2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) | 979-8-3503-9737-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICACCS57279.2023.10112890

Creating An Artificial Intelligence-Powered Image Classification Model For Specially Abled Persons

Kandula Srikanth
Department of Information Technology
Seshadri Rao Gudlavalleru Engineering College,
Andhra Pradesh, India.
srikanth.kandula15@gmail.com

Munni Venuthurumilli
Department of Computer Science
PB Siddhartha College of Arts and Science,
Vijayawada, Andhra Pradesh
India.
minni.j2ee@gmail.com

Nune VL Manaswini
Department of Computer Science Nalanda Degree College
Vijayawada, Andhra Pradesh,
India
manaswininvl@gmail.com

M V N N S S R K S Somayajulu
Department of Computer Science and Engineering
Krishna University Machilipatnam,
Andhra Pradesh, India.
m.somayajulu12@gmail.com

Abstract—**This article describes how to use Python's Keras deep learning package to build an Artificial Intelligence-powered image classification model. Speech, text, and other types of input are transformed into machine-readable form by the system using natural language processing (NLP) and speech recognition technology. Other assistive technologies can be adapted to and used with the system. The purpose of this study is to show how straightforward and simple it is to use Keras to create an accurate picture categorization model. Convolutional neural network (CNN) architecture, a kind of deep learning model best suited for image classification tasks, is built first in the procedure. Following that, the model is assembled and trained on a sizable dataset, where it discovers patterns and features in the photos. The model is tested and evaluated on fresh data after training to ensure accuracy. The outcomes demonstrate that Keras can efficiently and accurately construct an AI-powered picture classification model, making it a useful tool for resolving practical computer vision issues. This demonstrates the capabilities of AI in this area and offers a thorough manual for people wishing to create their own picture classification models using Keras.**

Keywords—Python; Flask; Computer Vision (CV2); Tensorflow; Neural Network.

I. INTRODUCTION

The "Creating an AI-Powered Image Classification Model Using Keras in Python" is a piece of technology that aims to improve the communication skills and general quality of life of people with disabilities. This system converts speech, text, and other forms of input into machine-readable format using natural language processing (NLP) and speech recognition technologies, enabling real-time communication and comprehension. Additionally, the system may include a virtual assistant feature that can be accessed through voice commands or texts. With this feature, the system will be able to help with reminders, appointment scheduling, setting alarms, and providing information. Especially abled people may find it simple to manage their day-to-day tasks with the help of this feature. The system may also include an emergency feature that enables the user to send an emergency message to a specific contact [1].

A different form of communication that can be very helpful to those who have mobility or communication impairments is made possible by tailoring the system to the individual needs of the user [2]. This system can be created as a web application or a mobile application, making it readily usable by users on various platforms. Because it works with laptops, smartphones, and other devices, users have the freedom to access the system from any location. It can also be combined with other assistive technology to improve the user's experience. As a result, people with disabilities may be able to easily control their surroundings and access a variety of resources.

Moreover, a real-time communication system can also use advanced AI technology like Computer Vision and Machine Learning to support users. For example, the system can use computer vision to recognize and interpret sign language, or use machine learning to improve its speech recognition capabilities over time. With this feature, the system can help users who have a different way of communicating [3]. Most research implementations for this mission use depth maps generated by depth camera and high-resolution images. The goal of this project is to see whether neurons networks can classify signed ASL letters using simple images of hands captured by a personal device such as laptop webcams.

Connecting with others who share their disabilities could be another advantage of a real-time AI-powered communication system for people with special needs. A social networking element of the system might be available, enabling users to connect and exchange data and resources. For those who are socially isolated or find it challenging to access local support networks and other

979-8-3503-9737-6/23/$31.00 ©2023 IEEE

resources, this may be especially helpful. Users can easily access these services from the comfort of their homes with the help of a feature that enables users to connect with professionals like healthcare workers, therapists, or social workers via secure video conferencing or messaging. Access to resources and support that can enhance their general well-being can be made available to people with special needs as a result [4].

Overall, a real-time communication system powered by AI has the potential to greatly benefit specially-abled individuals and improve their ability to communicate and participate in society. One key feature of this system is that it can be customized to the specific needs of the user. This means that the system can be configured to recognize specific accents or speech patterns or to provide speech output at faster pace. Furthermore, the system could also be integrated with transportation systems to allow for easy communication with public transportation services, such as bus or train schedules, or booking a ride-sharing service, like Uber or Lyft. This can make it easier for specially-abled individuals to get around and access different parts of the community.

II. Literature Review

According to S. S. Kakkoth et al. [5] describes real-time image-based hand gesture recognition. Gestures are still widely used alongside spoken language as a main form of communication. The goal of these gesture recognition systems is to make it possible for people with limited mobility or vision to do common tasks such as communicating basic information or manipulating machines by making use of just their hands. Here, a skin color detection approach is used to hand segmentation utilizing a standard camera and a system that relies on visual descriptors. Here, noise is suppressed by morphological processes. Along with this, we utilize a face recognition classifier that takes into account the presence of hair to get rid of the second greatest skin contour, the hand. The number of fingers is then calculated by forming a convex hull over the resultant shape. Instead of just using the centroid and the furthest distance determined from the centroid, combining the convex hull approach with geometric analysis leads to a higher detection rate of 94%. A geometric comparison between the convex defect spots, centroids, and skin points is used to determine which fingertips belong to which person. Different types of gestures exist depending on the amount of fingers used. In specifically, 10 hand motions were recognized by looking at the tips of people's fingers. Persons with impairments can use these gestures to communicate verbally and in writing.

According to Badiuzzaman Sabuj et al. [6] that millions of people around the world suffer from paralysis like having difficulty in walking. So, a new kind of helpful robot has been introduced. These patients can enhance the quality of their discharge by employing this robot and gesture-based control (glove or wheelchair grip). The suggested robot has two components: a robotic wheelchair and a gesture controller (RW). You may control your robot-based wheelchair with simple hand motions detected by its sensors. When a person and a robot work together, the patient may operate the robot with little effort and go around more independently. The elderly and the disabled can benefit greatly from this system since it can lessen the effort required to regulate RW. Our technology has a 94% success rate with almost no lag time.

According to Ren C. Luo et al [7] Thanks to advancements in technology, robots now play an essential part in society. As more and more robots enter the market, we may expect to see more and more. Intelligent robots for rescue and service are appearing in modern civilization. As a result, investigating how humans and robots can work together becomes crucial. Combining hand signal recognition is discussed in this article. Recognizing hand signals is a crucial step in human-robot interaction (HRI). For those who are unable to speak, sign language provides the most natural and immediate method of communicating. People who have mobility impairments can nevertheless communicate effectively with caretakers and robots by using hand and body movements. In this research, they present a method that integrates two existing recognizers to improve hand-he gesture recognition. Through a process known as CAR expression, these two recognizers collaborate to deduce hand gestures. They hope that these two recognizers will work together to improve the identifying process. This is accomplished through the recognizer's ability to decipher hand movements according to their anatomical structure. Support Vector Machine-based Recognizers (HSR) and Related Systems (SVM). Additionally, local binary patterns (LBPs) and raw data are used by the SVM to train the associated classifier. They also record the trained pictures from the Bosphorus hand database. A rule set that can transition between different types of recognizers In order to synthesize many techniques, a combinatorial methodology was created to recognize CAR expression. They have successfully proven the proof of concept and experimentally tested gesture recognition.

According to Kollipara Sai Varun et al. [8] that hand gestures play a leading role in today's industry. Developing gesture recognition or cognition helps people in many ways. Future approaches, such as gesture control and recognition, will be possible because to the advancement of automation technologies like machine learning, deep learning, neural networks, and computer vision. The usage of these gestures can be adapted to aid those with limited motor skills or other impairments in their ability to operate or communicate with electronic equipment. In this case, the model's evolution makes it easier to understand and use. As a result of these advancements, these models may be used with convolutional neural networks and back propagation to facilitate straightforward operations.

According to Kalpana Lamb et al. [9] there is a rise in the number of impaired patients in societies where the elderly and the sick cannot provide for themselves. Constant monitoring by a nurse may be necessary for some individuals, although this is not always feasible owing to social or financial factors. Currently available electric mattress. Only two of his bed shifts are recorded in the hospital's system (up and down). In order to reduce the burden on nursing staff and maximize the comfort of their disabled patients, they have suggested an automated bed position control system. With this system, they've taken an electronic bed that responds to hand gestures and added support for two more gestures: left and right. The patient's RS-232 communication connection transmits the gesture to the microcontroller when she uses it as input to the system. Accelerometers are employed as sensors to detect and signal patient falls, and a microcontroller interprets these data to automatically adjust the bed using DC motors.

According to M. Ali Qureshi et al. [10] it is an effective hand detection algorithm. When an input image of a hand is applied, the algorithm clarifies and subdivides parts of the hand and performs operations on the fingers of the recognized hand. They verified the performance of this algorithm by applying it to still images as well. The technique they've described can identify a wide variety of hand motions. It's also quite smart and can pick up motions in real time to aid the user.

According to Zhang Qiu-yu et al [11] technique for segmenting hand gestures using YCbCr colour space and K-Means Clustering was presented. The authors of this work compare and contrast the RGB, YCbCr, and HSV colour spaces. With respect to clustering, YCbCr colour space outperforms both RGB and HSV.

According to Chung-Ju Liao et al. [12] proposed that the system mounts and operates a mini DV. The center of the user's collar, allowing the user to show their collar hand in camera field of view move, stand, wear long-sleeves. Using the number of skin-tone pixels, a hand appears in the camera view box then different a technique for extracting a hand image from a complex image. They were motivated to address the issue of hands appearing beyond the frame by the backdrop camera position that had been put in place.

III. Proposed Work

A. System Architecture

This process describes the architecture of a system for real-time hand gesture recognition using AI for assistive technologies for disabled individuals. The system begins with an image dataset being uploaded and pre-processed for training. This pre-processing step includes the translation and testing of the data to ensure its accuracy.

Once the data is pre-processed, the system moves on to the training phase, where a model is created based on the pre-processed data. The training is then evaluated to check for its accuracy. The model created is then used for recognizing hand gestures in real-time.

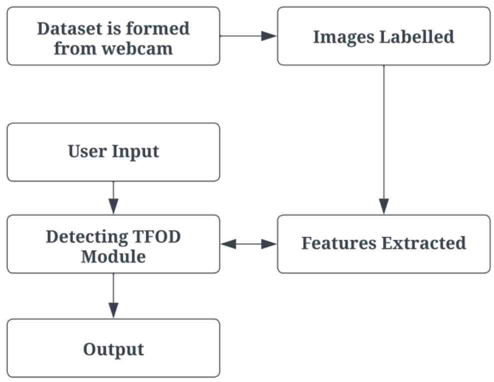

Fig. 1. Proposed System Architecture flowchart

Fig. 1 displays the flowchart of the proposed system architecture. The user provides input through a camera and the input is in the form of sine values, which are fed into the model. The model uses this input to predict the accurate hand gesture, and the predicted output is then displayed through a web application.

The system architecture is designed in such a way that it allows the system to recognize and interpret hand gestures in real-time, which can be used to control assistive technologies such as computer-controlled devices or smart home systems. Additionally, the architecture is designed to be efficient and accurate in recognizing hand gestures, which can greatly enhance the quality of life for those with disabilities.

B. Model Building

This process describes the steps to build a model for image classification using a dataset in Python. It starts by importing the dataset and necessary libraries. The dataset is then pre-processed using an image data generator to convert it into the appropriate format for training and testing. The preprocessed dataset includes 15750 images belonging to 9 classes for training and 2250 images belonging to 9 classes for testing. Each class is assigned an index, such as "a" being 0, "b" being 1, "c" being 2, and so on, to be used in the training and testing phase of the model building process. The model is then trained and tested with the preprocessed dataset, and it will be able to classify images into the respective classes then they process on with model creation.

C. Model Creation

This procedure details how to use the Keras library in Python [13] to build an image classification model. The Sequential class is used to construct a model with the help of layers, including an input layer, a hidden layer(s), and an output layer(s). After the optimizer, loss function, and evaluation metric have been defined, the model may be compiled. The model is then "fit" to the cleaned data using the fit () function. The accuracy of the model is measured and recorded after it has been trained for a set number of iterations (10 in this case). Eventually, you'll want to save the learned model somewhere for use later. The resulting model will have an impressively high prediction and classification accuracy of 0.9994% for pictures.

D. Deployment

This process describes the steps to deploy a trained model to IBM Cloud using the Watson Machine Learning (WML) and TensorFlow libraries in Python. The first step is to install the necessary libraries such as TensorFlow and Watson Machine Learning into the machine learning (ML) package.

Once the libraries are installed, the model is deployed to IBM Cloud using an API key for IBM Cloud Object Storage. This key is used to authenticate the connection and upload the data set files, which are saved in a streaming body format. The files are then extracted and shifted for pre-processing. The model is then compressed and saved in a compressed format.

The saved model is then deployed in the IBM Cloud using the Watson Machine Learning service. The deployment requires specifying a deployment space and generating an API key. The model is then saved in the TensorFlow specification format, and the Tensor ID is specified along with the deployment model ID. This allows the model to be accessed and used for predictions via an API endpoint.

Finally, the deployed model is evaluated for its accuracy and compared with the original model. The accuracy of the deployed model can be different from the original model due to the compression, but it still can be considered as a good trade-off between the performance and the size of the model. The compressed model will have smaller size, making it easier to deploy and use in production.

E. Keras

The algorithm used in this work is a Convolutional Neural Network (CNN). Convolutional Neural Networks are a type of deep learning algorithm [14] that is used for image classification tasks. It uses a series of convolutional and pooling layers to learn features from the input images and make predictions based on those features. In this work, the Keras library in Python was used to create the CNN model. The model was created using the Sequential class, and layers such as an input layer, one or more hidden layers, and an output layer were added. The model was then trained on a preprocessed dataset of images, and the accuracy of the model was checked and recorded. The final model was able to classify images with high accuracy and can be used for prediction [15].

F. Flask Application

This process describes the steps to create a web application to display the output of a hand gesture recognition model, using the Flask library in Python. Flask is a micro web framework that allows for the creation of web applications. The process begins by importing the necessary libraries and functions to create the application. The application also includes login and registration pages for user authentication.

Once the libraries are imported, the local system camera is accessed using the OpenCV library (CV2) to capture user input in the form of hand gestures. The previously created and trained model is loaded and used to predict the corresponding alphabet. Fig. 2 displays the working model of the trained model.

The code is then created to display the output on the web application. The server is started by running the command "python filename.py" which will display the HTTP address. This address is copied and pasted into the browser which will display the web page. The web page displays details of this work, and prompts the user to open the camera for authentication. After the user enters their username and password, the camera option is enabled. The user can then give input using hand gestures, captured by the local system camera, and the model will predict the corresponding alphabet in real-time. If the user changes the symbol of their hand gesture, the model will predict quickly and accurately, which indicates that the model is well-trained. There is also an option to view the American Sign Language (ASL) alphabets for reference.

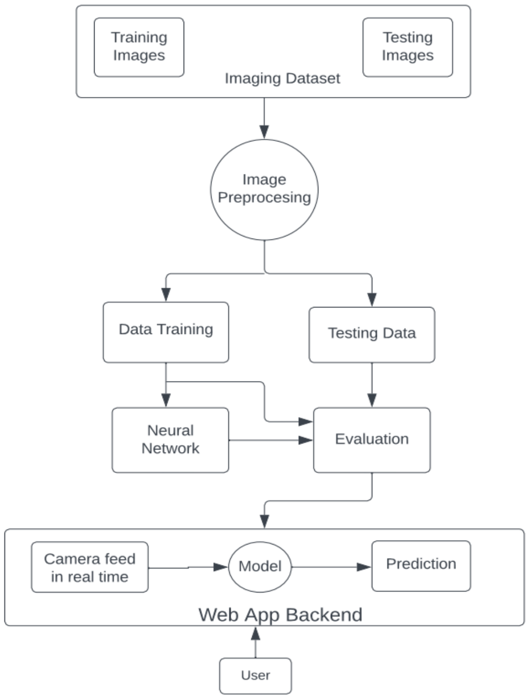

Fig. 2. Working of Model in Flowchart.

IV. Results

The hardware specifications Intel(R) Core (TM) i3-7020U CPU@ 2.30GHz, 8 GB RAM and 64-bit OS, x64-based processor. The software used for performing the code is VS Code and python was the technology used in this project which imports the libraries like cv2, numpy, keras, flask applications.

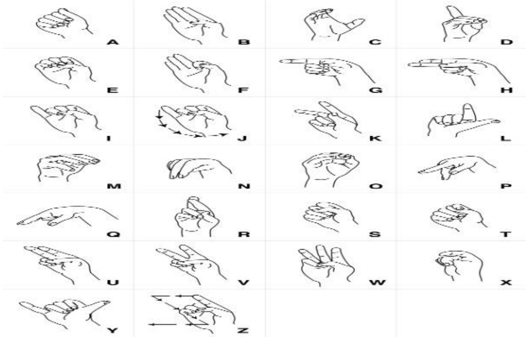

Fig. 3. ASL hand gestures.

Whenever the user displays a hand gesture the trained model will predict the alphabet correctly and displays us the output. Fig. 3 displays the ASL hand gestures symbols. This shows that the model is trained perfectly and helps us to communicate with the especially abled persons.

```
Epoch 10/10
31/31 [==============================] - 76s 2s/step - loss: 0.0024 - accuracy: 0.9991
```

Fig. 4. Output of the model.

Fig. 4 shows the trained model accuracy, loss, and delay time for the trained model. The accuracy value is 0.9991, which is considered to be high. This means that the model has a high ability to correctly classify images into the respective classes. The accuracy value was calculated after training the model for 10 epochs and testing it on a preprocessed dataset that includes 2250 images.

TABLE I. COMPARISON BETWEEN EXISTING METHODS AND OUR METHOD

	Accuracy
Proposed Work	0.9991
Existing Work [6]	0.9350
Existing Work [9]	0.9600

Comparing our method with the existing methods, our proposed method has a accuracy of 0.9991 which is more efficient and provides better accuracy than the existing methods as shown in the above Table I. Fig. 5 displays the accuracy comparison of our proposed method and existing methods in bar chart.

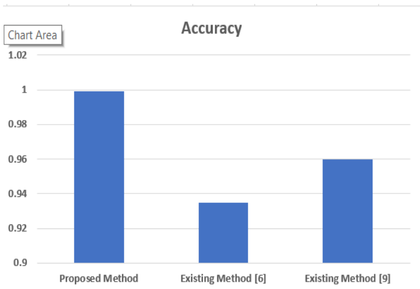

Fig. 5 Comparison in Bar Chart

Fig. 6. The model predicts the gestures as “F”.

Fig. 7. The model predicts the gestures as “C”.

Fig. 6 and Fig. 7 displays the predicted alphabet by the trained model. This project aims to improve communication for people who are deaf or hard of hearing by developing a web application that uses neural networks to classify letters in American Sign Language (ASL) by use simple hand images taken with a personal device such as a laptop webcam. This technology will allow people who are not fluent in sign language to communicate more easily with people who are deaf or hard of hearing. Additionally, learning sign language has been shown to have cognitive benefits such as improved attention span and creativity.

V. CONCLUSION

This research shows that the system that is designed to facilitate communication between deaf and hearing individuals by using sign language. The system uses a neural network to recognize hand gestures and translate them into English alphabets that can be understood by humans. This allows deaf individuals to communicate with others by making hand gestures, which are then recognized by the system and translated into text on a screen. The goal of the system is to bridge the communication gap between deaf individuals and the rest of society by providing a two-way communication channel. Having a technology that can translate hand sign language to its corresponding alphabet is a game changer in the field of communication and AI for the especially abled people such as deaf and dumb. With introduction of gesture recognition, the web app can easily be expanded to recognize letters beyond 'I', digits and other symbols plus gesture recognition can also allow controlling of software/hardware interfaces.

REFERENCES

[1] Seo Yul Kim et al., “A Hand Gesture Recognition Sensor Using Reflected Impulses,” IEEE Sensors Journal, vol. 17, May 2017, pp. 2975-2976.

[2] Siddharth S. Rautaray Anupam Agrawal, “Vision based hand gesture recognition for human computer interaction: a survey,” Springer Science Business Media Dordrecht 2012.

[3] Hong Cheng, Lu Yang, Member, and Zicheng Liu, “Survey on 3D Hand Gesture Recognition,” IEEE transactions on circuits and systems for video technology, vol. 26, no. 9, sept 2016.

[4] A. Barkoky and N.M. Charkari, “Static Hand Gesture Recognition of Persian Sign Numbers using Thinning Method,” ICMT 2011

[5] S. S. Kakkoth and S. Gharge, "Real Time Hand Gesture Recognition & its Applications in Assistive Technologies for Disabled," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-6.

[6] B. Sabuj, M. J. Islam and M. A. Rahaman, "Human Robot Interaction Using Sensor Based Hand Gestures For Assisting Disable People," 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 2019, pp. 1-5.

[7] R. C. Luo and Y. -. C. Wu, "Hand gesture recognition for Human-Robot Interaction for service robot," 2012 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Hamburg, Germany, 2012, pp. 318-323.

[8] K. S. Varun, I. Puneeth and T. P. Jacob, "Hand Gesture Recognition and Implementation for Disables using CNN’S," 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2019, pp. 0592-0595.

[9] K. Lamb and S. Madhe, "Automatic bed position control based on hand gesture recognition for disabled patients," 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), Pune, India, 2016, pp. 148-153.

[10] M. A. Qureshi, A. Aziz, M. A. Saeed, M. Hayat and J. S. Rasool, "Implementation of an efficient algorithm for Human hand gesture identification," 2011 Saudi International Electronics, Communications and Photonics Conference (SIECPC), Riyadh, Saudi Arabia, 2011, pp. 1-5.

[11] Zhang Qiu-yu, Lu Jun-chi, Zhang Mo-yi, Duan Hong-xiang and Lv Lu, “Hand Gesture Segmentation Method Based on YCbCr Color Space and K-Means Clustering,” International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 8, No. 5 (2015), pp. 105-116.

[12] Chung-Ju Liao, Shun-Feng Su, Ming-Chang Chen, “Vision-based Hand Gesture Recognition System for a Dynamic and Complicated Environment,” IEEE Int. Conf. Syst, Man, & Cybernetics, 2015.

[13] S. Navaneethan, P. Siva Satya Sreedhar, S. Padmakala and C. Senthilkumar, "The human eye pupil detection system using bat optimized deep learning architecture," Computer Systems Science and Engineering, vol. 46, no.1, pp. 125–135, 2023. 7.

[14] Jyoti A. Dhanke, Rajesh Kumar Maurya, S. Navaneethan, Dinesh Mavaluru, Shibili Nuhmani, Nilamadhab Mishra, Ellappan Venugopal, "Recurrent Neural Model to Analyze the Effect of Physical Training and Treatment in Relation to Sports Injuries", Computational Intelligence and Neuroscience, vol. 2022, Article ID 1359714, 9 pages, 2022.

[15] N Nandhagopal, S Navaneethan, V Nivedita, A Parimala and D Valluru, "Human Eye Pupil Detection System for Different IRIS Database Images", Journal of Computational and Theoretical Nanoscience, vol. 18, no. 4, pp. 1239-1242, 2021.