iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: https://doi.org/10.22214/ijraset.2023.57432

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XII Dec 2023- Available at www.ijraset.com

Sign Language Detection

Mohit Titarmare[1], Gaurav Wankar[2], Gaurav Thapliyal[3], Yash Raut[4], Dr. Rahila Sheikh[5]
[1,2,3,4]*Student,* [5]*Guide, Department of Computer Science Engineering, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, Maharashtra, India*

Abstract: ***Sign language is a crucial means of communication for individuals with hearing impairments, serving as a rich and expressive form of language. As the world becomes increasingly interconnected, there is a growing need for technology to facilitate effective communication between sign language users and the broader community. Sign language detection, a branch of computer vision and artificial intelligence, plays a pivotal role in bridging these communication gaps.***
This abstract provides an overview of the key aspects related to sign language detection. The primary focus is on the technological advancements and applications that contribute to the seamless integration of sign language into various aspects of daily life. The paper explores the challenges associated with recognizing and interpreting sign language gestures, considering factors such as variations in signing styles, lighting conditions, and diverse signing communities.
The proposed paper delves into the methodologies employed in sign language detection systems, including image and video processing techniques, deep learning algorithms, and sensor technologies. It highlights the significance of dataset diversity and the role of machine learning in enhancing the accuracy and robustness of sign language detection models. Additionally, the paper discusses the potential integration of sign language detection into mainstream devices, applications, and assistive technologies to empower individuals with hearing impairments.

I. INTRODUCTION

Sign language is a vital and expressive mode of communication for millions of people worldwide who are deaf or hard of hearing. It serves as a unique linguistic system with its grammar, syntax, and vocabulary, allowing individuals to convey thoughts, emotions, and information through visual-gestural means. While sign language is an essential part of the daily lives of many, there exist communication challenges when interacting with individuals who do not understand or use sign language.

In recent years, technological advancements in computer vision and artificial intelligence have paved the way for innovative solutions to address these communication barriers. One such area of focus is the development of sign language detection systems, which aim to recognize and interpret sign language gestures, thereby facilitating more seamless communication between sign language users and the broader community.

The significance of sign language detection becomes particularly evident in a world that is becoming increasingly digital and interconnected. Individuals with hearing impairments face obstacles in various domains, from accessing education and employment opportunities to participating in social interactions. Sign language detection technology holds the promise of breaking down these barriers, offering a bridge between the sign language community and the wider society.

II. RELATED WORK

Characterization of sign language is between two parameter one being manual and other non-manual. The manual parameter consists of motion, location, hand shape, and hand orientation. The non-manual parameter includes facial expression, mouth movements, and motion of the head [2]. Sign language does not include the environment which kinesics does. Few terms are use in the sign language like signing space, which refers to signing taking place in 3D space and close to truck and head. Signs are either one-handed or two-handed. When only the dominant hand is in use to perform the signs they are denoted as one-hand signs else when the non-dominant hand also comes in the phase it is termed as two- handed signs [3]. Sign language when evolved is different from spoken language so the grammar of the sign language is primarily different from spoken language. Inspoken language, the structureof the sentence is one- dimensional; one word followed by another, while in sign language, a simultaneous structure exists with aparallel temporal and spatial configuration. As based on these characteristics, the syntax of sign language sentence is not as strict as in spoken language. Formation of a sign language sentence includes or refers to time, location, person, base. In spoken languages, a letter represents a sound. For deaf nothing comparable exists. Hence the people, who are deaf by birth or became deaf early in their lives, have very limited vocabulary of spoken language and faces great difficulties in reading and writing.

©IJRASET: All Rights are Reserved

III. LITERATURE REVIEW

Khan and Ibraheem researched the essential parts of communication via hand gestures and distinguished the methods that could be helpful to structure sign language vocabulary arrangements for gesture-based communication. The main aim was to report the significance of unaddressed problems, related difficulties, and likely arrangements in the practical implementation of sign language translation.

Mohandes et al. presented a sign language recognition system for Arabic sign language. An effort was made to use a color-based approach where the subject wore colored gloves. A Gaussian skin color model was used to detect the face. The centroid of the face was taken as a reference point to track the movement of hands. The feature set included geometric values such as centroid, angle, and area of the hands. The recognition stage was implemented using the hidden Markov model.

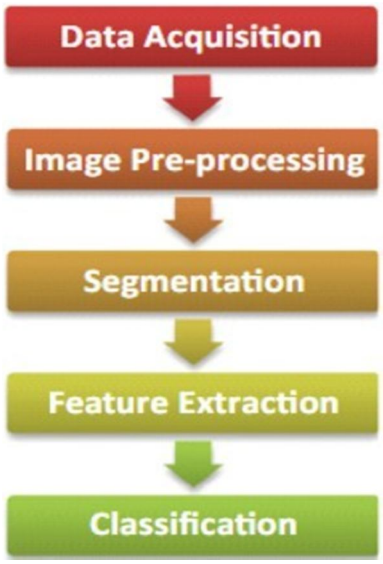
Ross and Govindarajan utilized fusion based on feature level and evaluated on two biometrics systems such as face and hand biometrics system. The data related to the feature level and match level was consolidated. The strategy was examined by combining two types, i.e., intermodal and its fusion scenarios of the classifier such as strong and weak classifiers.

Jiang et al. used RGB and depth image datasets for extricating shape features of input images. The size of the shape features vector was reduced with the application of discriminate analysis, which upgraded the discriminative capacity of the shape features by selecting an adequate number of features. The concepts of multimodal and method of image extraction were thoroughly explained.

Zhu provided a two-phase strategy of feature fusion for bimodal biometrics. During the first phase, linear discriminant analysis was performed to measure the transform features. In the subsequent stage, complex vectors were considered as a transform feature and had the flexibility to add more input compared to the regular fusion method.

The classification of different comparison parameters used in this review. The strategy that we have followed for Systematic Literature Review includes acquisition mode, static/dynamic signs, signing mode, single/double handed signs, techniques used and average accuracy as their parameters. On the basis of these parameters, the review for different sign languages like American Sign Language (ASL), Indian Sign Language (ISL), Arabic Sign Language (ArSL), Chinese Sign Language (CSL), Persian, Brazilian, Greek, Irish, Malaysian, Mexican, Taiwanese, Thai, German, Japanese, South African, Sri Lankan, Auslan, Bangladeshi, Ecuadorian, Ethiopian, Farsi, Italian, Polish, Spanish and Ukrainian Sign Languages have been analyzed and documented respectively.

IV. METHODOLOGY



A. Acquisition of Data (Camera Interfacing)

This is a primary and essential step in sign recognition whole process. Camera interfacing is necessary task to capture images with the help of Webcam. Now a days lots of Laptops are coming with inbuilt camera system so that's helps lot for capturing images to process it further. Gestures can be captured by inbuilt camera to detect hand movements and position. Capturing 30fps will be sufficient to process images; more input images may lead to higher computational time and will make system slow and vulnerable.

B. Image Processing

Image pre-processing contains removing unwanted noise, adjusting brightness and contrast of the image, cropping the image as per requirement [li]. In this process contains image enhancement, segmentation and color filtering process

C. Image Enhancement and Segmentation

As images captured by webcam is RGB images, but RGB images are very much sensitive for various light conditions therefore RGB information convert into YCbCr. Where Y is luma component which denotes luminance information of image, and Cb , Cr are chromo components which give color information of images red difference and blue difference. Luminance component may create problems so only chrominance components get process further. After that YCbCr image converted to binary image.

D. Feature Extraction

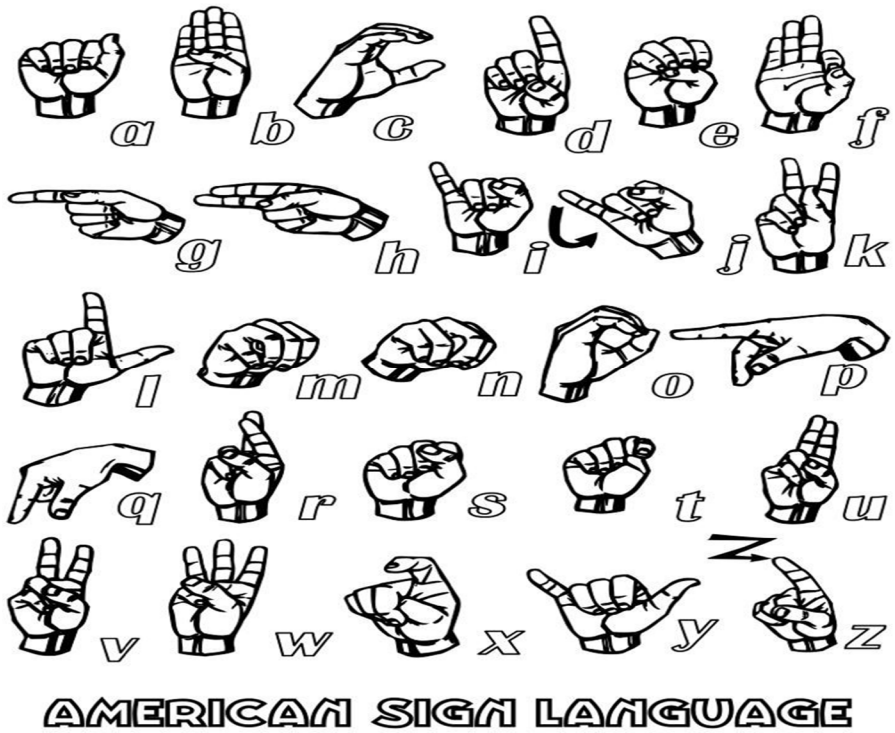
Feature extraction plays a crucial role in sign language detection systems, as it involves capturing relevant information from the input data (such as images or video frames) to represent the distinctive characteristics of sign language gestures. Effective feature extraction contributes to the accuracy and robustness of the model.

E. Classification

Classification in sign language detection involves the process of assigning a specific label or category to a given input, which typically represents a sign language gesture. The goal is to train a model to accurately recognize and classify different signs based on the extracted features.

V. AMERICAN SIGN LANGUAGE (ASL)

American Sign Language (ASL) is the non-verbal way of communication based on English language. Which can be expressed by movements of the hands and face.It is the primary language of many North Americans who are deaf and find difficulties in hearing . It is not a universal sign language. Different sign languages are used in different countries or regions. For example, British Sign Language(BSL) is a different language compared to ASL so the person who knows ASL may not understand BSL.ASL is forth most commonly used Language in US.

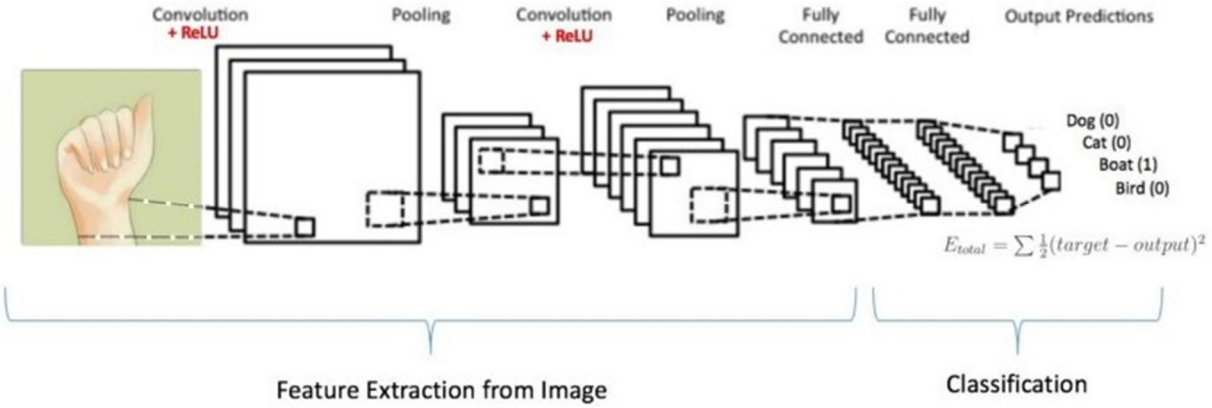


ASL is a language completely segregated and different from English. ASL contains all the significant features of language, with its own rules for pronunciation, word formation, and word order. While every language has ways of indicating different functions, such as asking question instead of making a statement.

VI. ALGORITHM USED

A. *Convolutional Nueral Network*

Convolutional Neural Network is a deep learning technique which is developed from the inspiration of visual cortex which are the fundamental blocks of human vision. It is observed from the research that, the human brain performs a large-scale convolutions to process the visual signals received by eyes, based on this observation CNNs are constructed and observed to be outperforming all the prominent classification techniques .Two major operations performed in CNN are convolution (wT□ X) and pooling(max()) and these blocks are wired in a highly complex fashion to mimic the human brain .The neural network is constructed in layers, where the increase in the number of layers increases the network complexity and is observed to improve the system accuracy .The CNN architecture consists of three operational blocks which are connected as a complex architecture. The functional blocks of Convolutional Neural Network:



1) *Convolutional Layer:* In convolution layer we take a small window size [typically of length 5*5] that extends to the depth of the input matrix. The layer consist of learnable filters of window size. During every iteration we slid the window by stride size [typically 1], and compute the dot product of filter entries and input values at a given position. As we continue this process well create a 2-Dimensional activation matrix that gives the response of that matrix at every spatial position. That is, the network will learn filters that activate when they see some type of visual feature such as an edge of some orientation or a blotch of some color
2) *Max Pooling Layer:* We use pooling layer to decrease the size of activation matrix and ultimately reduce the learnable parameters. There are two type ofpooling :

a) *Max Pooling:* In max pooling we take a window size [for examplewindow of size 2*2], and only take the maximum of 4 values. Well lid this window and continue this process, so well finally get a activation matrix half of its original Size.
b) *Average Pooling:* In average pooling we take average of all Values in a window

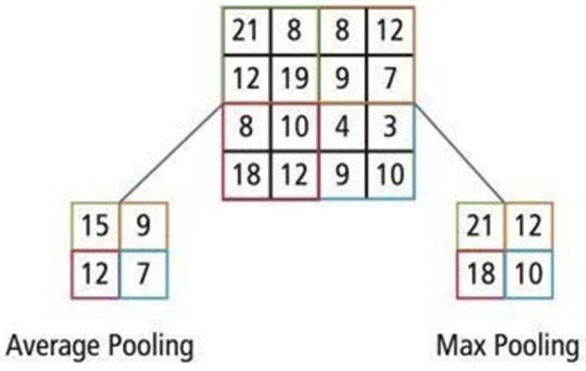


Figure 5.3: Types of pooling

3) *Fully-Connected Layer:* In convolution layer neurons are connected only to a local region, while in a fully connected region, well connect the all the inputs to neurons.

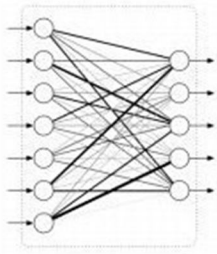


Figure 5.4: Fully Connected Layer

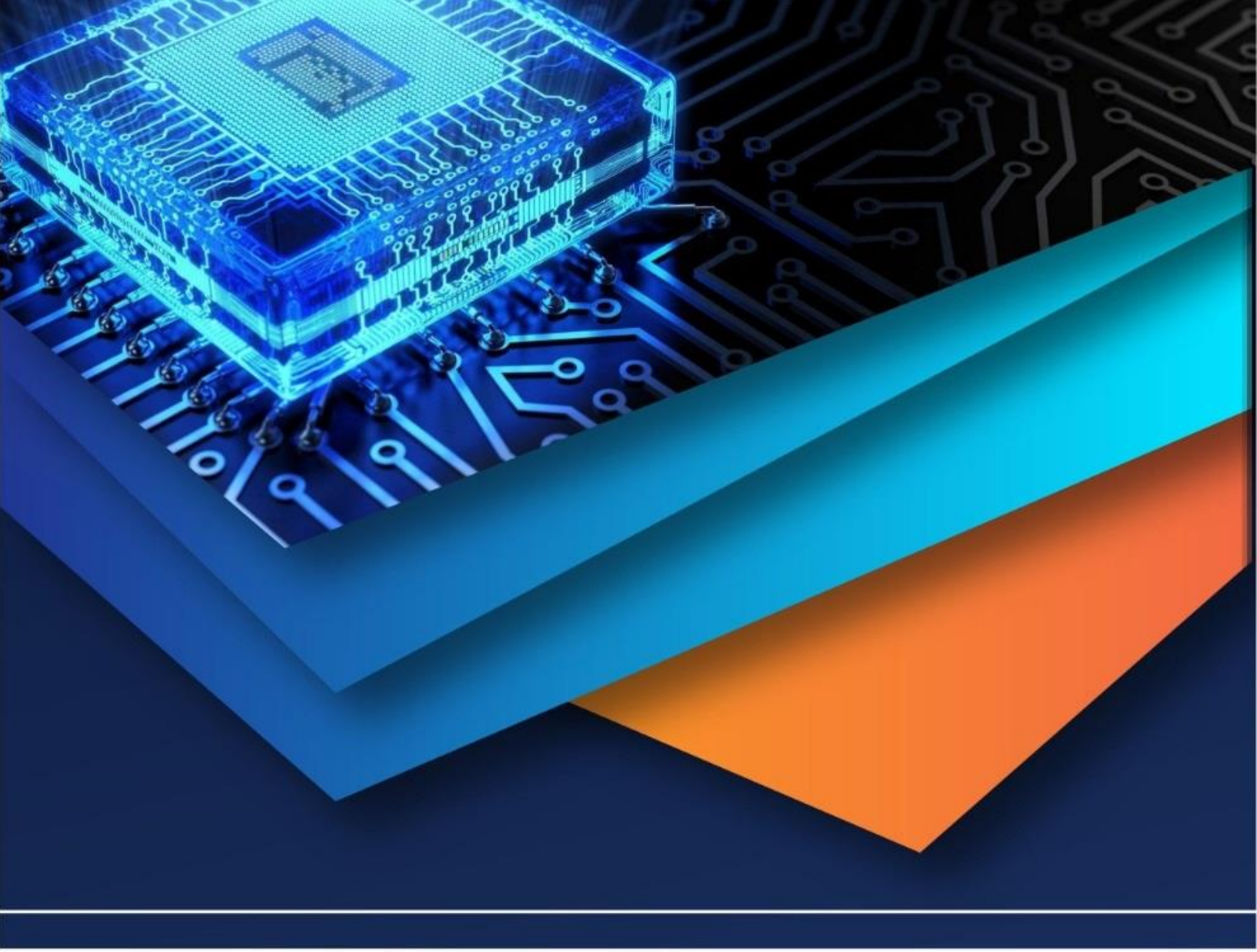
4) *Final Output Layer:* After getting values from fully connected layer, well connect them to final layer of neurons[having count equal to total number of classes], that will predict the probability of each image to be in different classes

VII. CONCLUSION

Sign Language Recognition System has been developed from classifying only static signs and alphabets, to the system that can successfully recognize dynamic movements that comes in continuous sequences of images. Researcher nowadays are paying more attention to make a large vocabulary for sign language recognition systems. Many researchers are developing their Sign Language Recognition System by using small vocabulary and self-made database. Large database build for Sign Language Recognition System is still not available for some of the country that involved in developing Sign Language Recognition System.The neural networks are one of the more powerful tools in the identification system and pattern recognition. The system presents a performance pretty good to identify the static images of the sign alphabetic language. The system shows that the first stage can be useful for deaf persons or with speech disability for communicating with the rest of the people who do not know the language. In this work, the developed hardware architecture is used as image recognizing system but it is not only limited to this application, it means, the design can be employed to process other type of signs. As future work, it is planned to add to the system a learning process for dynamic signs, as well as to prove the existing system with images taken in different position. Several applications can be mention for this method: finding and extracting information about human hands, which can be apply in sign language recognition that it is transcribed to speech or text, robotics, game technology, virtual controllers and remote control in the industry and others

REFERENCES

[1] T. Yang, Y. Xu, and “A. , Hidden Markov Model for Gesture Recognition”, CMU-RI-TR-94 10, Robotics Institute, CarnegieMellon Univ., Pittsburgh, PA, May 1994.
[2] Pujan Ziaie, Thomas M uller , Mary Ellen Foster , and Alois Knoll“A Na ıve Bayes Munich,Dept. of Informatics VI, Robotics and EmbeddedSystems,Boltzmannstr. 3, DE-85748 Garching, Germany
[3] https://docs.opencv.org/2.4/doc/tutorials/imgproc/gausian_median_blur_b ilateral_filter/gausian_median_blur_bilateral_filter.html
[4] Mohammed Waleed Kalous, Machine recognition of Auslan signs using PowerGloves: Towards large-lexicon recognition of sign language.
[5] aeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional Neural-Networks-Part-2/
[6] http://www-i6.informatik.rwth-aachen.de/~dreuw/database.php
[7] Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language Recognition Using Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol 8925. Springer, Cha
[8] Zaki, M.M., Shaheen, S.I.: Sign language recognition using a combination of new vision based features. Pattern Recognition Letters 32(4), 572–577 (201



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)