, and Y. Bengio, "Object recognition with gradient-based learning," in Shape, contour and grouping in computer vision. Springer, 1999, pp. 319–345.
[4] E. Stergiopoulou and N. Papamarkos, "Hand gesture recognition using a neural network shape fitting technique," Engineering Applications of Artificial Intelligence, vol. 22, no. 8, pp. 1141–1158, 2009.
[5] T.-N. Nguyen, H.-H. Huynh, and J. Meunier, "Static hand gesture recognition using artificial neural network," Journal of Image and Graphics, vol. 1, no. 1, pp. 34–38, 2013.
[6] Q. Chen, N. D. Georganas, and E. M. Petriu, "Hand gesture recognition using haar-like features and a stochastic context-free grammar," IEEE transactions on instrumentation and measurement, vol. 57, no. 8, pp. 1562–1571, 2008.
[7] P. Molchanov, S. Gupta, K. Kim, and J. Kautz, "Hand gesture recognition with 3d convolutional neural networks," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2015, pp. 1–7.
[8] C. J. L. Flores, A. G. Cutipa, and R. L. Enciso, "Application of convolutional neural networks for static hand gestures recognition under different invariant features," in 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON). IEEE, 2017, pp. 1–4.
[9] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in null. IEEE, 2004, pp. 28–31.
[10] Z. Zivkovic and F. Van Der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," Pattern recognition letters, vol. 27, no. 7, pp. 773–780, 2006.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.165

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING