EAST DELTA UNIVERSITY

Project Report on

"American Sign Language Recognition Using Image Processing"

Project carried out

By

Name	ID
Mahzuzah Afrin	201000612
Pranta Chakraborty	201003412
Rajasree Das Chaiti	201007712
Sumaia Farhana	201012512

Submitted to

Muhammad Asif Hasan

Lecturer

School of Science, Engineering & Technology

Table of Contents

Table of Figures

I. Abstract

There are many people in this world who can't hear and speak properly. Sign language is created for them to communicate with others. In this manuscript, we discuss about Recognition of American Sign Language using machine learning algorithms. In the computer vision field, sign language images are classified in many ways. In this paper, we discuss about classifying sign language images using many algorithms with the dataset which have 36 different classes and was collected from. The dataset provides a set of 26 sign alphabets and 10 numeric gestures (0 to 9). Some preprocessing methods, feature extraction with Hog feature is applied before model training and evaluation. To get the accuracy we applied SVM, KNN and Random Forest and Decision Tree. We get the best evaluation by using SVM which is 96%. In our proposed paper, we also discuss about related works, weakness and future work.

.

II. Introduction

People who are hearing impaired cannot properly hold social interaction and regular communication with other people in the society. It is shown that they often encounter various problems in communicating with other people through gestures because majority of people can recognize only a small part of them. As they are not capable of communicating verbally like hearing people, deaf persons and those who have hearing impairment must always rely on some form of visual conversation.

Sign language is a fundamental way of communication for individuals who have hearing loss since it serves as a connection between the community of hearing impaired and the rest of the society. Not only the deaf community in North America, but also other 20 countries in which English is the common communication language benefits greatly from the use of American Sign Language (ASL), which makes communication among the community and out of the community easier. In today's world of advanced technology, using image processing and various machine learning algorithms to improve the working method or functionality of sign language recognition system is getting popular day by day. There are hand gestures that call for one hand or both hands created

for different languages using various hand positions and movements. For American Sign Language, there are 36 frequently used symbols where 10 are digits and 26 are English alphabets . Over the past few years, people have contributed in taking necessary steps in making American Sign Language more accessible for the deaf community. In 2021, a paper was done on American sign language recognition using Support Vector Machine and CNN (SVM) [2]. Another paper also proposed a CNN based sign language recognition system in 2021 where the creators applied several image processing techniques to their dataset [3].

This paper explores a American Sign Language dataset which have 36 different classes. The proposed paper consists of three steps – applying image processing steps to the dataset, feature extraction of the hand signs and applying machine learning algorithms. The dataset is analyzed using a combination of image processing methods and machine learning algorithms. Image processing methods include grayscale conversion, thresholding, skin masking and canny edge detection. The ML algorithms implemented for classification are Support Vector Machines (SVM), Random Forest, Decision Tree and k-Nearest Neighbors (k-NN) and Decision Tree. In summary, as a result of this study significance progress has been made in the field of American Sign Language using Image Processing.

III. Related Works

1. In 2020, Salma A. Essam El-Din and Mohamed A. Abd El-Ghany proposed a system that has 88% accuracy of reading dynamic signs and 95% accuracy of reading static signs. We all know that sign language is essential for those who have difficulty of listening or speaking. Since not everyone is familiar with Sign language so this will cause a gap between people who know and who don't. The researcher has been created a glove to bridge this gap. The glove links to a device on the arm and the fingers of the glove include sensors. Machine learning is used to interpret the signs better also the solution is user-friendly and reasonably priced. People with speech and hearing problems may communicate easily for this innovation.

2. In September 2020, Jayshree Maloo, Aishwariya Ramesh et.al [] proposed a study that examines many processes which translates sign language into spoken or written speech. After that the researcher of this research paper settled on the best one. They developed an Android

app that converts the sign languages into text or voice. To put it simply they worked up how to create an app that transforms ASL into words that can be heard or read.

3. In 14 October, 2022 Surya Narayan S, Surendar K et.al [] developed a model that uses pictures and a computer learning instead of fancy sensors. Computers are used in this study to communicate with people that has hearing and speaking difficulties by interpreting the sign languages. They are concentrating the whole ASL languages that consists numbers and letters. While some sign are moving others are stationary. By watching the movements of hands and fingers they are instructing the computer to make distinguish between the movements. In this way they are facilitating communication.

4. In 2022, Roshnee Matlani, Roshan Dadlani, Shurti Mishra et.al [] wants to utilize a method that will help people has trouble hearing and speaking to assist them, that's why they employing a unique Machine Learning (ML) technique which train computer how to understand motion and interpret it into spoken words or handwritten materials. Additionally, to be created from spoken words or written text they are making feasible for movements. Deaf people can hear more effectively in this approach. Their research is a sort of link between many forms of communication.

5. People who are hearing impaired or hard of hearing technology have advanced to assist them. They can't access voice assistants like Siri or Google Assistant. To solve these problems in November 2021, Sadaf Ikram and Namrata Dhanda developed a method using some modern methodology like Convolutional Neural Networks (CNN), Computer Vision and Deep Learning (DL). To recognize hand signals they created an app with accurate findings. They taught the software to recognize hand moments using camera images. The software is capable of real-time sign detection even it is using a camera. To comprehend other hand motions, phrases, or even sentences this research might be developed further in the future.

IV. Methodology

Many pre-processing techniques and feature extraction have been used to generate the best level of accuracy for our proposed work. Our developed system consists of several methods such as data pre-processing, extracting feature from the images, model training and evaluation.

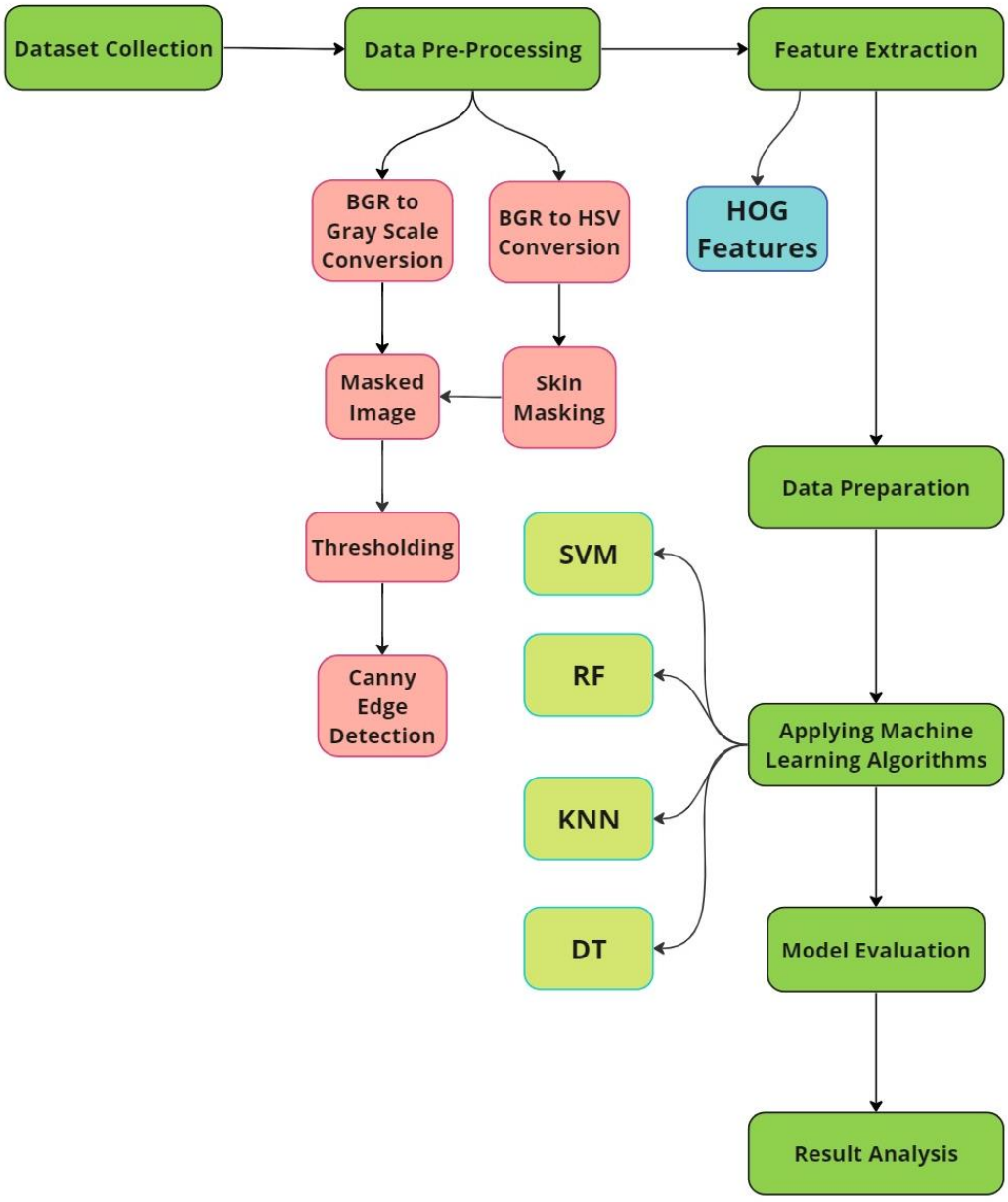

Figure 1: Flowchart of methodology

- **Datset Collection:**

  The dataset was collected from ‘kaggle’[3]. This dataset consists of different types of colored hand images including alphabets and digits which are called American Hand Sign Language. The dataset contains 36 classes where 10 classes are related to the digits and 26 classes from them are English alphabets. There are 2515 images on this dataset.

Figure 2: American Sign Alphabets

- **Data Pre-Processing:**

  Several pre- processing methods were applied to analyze the images to get the best result. Such as-

I. **BGR to Gray Scale Conversion:** We transformed the BGR color images into gray scale to analyze the images easily, also for the feature extraction. While retaining necessary information similar to hand gestures the transformation makes the data simple. cv2.cvtColor function have been used for this conversion.

Figure 3: Output of Grayscale Conversion

II. **BGR to HSV conversion:** The HSV is a color space that stands for Hue, Saturation, and Value. This BGR to HSV conversion was applied to our images to separate hue and saturation that is the color information from value which is the brightness information of the images to facilitate further image color analysis for skin masking.

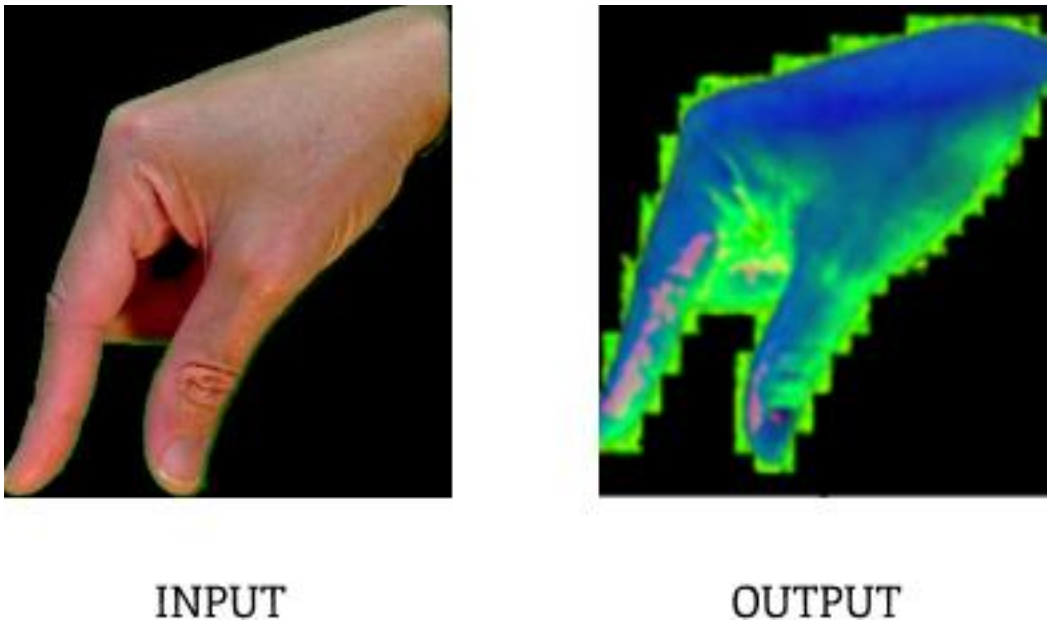

Figure 4: Output of HSV conversion

III. **Skin Masking:** The HSV conversion generates a mask for separating regions of skin color. Regions of skin color can be separated by using upper skin and lower skin arrays. The arrays are threshold to the HSV color model. Then the following processes after this step are creating a gray mask and applying the mask for further analysis.

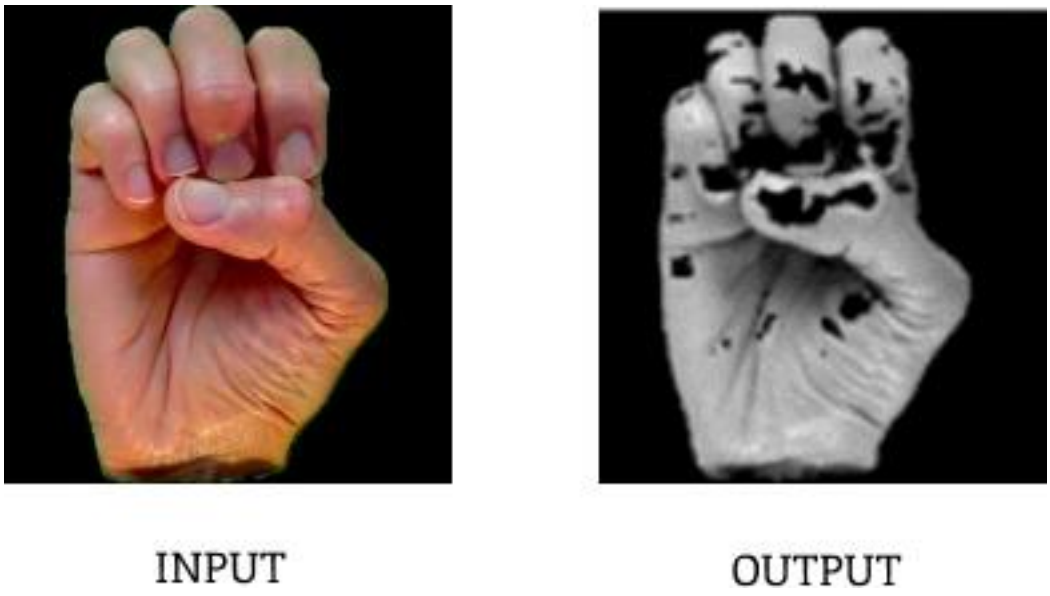

Figure 5: Output of Skin Masking

IV. **Thresholding:** Thresholding is a method to create binary image from grayscale image where pixels that are greater than the threshold (skin region) are white and pixels that are less than the threshold are black. This method is used to makes it easier to for applying further processes like canny edge detection and feature extraction to the images.

Figure 6: Output of Binary Thresholding

V. **Canny Edge Detection:** To detect and identify the boundaries of objects in images canny edge detection is a useful technique. The process of canny edge detection was concluded by using the cv2.Canny function. The threshold binary images that we got from the resultant pre-processing steps being used in this process.

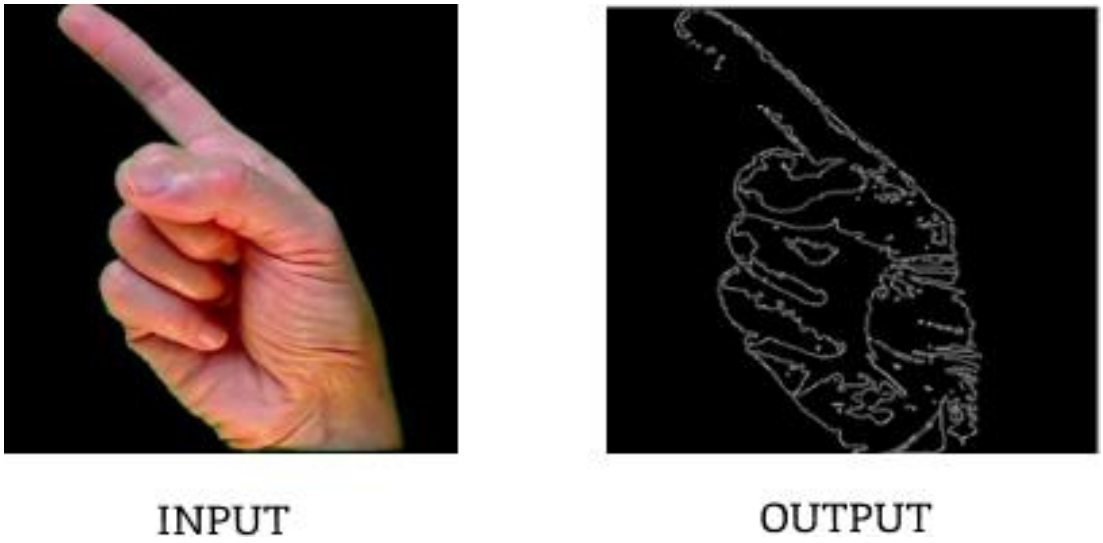

Figure 7: Output of Canny edge detection

- **Feature Extraction:** While storing the necessary information in the actual dataset, feature extraction covert the raw characteristics data into numerical.

  HOG Feature: HOG stands for Histogram of Oriented Gradient. In this process, the grayscale images of the dataset that have gone through other neccessary pre-processsing steps are used as input for extraction relevant features. Gradient information from each cell and gradient from each block per cell is computed by HOG. All the featured vectors, resulted from HOG were then used as inputs for evaluating the dataset using machine learning algorithms.

- **Dataset Preparation:** The actual dataset was prepared to feed it to the machine learning algorithms. It contains various steps such as-
  a. Loading the dataset which read each image from the dataset directory for applying the preprocessing steps and HOG feature extraction to the images.
  b. Assigning labels or classes to them.
  c. Splitting the images into train and test images using sklearn's "train_test_split" function.
  d. Converting the images to NumPy array.
  e. Encoding the labels which are the process of converting the categorical labels to numerical values.

- **Apply Machine Learning Algorithms:** Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest, and Decision Tree classifiers are the machine learning algorithms used in our model to correctly recognize the American Sign Language. These classifiers applied on the pre-processed images (HOG feature extracted) to make predications and learn the pattern of the data. We train the classifier using ‘sklearn’ library for example: SVM from sklearn.svm, RF from sklearn.ensemble, KNN from sklearn.neighbors and DT from sklearn.tree. We predicted the test data using (test_data) after training all the classifiers. After that, store the predicted value for each classifier by using different variables.

- **Result analysis:** The support vector machine algorithm achieved an accuracy of 96%. 88% accuracy was achieved after evaluating the model with k-nearest neighbor. The random forest algorithm achieved an accuracy of 89%. The test accuracy of the model after evaluating it with decision tree is 48%. In comparison with others models Support Vector Machine (SVM) performs well.

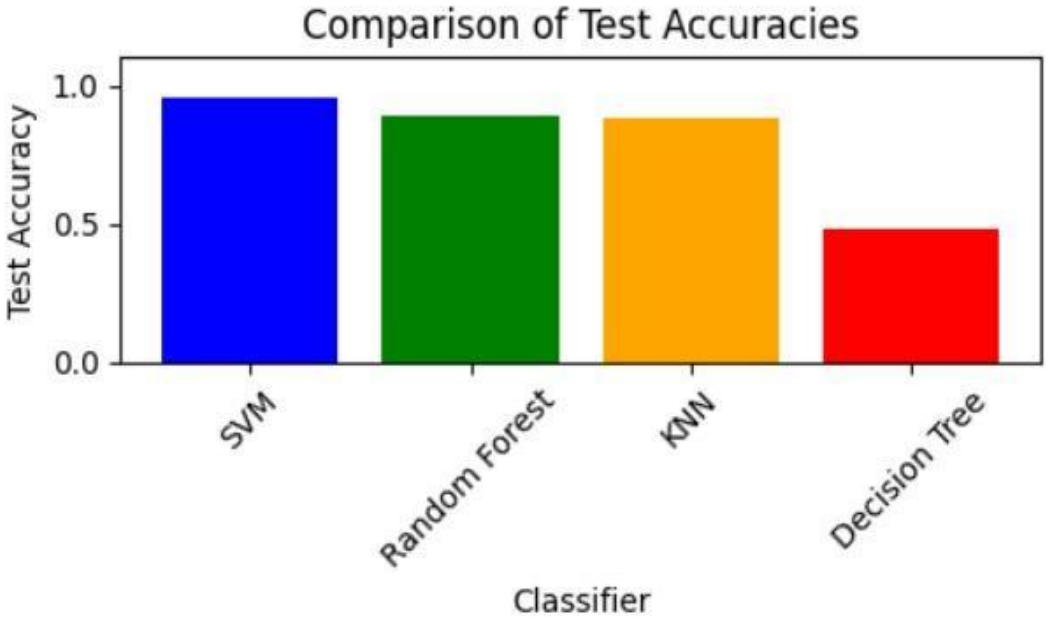

Figure 8: Comparison of test accuracies

V. Conclusion

Sign language is one of the most important and widely used forms of communication for persons who are deaf and dumb (D&D). Communication with the broader public is challenging for the D&D community. The American Sign Language (ASL) users are also faces the same challenges. We work on a dataset which has 36 classes where 26 of them are ASL alphabets, 10 of them are digits and a total of 2515 distinct hand signals. For pre-processing the dataset, we applied some image processing techniques such as- BGR to gray scale conversion, BGR to HSV conversion, skin masking, thresholding and canny edge detection. To detect the edge canny edge detection is also applied in this research. In order to visually show how pictures are transformed into their HOG representations, hog visualization techniques are utilized. These approaches can be helpful in learning how HOG functions and its possible applications. To extract the features from our dataset we applied HOG features extraction techniques. After that we applied four algorithms from machine learning which are SVM, Random Forest, KNN and Decision Tree. According to result of this algorithm we can say that Support Vector Machine gives the most promising result among the others algorithms with an accuracy of 96%. We plot a bar chart to show how these algorithms predict the classification correctly or incorrectly. With the help of the figure, we can see that Decision Tree have largely classified occurrences erroneously, yielding lower accuracy ratings when compared to SVM, KNN and Random Forest. There are also some limitation we faced while developing the model such as the canny edge detection detect the edges of the object very slightly not properly, also the dataset is fairly small. Working on the large dataset will improve our resultant output.

VI. Reference:

1. V. Jain, A. Jain, A. Chauhan, S. S. Kotla, and A. Gautam, "American Sign Language recognition using Support Vector Machine and Convolutional Neural Network," *Int. J. Inf. Technol.*, vol. 13, no. 3, pp. 1193–1200, 2021.
2. J. J. Raval and R. Gajjar, "Real-time sign language recognition using computer vision," in *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*, 2021.
3. https://www.kaggle.com/datasets/ayuraj/asl-dataset
4. El-Din, Salma A. Essam, and Mohamed A. Abd El-Ghany. "Sign Language Interpreter System: An Alternative System for Machine Learning." In 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), 332–37. Giza, Egypt: IEEE, 2020. https://doi.org/10.1109/NILES50944.2020.9257958.
5. Ikram, Sadaf, and Namrata Dhanda. "American Sign Language Recognition Using Convolutional Neural Network." In 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), 1–12. Kuala Lumpur, Malaysia: IEEE, 2021. https://doi.org/10.1109/GUCON50781.2021.9573782.
6. J, Manikandan, Brahmadesam Viswanathan Krishna, Surya Narayan S, and Surendar K. "Sign Language Recognition Using Machine Learning." In 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 1–5. Chennai, India: IEEE, 2022. https://doi.org/10.1109/ICSES55317.2022.9914155.
7. Matlani, Roshnee, Roshan Dadlani, Sharv Dumbre, Shruti Mishra, and Mrs. Abha Tewari. "Real-Time Sign Language Recognition Using Machine Learning and Neural Network." In 2022 International Conference on Electronics and Renewable Systems (ICEARS), 1381–86. Tuticorin, India: IEEE, 2022. https://doi.org/10.1109/ICEARS53579.2022.9752213.
8. Tiku, Kohsheen, Jayshree Maloo, Aishwarya Ramesh, and Indra R. "Real-Time Conversion of Sign Language to Text and Speech." In 2020 Second International

Conference on Inventive Research in Computing Applications (ICIRCA), 346–51. Coimbatore, India: IEEE, 2020. https://doi.org/10.1109/ICIRCA48905.2020.9182877.