Visvesvaraya Technological University

Jnana Sangama, Belagavi - 590018

A Project Work Phase-I (17CSP78)

Report on

"ALPHANUMERIC CHARACTER RECOGNITION IN AUDIO/TEXT BASED CAPTCHA"

Project Report submitted in partial fulfilment of the requirement for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

AAFREEN HUSSAIN	**1KS17CS001**
AKSHITHA B.S	**1KS17CS004**
ANOOP P S	**1KS17CS008**

Under the guidance of

Mrs. Soungandhika Narayan
Assistant Professor
Department of Computer Science & Engineering
K.S.I.T, Bengaluru-560109

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

K. S. Institute of Technology

#14, Raghuvanahalli, Kanakapura Road, Bengaluru - 560109

2020 - 2021

K. S. Institute of Technology

#14, Raghuvanahalli, Kanakapura Road, Bengaluru - 560109

Department of Computer Science & Engineering

CERTIFICATE

Certified that the Project Work Phase-I (17CSP78) entitled **"Alphanumeric Character Recognition in Audio/Text Based Captcha"** is a bonafide work carried out by:

AAFREEN HUSSAIN	**1KS17CS001**
AKSHITHA B.S	**1KS17CS004**
ANOOP P S	**1KS17CS008**

in partial fulfilment for VII semester B.E., Project Work in the branch of Computer Science and Engineering prescribed by **Visvesvaraya Technological University, Belagavi** during the period of September 2020 to January 2021. It is certified that all the corrections and suggestions indicated for internal assessment have been incorporated. The Project Work Phase-I Report has been approved as it satisfies the academic requirements in report of project work prescribed for the Bachelor of Engineering degree.

……	……..……	……..……
Signature of the Guide	**Signature of the HOD**	**Signature of the Principal & CEO**
[Mrs.Sougandhika Narayan]	[Dr. Rekha B. Venkatapur]	[Dr. K.V.A. Balaji]

DECLARATION

We, the undersigned students of 7th semester, Computer Science & Engineering, KSIT, declare that our Project Work Phase-I entitled "**Alphanumeric Character Recognition in Audio/Text Based Captcha**", is a bonafide work of ours. Our project is neither a copy nor by means a modification of any other engineering project.

We also declare that this project was not entitled for submission to any other university in the past and shall remain the only submission made and will not be submitted by us to any other university in the future.

Place: Bangalore
Date:

Name and USN	**Signature**
AAFREEN HUSSAIN(1KS17CS001)	**…….**
AKSHITHA B.S(1KS17CS004)	**…….**
ANOOP P S(1KS17CS008)	**…….**

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task will be incomplete without the mention of the individuals, we are greatly indebted to, who through guidance and providing facilities have served as a beacon of light and crowned our efforts with success.

First and foremost, our sincere prayer goes to almighty, whose grace made us realize our objective and conceive this project. We take pleasure in expressing our profound sense of gratitude to our parents for helping us complete our Project Work Phase-I successfully.

We take this opportunity to express our sincere gratitude to our college **K.S. Institute of Technology,** Bengaluru for providing the environment to work on our project.

We would like to express our gratitude to our **MANAGEMENT,** K.S. Institute of Technology, Bengaluru, for providing a very good infrastructure and all the kindness forwarded to us in carrying out this project work in college.

We would like to express our gratitude to **Dr. K.V.A Balaji, Principal & CEO**, K.S. Institute of Technology, Bengaluru, for his valuable guidance.

We like to extend our gratitude to **Dr. Rekha.B.Venkatapur**, **Professor and Head**, Department of Computer Science & Engineering, for providing a very good facilities and all the support forwarded to us in carrying out this Project Work Phase-I successfully.

We also like to thank our Project Coordinators, **Mr. K Venkata Rao, Associate Professor, Mrs. Vaneeta M, Associate Professor, Mr. Raghavendrachar S, Asst. Professor, Mr. Aditya Pai H, Asst. Professor, and Mrs. Sneha K, Asst. Professor, Department of Computer Science & Engineering** for their help and support provided to carry out the Project Work Phase-I successfully.

Also, we are thankful to **Mrs. Sougandhika Narayan,** Assistant Professor, for being our Project Guide, under whose able guidance this project work has been carried out Project Work Phase-I successfully.

We are also thankful to the teaching and non-teaching staff of Computer Science & Engineering, KSIT for helping us in completing the Project Work Phase-I work.

AAFREEN HUSSAIN
AKSHITHA B.S
ANOOP P S

ABSTRACT

CAPTCHAs are computer-generated tests that humans can pass but current computer systems cannot. CAPTCHAs provide a method for automatically distinguishing a human from a computer program, and therefore can protect Web services from abuse by so-called “bots.” Most CAPTCHAs consist of distorted images, usually text, for which a user must provide some description. Unfortunately, visual CAPTCHAs limit access to the millions of visually impaired people using the Web. Audio CAPTCHAs were created to solve this accessibility issue. Briefly, audio CAPTCHAs are sound files which consist of human sound under heavy noise where the speaker pronounces a bunch of digits consecutively. Generally, these sound files are composed of a set of words to be identified, layered on top of noise and some periodic and non-periodic noises to get difficult to recognize them with a program but not for a human listener. However, with the advancements in deep learning, it becomes easier to build deep learning models that can efficiently recognize text, image, and audio-based CAPTCHAs. So, we gather numerous randomly generated captcha files to train our neural network model and test the model’s ability to recognize characters from audio and image captcha files. The objective of this project is to identify alpha-numeric characters with the use of neural networks. We construct suitable neural network and train it suitably. Our model will be able to extract the characters one by one and map the target output for training purpose. Further, the performance of our model will be evaluated based on various performance metrics like accuracy, sensitivity, specificity, precision, recall.

Keywords: CAPTCHA Analysis, Deep Learning, Support, Image Processing, Melspectogram, Neural Network, performance metrics.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Overview of the project:

CAPTCHA was concocted in 2000 at Carnegie Mellon University by John Langford, Nicholas J. Hooper what's more, Luis Von Ahn [8]. CAPTCHA is an acronym for "completely Automated Public Turning Test to tell Humans and computer Apart". The advance of Web, Web security has turned into an essential issue. There are an excessive number of malevolent dangers over the Internet which may trade off your framework without any secure application which gives insurance against such dangers. One such danger is the Bot. CAPTCHA is one such shield which can be utilized as an insurance from these malignant projects like Bot.

CAPTCHA is a verification procedure in view of test reaction verification. CAPTCHA furnishes an instrument with the help of which a client's can secure themselves for spam and secret key decoding by taking a straightforward test. In this test a client will see either a picture or a content which are regularly misshaped. The client should enter the example precisely as appeared to him if the CAPTCHA depends on content. In the event that the CAPTCHA depends on picture the client should enter the right name of the picture which accurately symbolizes.

However, with the advancements in deep learning, it becomes easier to build deep learning models that can efficiently recognize text, and audio-based CAPTCHAs. So we gather numerous randomly generated captcha files to train our neural network model and test the model's ability to recognize characters from audio and text-based captcha files.

1.2 Categories of Captcha:

A CAPTCHA may come in different structures like content based or picture based CAPTCHA. In recent years, many types of CAPTCHAs have been developed. Some are based on Optical Character Recognition (OCR) such as text CAPTCHA, whereas others are based on Non-Optical Character Recognition (Non-OCR) which uses multimedia, such as voice and video.

The CAPTCHAs can be classified into different types depend on what is distorted that is whether characters, digits, or images. These types are given below:

1.2.1 Text-based captchas:

Text, Image - based CAPTCHA is the most common way of usage of CAPTCHAs. These consist of distorted images, mostly test images, that a user must write some description about that image [1]. Usually these are recognized very easily by humans but are difficult to be understandable to machines or robots. While numerous alternatives to text-based image captchas have been proposed many websites and applications still use text-based captchas as a security and authentication mechanism [8]. Due to the wide deployment of text-based captchas, a compromise on the scheme can have significant implications and could result in serious consequences.

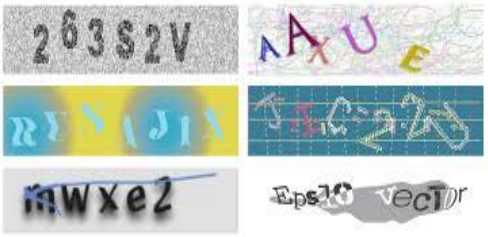

Fig 1.2.1 Text-based Captcha

1.2.2 Audio-based captchas:

Due to improper visual entries of text-based CAPTCHA, AUDIO CAPTCHAs were invented. These CAPTCHAs depend on the sound-based frameworks. Briefly, audio CAPTCHAs are sound files which consist of human sound under heavy noise where the speaker pronounces a bunch of digits, characters consecutively. Generally, in those audio files, there are some periodic and non-periodic noises to get difficult to recognize them. A web client has to properly diagnose the digits or characters pronounced in the audio file to elapse the CAPTCHA.

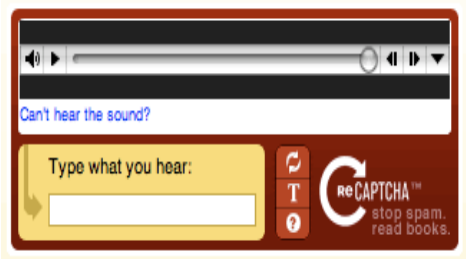

Fig 1.2.2 Audio Captcha

1.2.3 Image-based captchas:

CAPTCHAs are challenge-tests in which the clients need to figure those pictures that have some comparability. For instance: visual riddles. In picture based CAPTCHAs client is required to character picture. The upside of picture based CAPTCHA is that example acknowledgment is hard AI issue and along these lines it is hard to break this test utilizing design acknowledgment system.

Fig 1.2.3 Image Based Captcha

1.2.4 Video-based captchas:

Video CAPTCHA is a more current and less normally observed CAPTCHA framework. In video-based CAPTCHAs, three words (labels) are given to the client which portrays a video. The client's label must match to an arrangement of consequently produced ground truth labels at that point just the test is said to be passed. The term video CAPTCHA is utilized to any CAPTCHA that uses a video as its way to display data to a client . In spite of the fact that video CAPTCHA is constrained, both business and scholarly application do exist.

Fig 1.2.4 Video Based Captcha

1.3 Definitions

- **Deep Learning** - Deep learning is an AI function that mimics the workings of the human brain in processing data for use in detecting objects, recognizing speech, translating languages, and making decisions. Deep learning AI is able to learn without human supervision, drawing from data that is both unstructured and unlabeled.
- **Neural Networks -** are a series of algorithms that mimic the operations of a human brain to recognize relationships between vast amounts of data. They are used in a variety of applications in financial services, from forecasting and marketing research to fraud detection and risk assessment.
- **Tensor Flow -** TensorFlow is a free and open-source software library for machine learning. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. Tensorflow is a symbolic math library based on dataflow and differentiable programming.
- **Python -** Python is an interpreter, object-oriented, high-level programming language with dynamic semantics, created by Guido van Rossum and first released in 1991. Python features a dynamic type system and automatic memory management and supports multiple programming paradigms, including object-oriented, imperative, functional programming, and procedural styles. It has an expansive and extensive standard library its abnormal state worked in information structures, joined with dynamic writing make it extremely alluring for Rapid Application Development.
- **Anaconda -** Anaconda is a free and open source distribution of the Python and R programming languages for data science and machine learning related applications(large-scale data processing, predictive analytics, scientific computing), that aims to simplify package management and deployment. Package versions are managed by the package management system conda.

CHAPTER 2

LITERATURE SURVEY

Although CAPTCHAs are designed to block automated bots. CAPTCHAs are themselves automated. They're programmed to pop up in certain places on a website, and they automatically pass or fail users.

2.1 Breaking of Audio Captchas

We have considered the exploration papers identified with the Audio Captcha Recognition Using RastaPLP Features by SVM as given beneath by [1] Ahmet Faruk Çakmak, et.al. In their paper, authors have used the technique relative spectral transform-perceptual linear prediction (RASTA-PLP) and PLP. By using RASTA-PLP, they could be able to train classifiers to identify words and digits all by itself of who pronounce them. In the sample files, they also added vocal or music noise that makes the problem more challenging. In this problem, they gathered 900 audio CAPTCHAs from various websites: google.com, digg.com, and recaptcha.net. Each of the CAPTCHAs annotated with the information regarding digit locations provided by the manual transcriptions. They randomly selected 800 samples for training and used the remaining 100 for testing set.

First, the audio file was divided into segments of noise or words to be able to reveal the audio CAPTCHA. In the test files, the words include digits. Next, they managed to identify where the locations of the digits in the CAPTCHA file starts and finishes. Following the audio files are converted into Mel- Spectrogram.

In [8], spectrogram is a visual depiction of a signal's frequency composition over time. The Mel spectrogram is used to provide our models with sound information similar to what a human would perceive. The raw audio waveforms are passed through filter banks to obtain the Mel spectrogram.

Whenever the frequency of the digit pronounced matched with the frequency of the digit in the audio files then the digit was recognized at that particular interval of time. According to their tests, 98% accuracy was obtained for individual digit recognition precision, and around 89% accuracy for the entire digit recognition.

2.2 Breaking of Text Captchas

In [2] Guixin Ye et.al have developed a captcha solver based on the generative adversarial network (GAN). Unlike other machine learning techniques, which uses a large volume of data they have used fewer real captchas.

They achieved by first learning a captcha synthesizer to automatically generate synthetic captchas that is captchas which is visually similar to the target ones. GAN consists of two models: a generative network for creating synthetic examples and a discriminative network to distinguish the synthesized examples from the real ones. They used backpropagation to train both networks, so that over the training iterations, the generator produces better synthetic samples, while the discriminator becomes more skilled at flagging synthetic samples. They pre-processed using Pix2Pix image to image translation framework. This transform an image from one style to another. The training goal is to learn a generator to remove security features and standardize the font style.

Following this data was given to a base solver for a target captcha scheme. This base solver was based on Convolutional Neural Network.

Finally, they applied transfer learning to refine the base solver by using small set of manually labelled captchas that are collected from target websites.

In [6] Md Fazuel Kader et.al, an artificial neural network based color and size invariant character recognition system was proposed which was able to recognize English characters (A~Z) and numbers (0~9) successfully. The feed-forward network has two layers: one is input layer and another is output layer. No hidden layer is used. A supervised manner was used to train the neural network.

2.3 Analysis

- **Construction of Standard Test Database for Text-Based CAPTCHA** - A rich and high quality text-based CAPTCHA image database is the necessary foundation for the research of text-based CAPTCHA breaking. At present, the researchers get CAPTCHA images mainly by web access and software generation. However, due to the diversity and timeliness of text-based CAPTCHA, it has not been possible to construct a common image database in the field of text-based CAPTCHA recognition.

  It is necessary to collect, classify, organize, and establish the text-based CAPTCHA images database. The database can provide the reliable training and testing data for research work and also provide the premise and basis of unified evaluation for various methods in this field.
- **Multi-type CAPTCHA Recognition**- At present, only when training set and test set belong to the same type, the classifier can effectively recognize CAPTCHAs. In fact, there are a variety of character changes in a CAPTCHA. Therefore, it is an arduous and important task to design a reasonable classifier to recognize various types of CAPTCHAs.
- **Segmentation-Free** CAPTCHA Recognition- After more than ten years of development, the text-based CAPTCHA breaking has achieved a high success rate in individual character. However, the breaking success rate of the CAPTCHA string is generally low, and the results are less.
- **Application of Deep Learning Model** - At present, in the field of CAPTCHA recognition, deep learning model can achieve better results than traditional methods. Furthermore, the study of the interrelationships and fusion applications between the various deep learning models is not thorough. We hope that newer and better deep learning models are proposed to make a breakthrough in CAPTCHA recognition, which will certainly promote the development in this field.
- **Misrecognition of Confusable Characters**-When using the deep learning network to extract character features automatically, the characters with similar features are easily confused. It has practical significance to improve the precision of feature extraction and the training methods in the deep learning network.

CHAPTER 3

PROBLEM IDENTIFICATION

3.1 Problem Statement

- Audio CAPTCHAs are sound files which consist of human sound under heavy noise where the speaker pronounces a bunch of digits consecutively.
- Generally, in those sound files, there are some periodic and non-periodic noises to get difficult to recognize them with a program but not for a human listener.
- We gathered numerous randomly collected audio files to train and then test them using our deep learning model to be able to extract digits out of each conversation.

3.2 Project Scope

- Each captcha can be used only once for validation. Hence there is a need for n number of captchas. Therefore, in our project, we generate n number of captchas using python script.
- The application asks for a media file and generates an api request which loads image/audio captcha thus predicting the appropriate result.

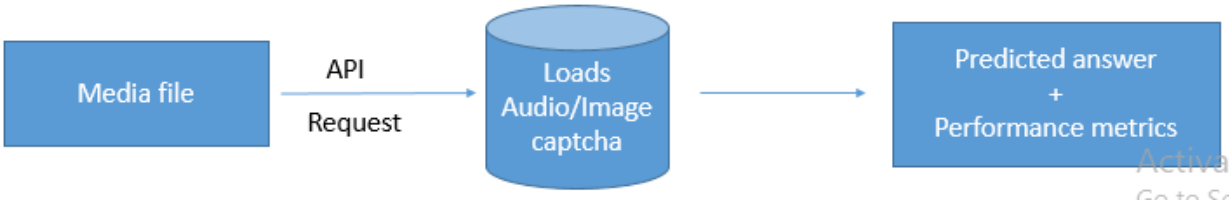

Fig.3.2 Prediction process

CHAPTER 4

GOALS AND OBJECTIVES

4.1 Project Goals

- The scope of our project covers gathering numerous randomly generated audio and image-based CAPTCHA files to train our deep learning model and then test the model's ability.
- The objective of this project is to identify alpha-numeric characters with the use of neural networks. We construct suitable neural network and train it suitably. Our model will be able to extract the characters one by one and map the target output for training purpose.
- Further, we will also be evaluating the performance of our model based on various performance metrics like accuracy, sensitivity, specificity, precision, recall, etc.

4.2 Project Objectives

- Generation of annotated datasets of Audio and Text Captchas.
- The objective of this project is to identify alpha-numeric characters with the use of neural networks. We construct suitable neural network and train it suitably. Our model will be able to extract the characters one by one and map the target output for training purpose.
- Further, we will also be evaluating the performance of our model based on various performance metrics like accuracy, sensitivity, specificity, precision, recall, etc.

CHAPTER 5

SYSTEM REQUIREMENT SPECIFICATION

A software requirements specification (SRS) is a comprehensive description of the intended purpose and environment for software under development. The SRS fully describes what the software will do and how it will be expected to perform. Software requirements specification permits a rigorous assessment of requirements before design can begin and reduces later redesign. It should also provide a realistic basis for estimating product costs, risks, and schedules.

The software requirements specification document enlists enough and necessary requirements that are required for the project development. To derive the requirements we need to have clear and thorough understanding of the products to be developed or being developed. This is achieved and refined with detailed and continuous communications with the project team and customer till the completion of the software.

5.1 Hardware Requirements

We use the following hardware requirements:

- Processor : Intel i3
- RAM : 4GB
- Hard-Disk : 500GB
- GPU (Preferably Nvidia GTX710 and its equivalent)

5.2 Software Requirements

We use the following software requirements:

- Coding Language: Python 3.5(Python **(packages -** numpy, pandas, matplotlib, tensorflow, keras)
- Jupyter Notebook
- IDE: Anaconda
- Operating System: Windows 7 and above

CHAPTER 6

METHODOLOGY

Following are the steps for Character Recognition:

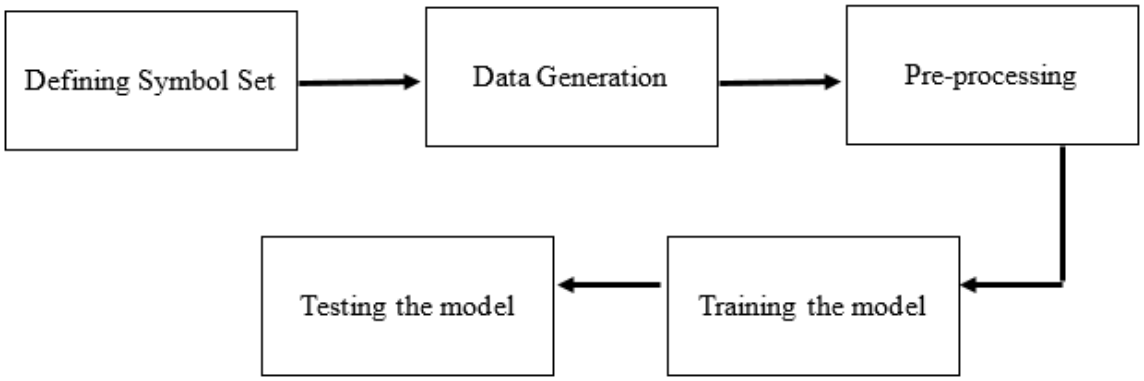

Fig.6.1 Methodology

- **Defining symbol set** - List of alpha-numeric characters grouped in a text file. We generate numerous captchas on randomly choosing characters from this text file.

- **Data Generation** - We develop a script which generates certain amount of captchas from the set of symbols based on the dimensions, length of the symbols specified. And this generated captchas are scrambled. Without scrambling the name of the image is the captcha text. To train and validate a neural network, we need two sets of data: a big training set, and a smaller validation set. The network is trained on the training set, and tested on the validation set, so it is very important that there are no images that are in both sets. To generate the training data, the "ground truth" classification for each training example image must be known. This means that for training, the names of the captchas cannot be scrambled, because otherwise the training process has no way to check if the answer from the CNN for some captcha is right or wrong! Make sure not to use the `--scramble` option when generating the training or validation datasets.

- **Pre-processing** - Data Preprocessing is a Data Mining technique that involves transforming raw data into an understandable format.

  The various Data pre-processing techniques involved are as follows:

  - **Dimension reduction of the audio using Mel Spectrogram -**

    An acoustic time-frequency representation of a sound. A spectrogram is a visual depiction of a signal's frequency composition over time. The Mel scale provides a linear scale for the human auditory system .The Mel spectrogram is used to provide our models with sound information similar to what a human would perceive. The raw audio waveforms are passed through filter banks to obtain the Mel spectrogram. Our models look to learn features from this representation.

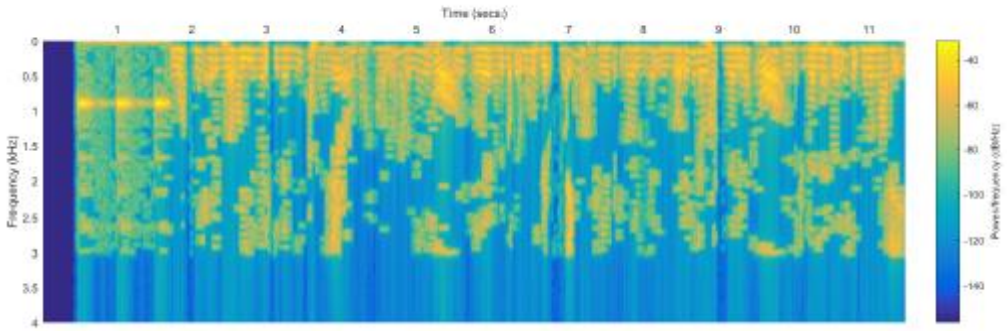

Fig.6.2 Mel-spectrogram of an audio

  - **Converting amplitude into decibels –**

    Decibels are a relative unit, they express the power of your signal relative to some reference power. This is used to convert an amplitude spectrogram to decibel spectrogram. This can be done using librosa.amplitude_to_db.

  - **Plotting audio as a spectrogram in greyscale –**

    A grayscale image is a data matrix whose values represent intensities of one image pixel. While grayscale images are rarely saved with a color map, MATLAB uses a color map to display them.

- **Auto-Contrasting of an image –**

  Contrast is the difference in luminance or color that makes an object (or its representation in an image or display) distinguishable. The maximum contrast of an image is the contrast ratio or dynamic range.

  autocontrast() method maximizes (normalize) image contrast. This function calculates a histogram of the input image, removes cutoff percent of the lightest and darkest pixels from the histogram, and remaps the image so that the darkest pixel becomes black (0), and the lightest becomes white (255).

➢ Training the model –

Training a model simply means learning (determining) good values for all the weights and the bias from labeled examples. The process of training an ML model involves providing an ML algorithm (that is, the learning algorithm) with training data to learn from. The term ML model refers to the model artifact that is created by the training process.

The training data must contain the correct answer, which is known as a target or target attribute. The learning algorithm finds patterns in the training data that map the input data attributes to the target (the answer that you want to predict), and it outputs an ML model that captures these patterns.

Here we build a Neural network model on the available training set.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area. The objective of the Convolution Operation is to extract the high-level features such as edges, from the input image. We choose CNN mainly because of its ability to analyze spatial invariant features and using a relatively small number of parameters.

The **convolution layer** uses filters to translate over the input and then takes the inner product before adding the bias. Each filter has its own set of weights and bias. The weights and bias are the only parameters to train. Each layer can have multiple filters to learn different features. This gives CNN the benefits of relatively small number of parameters to learn and being able to learn spatial invariant features.

The **pooling layer** is used to reduce the dimensions of the subsequent layers. The commonly used pooling techniques are maxpooling and average pooling, where maxpooling takes the maximum value of the pooling window and average pooling takes the average value.

For hidden layers in the model, we use Relu **activation function**. Basically, the activation parameter to the Conv2D class is simply a convenience parameter which allows you to supply a string.

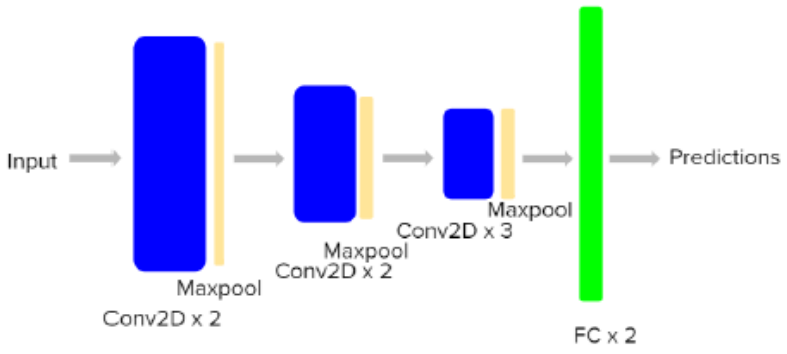

Fig.6.3 CNN-Architecture

➢ Testing the model –

The next step after implementing a machine learning algorithm is to find out how effective is our model based on metric and datasets. Different performance metrics are used to evaluate different Machine Learning Algorithms.

The various performance metrics which will be evaluated are as follows:

- **Accuracy** - is the number of correctly predicted data points out of all the data points.
- **Precision** - is a metric that quantifies the number of correct positive predictions made. Precision, therefore, calculates the accuracy for the minority class. It is calculated as the ratio of correctly predicted positive examples divided by the total number of positive examples that were predicted.
- **Sensitivity** - is a measure of the proportion of actual positive cases that got predicted as positive.
- **Specificity** - is defined as the proportion of actual negatives, which got predicted as the negative.
- **Recall** - is the fraction of the total amount of relevant instances that were actually retrieved.

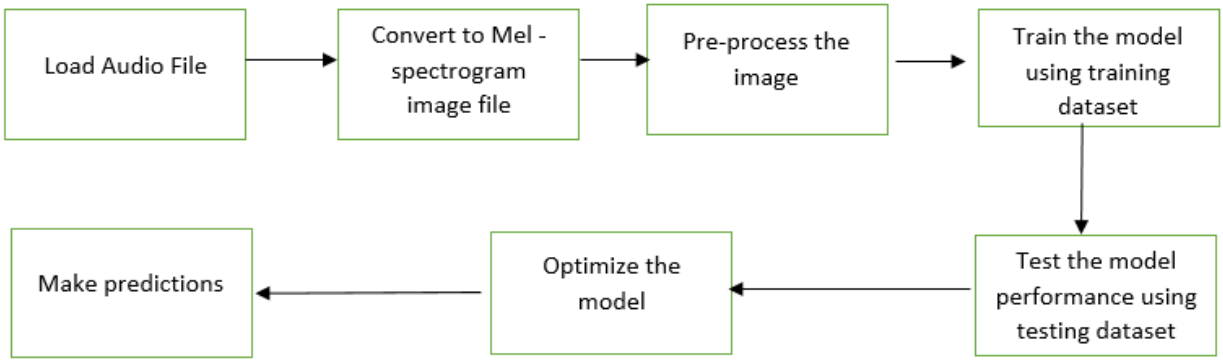

Fig.6.4 Block Diagram

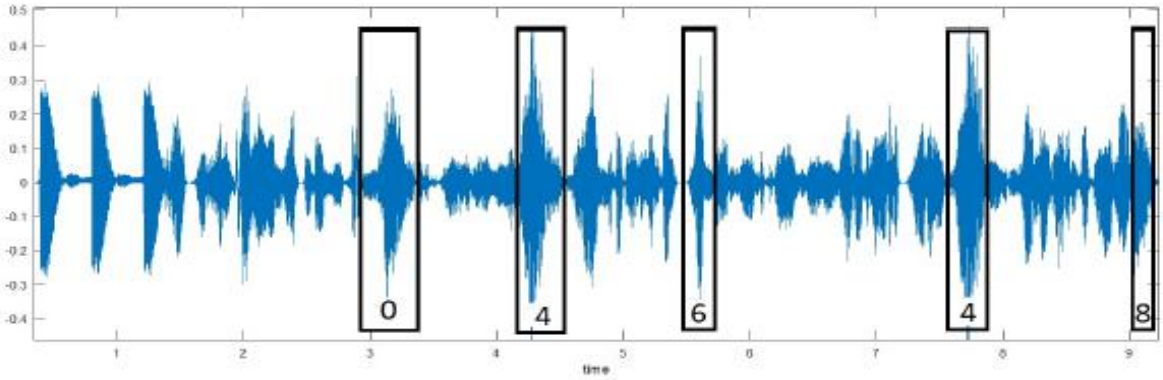

Fig.6.5 Audio Captcha Recognition

CHAPTER 7

APPLICATIONS

Following are the several applications:

- Irregular text recognition –

  Despite great progress recently, it remains a challenging task because the potential irregularity and diversity of text shapes and layouts in the wild, which can be curved, oriented or distorted, make the misalignment between the output character sequence and the two-dimensional input image.

- Number plate recognition –

  A technology that uses optical character recognition on images to read vehicle registration plates to create vehicle location data.

Fig.7.1 Number plate recognition

- Handwriting recognition –

  Handwriting recognition systems use pattern matching to convert handwritten letters into corresponding computer text or commands in real time.

- Optical character recognition –

  Optical Character Recognition, or OCR, is a technology that enables you to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera into editable and searchable data.

Chapter 8

CONTRIBUTION TO SOCIETY AND ENVIRONMENT

Captchas for Authentication/Verification system

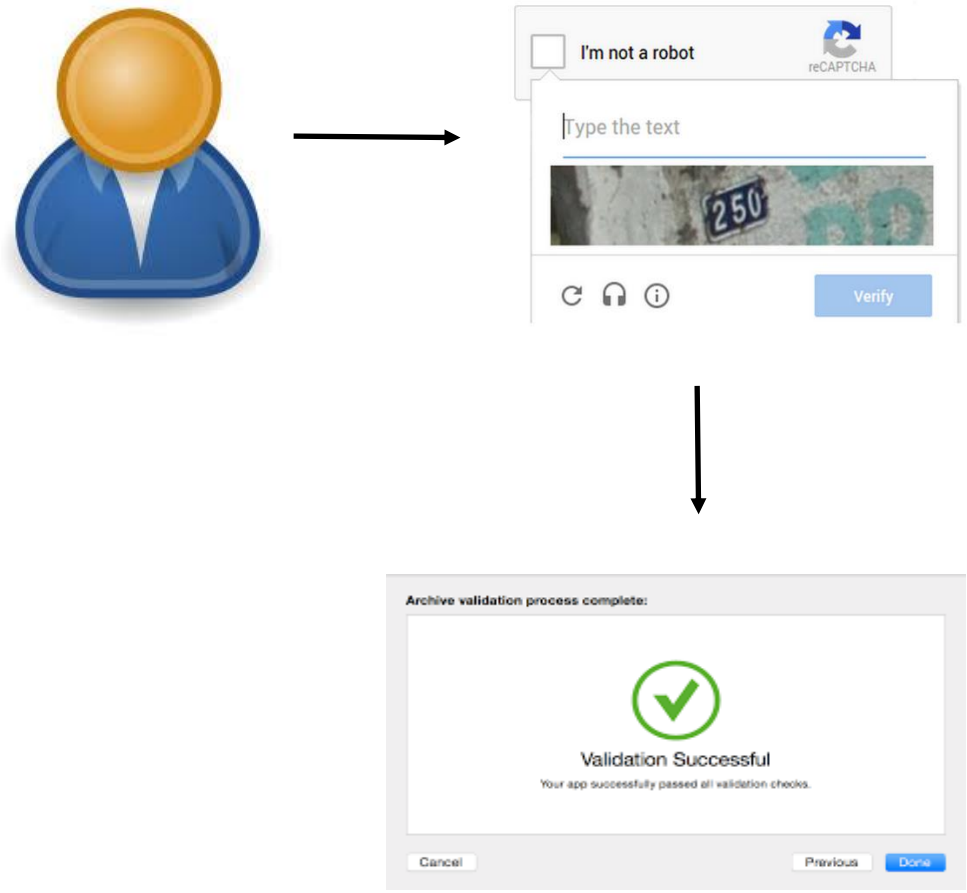

Fig.8.1 Authentication Process

REFERENCES

[1] Cakmak, Ahmet & Balcilar, Muhammet. (2019). Audio Captcha Recognition Using RastaPLP Features by SVM.

[2] Guixin Ye, Zhanyong Tang, Dingyi Fang, Zhanxing Zhu, Yansong Feng, Pengfei Xu, Xiaojiang Chen, and Zheng Wang. 2018. Yet Another Text Captcha Solver: A Generative Adversarial Network Based Approach. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security* (*CCS '18*). Association for Computing Machinery, New York, NY, USA, 332–348.

[3] Mori, G. & Malik, J.. (2003). Recognizing objects in adversarial clutter: breaking a visual CAPTCHA. IEEE Conf Comput Vision Pattern Recogn. 1. I-134. 10.1109/CVPR.2003.1211347.

[4] Tam, Jennifer & Simsa, Jirí & Hyde, Sean & Ahn, Luis. (2008). Breaking audio CAPTCHAs. Advances in Neural Information Processing Systems. 1625-1632.

[5] Chen, Jun & Luo, Xiangyang & Guo, Yanqing & Zhang, Yi & Gong, Daofu. (2017). A Survey on Breaking Technique of Text-Based CAPTCHA. Security and Communication Networks. 2017. 1-15. 10.1155/2017/6898617.

[6] Kader, Md Fazlul & Kaushik, Deb. (2012). Neural Network-Based English Alphanumeric Character Recognition. International Journal of Computer Science, Engineering and Applications. 2. 10.5121/ijcsea.2012.2401.

[7] Hasan, Walid. (2016). A Survey of Current Research on CAPTCHA. International Journal of Computer Science & Engineering Survey. 7. 1-21. 10.5121/ijcses.2016.7301.

[8] Zhang, B., Leitner, J. and Thornton, S., n.d. Audio Recognition using Mel Spectrograms and Convolution Neural Networks. San Diego: Boyang Zhang, p.5.

[9] Hermansky, Hynek & Morgan, Nathaniel & Bayya, A. & Kohn, P.. (1992). RASTA-PLP speech analysis technique. 1. 121 - 124 vol.1. 10.1109/ICASSP.1992.225957.

[10] Sinha, Anvesh & Tarar, Sandhya. (2016). Review Paper on Different CAPTCHA Techniques. www.ijcst.com. 7. 174-176.

[11] Shinde, Vishal and Prof. Vijay Rathi. “DIFFERENT TYPES OF CAPTCHA : A LITERATURE SURVEY.” (2018).

APPENDIX I

SURVEY ON ALPHA-NUMERIC CHARACTER RECOGNITION IN AUDIO/ TEXT-BASED CAPTCHA

Anoop P S
Computer Science and Engineering
K. S. Institute of Technology
Bengaluru, Karnataka
anoop.purohith.04@gmail.com

Akshitha B S
Computer Science and Engineering
K. S. Institute of Technology
Bengaluru, Karnataka
akshithabsyadav@gmail.com

Aafreen Hussain
Computer Science and Engineering
K. S. Institute of Technology
Bengaluru, Karnataka
aafreenhussain1999@gmail.com

Mentor/Co-Author:
Prof.Sougandhika Narayan
Assistant Professor,Dept of CSE
K. S. Institute of Technology
Bengaluru, Karnataka
sougandhikanarayan@ksit.edu.in

***Abstract* - CAPTCHAs are computer-generated tests that humans can pass but current computer systems cannot. CAPTCHAs provide a method for automatically distinguishing a human from a computer program, and therefore can protect Web services from abuse by so-called "bots." Most CAPTCHAs consist of distorted images, usually text, for which a user must provide some description. Unfortunately, visual CAPTCHAs limit access to the millions of visually impaired people using the Web. Audio CAPTCHAs were created to solve this accessibility issue. Briefly, audio CAPTCHAs are sound files which consist of human sound under heavy noise where the speaker pronounces a bunch of digits consecutively. Generally, these sound files are composed of a set of words to be identified, layered on top of noise and some periodic and non-periodic noises to get difficult to recognize them with a program but not for a human listener. However, with the advancements in deep learning, it becomes easier to build deep learning models that can efficiently recognize text, image, and audio-based CAPTCHAs. So, we gather numerous randomly generated captcha files to train our neural network model and test the model's ability to recognize characters from audio and image captcha files. The objective of this project is to identify alpha-numeric characters with the use of neural networks. We construct suitable neural network and train it suitably. Our model will be able to extract the characters one by one and map the target output for training purpose. Further, the performance of our model will be evaluated based on various performance metrics like accuracy, sensitivity, specificity, precision, recall.**

I. INTRODUCTION

CAPTCHA was concocted in 2000 at Carnegie Mellon University by John Langford, Nicholas J. Hooper what's more, Luis Von Ahn [8]. CAPTCHA is an acronym for "completely Automated Public Turning Test to tell Humans and computer Apart" . The advance of Web, Web security has turned into an essential issue. There are an excessive number of malevolent dangers over the Internet which may trade off your framework without any secure application which gives insurance against such dangers. One such danger is the Bot. A Bot is a malevolent program which has the ability to run mechanized errands over the system and in this manner making issue in the system. CAPTCHA is one such shield which can be utilized as an insurance from these malignant projects like Bot.

The Bot operation is like invert "TURING TEST" where the program demonstrations like judge and the other individual acts like client. CAPTCHA is likewise called as a test reaction test which gives a test to the clients, when the client gives adjust reply he is considered as human generally a web bot. CAPTCHA is a verification procedure in view of test reaction verification. CAPTCHA furnishes an instrument with the help of which a client's can secure themselves for spam and secret key decoding by taking a straightforward test. In this test a client will see either a picture or a content which are regularly misshaped. The client should enter the example precisely as appeared to him if the CAPTCHA depends on content. In the event that the CAPTCHA depends on picture the client should enter the right name of the picture which accurately symbolizes.

A CAPTCHA may come in different structures like content based or picture based CAPTCHA. In recent years, many types of CAPTCHAs have been developed. Some are based on Optical Character Recognition (OCR) such as text CAPTCHA, whereas others are based on Non-Optical Character Recognition (Non-OCR) which uses multimedia, such as voice and video.

However, with the advancements in deep learning, it becomes easier to build deep learning models that can efficiently recognize text, and audio-based CAPTCHAs. So we gather numerous randomly generated captcha files to train our neural network model and test the model's ability to recognize characters from audio and text-based captcha files.

II. CATEGORIES OF CAPTCHA

The CAPTCHAs can be classified into different types depend on what is distorted that is whether characters, digits, or images .These types are given below:

i. Text-based captchas:

Text, Image - based CAPTCHA is the most common way of usage of CAPTCHAs. These consist of distorted images, mostly test images, that a user must write some description about that image [1]. Usually these are recognized very easily by humans but are difficult to be understandable to machines or robots. While numerous alternatives to text-based image captchas have been proposed many websites and applications still use text-based captchas as a security

and authentication mechanism [8]. Due to the wide deployment of text-based captchas, a compromise on the scheme can have significant implications and could result in serious consequences.

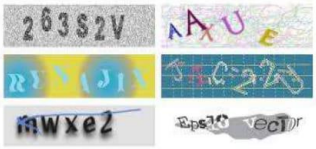

Fig 1. Text-based Captcha

ii. Audio-based captchas :

Due to improper visual entries of text-based CAPTCHA, AUDIO CAPTCHAs were invented. These CAPTCHAs depend on the sound-based frameworks. Briefly, audio CAPTCHAs are sound files which consist of human sound under heavy noise where the speaker pronounces a bunch of digits, characters consecutively. Generally, in those audio files, there are some periodic and non-periodic noises to get difficult to recognize them. A web client has to properly diagnose the digits or characters pronounced in the audio file to elapse the CAPTCHA.

Fig 2. Audio Captcha

III. LITERATURE SURVEY

This section provides a comprehensive review of the techniques used for analysis of text-based captchas.

i. Breaking of Text-Based Captchas

In [2] Guixin Ye et.al have developed a captcha solver based on the generative adversarial network (GAN). Unlike other machine learning techniques, which uses a large volume of data they have used fewer real captchas. They achieved by first learning a captcha synthesizer to automatically generate synthetic captchas that is captchas which is visually similar to the target ones. GAN consists of two models: a generative network for creating synthetic examples and a discriminative network to distinguish the synthesized examples from the real ones. They used backpropagation to train both networks, so that over the training iterations, the generator produces better synthetic samples, while the discriminator becomes more skilled at flagging synthetic samples.They pre-processed using Pix2Pix image to image translation framework. This transform an image from one style to another. The training goal is to learn a generator to remove security features and standardize the font style. Following this data was given to a base solver for a target captcha scheme. This base solver was based on Convolutional Neural Network.

Finally, they applied transfer learning to refine the base solver by using small set of manually labelled captchas that are collected from target websites.

In [3] Greg Mori et.al developed an object detection technique in GIMPY and EZ-GIMPY captchas. GIMPY captchas are sequence of characters which are presented as distorted and corrupted images by adding black and white lines and making non-linear modifications and asking the user to type the characters correctly. These captchas are basically word recognition in the presence of clutter. They used a database of images of known objects and the task was to find these objects in a cluttered environment. They used Shape context descriptor to match handwritten digits and 3D-objects.
They described two-stage approach to object recognition namely,

- Fast pruning : Given a query image, we should be able to quickly retrieve a small set of likely candidate shape and location pairs from a potentially very large collection of stored shapes.
- Detailed matching: Once we have a small set of candidate shapes, we can perform a more expensive and more accurate matching procedure to find the best matching shapes to the query image.

There are two important types of data available to us in solving word recognition tasks – lexical information and visual cues.They achieved 92% of success rate of identifying words in EZ-Gimpy captcha and whereas in Gimpy Captcha they achieved a success rate of 33%.

In [7], Jun Chen et.al described the techniques used for breaking of text-based captchas, which is as follows:

- Pre-processing - The pre-processing of existing CAPTCHA breaking methods mainly includes image binarization, image thinning, denoising. Image binarization is to highlight interesting objects contour and to remove noises in background. Image thinning is to process the character's contour as skeleton. It must not change the character's adhesion. Its purpose is to highlight image contour and to simplify subsequent processing. Denoising is removing some interference lines.
- Segmentation - The segmentation methods based on individual characters segment, character projection, connected components, character contour, character width.
- Combination Methods – The Combination methods based on redundancy is where each character fragment is labelled in order from top to bottom and left to right.
- Recognition Methods – The Recognition methods based on template matching is to compare similarity of each pixel between characters and every template and to find the highest similarity. The matching recognition methods based on global property is traverse scanning. Within search area, the optimal match point to each pixel is found by regional correlation matching calculation. The matching based on Neural network ,For the principle of parallel distributed operation in large number of neurons, the efficient learning algorithms, and the ability to imitate human cognitive systems, the neural network is very suitable to solve problems such as speech recognition and text recognition.
- Post-processing: In post-processing stage, the final results reliability is ensured by simplification, selection, and optimization.

This section provides a comprehensive review of the techniques used for analysis of Audio-based captchas.

ii. **Breaking of Audio-Based Captchas**

We have considered the exploration papers identified with the Audio Captcha Recognition Using RastaPLP Features by SVM as given beneath by [1] Ahmet Faruk Çakmak, et.al.In their exploration paper, authors have used the technique relative spectral transform-perceptual linear prediction (RASTA-PLP) and PLP. By using RASTA-PLP, they could be able to train classifiers to identify words and digits all by itself of who pronounce them. They could identify particular digits during the existence of noise. In the sample files, they also added vocal or music noise that makes the problem more challenging. In this problem, they gathered 900 audio CAPTCHAs from various websites: google.com, digg.com, and recaptcha.net. Each of the CAPTCHAs annotated with the information regarding digit locations provided by the manual transcriptions. For each type of CAPTCHA, they randomly selected 800 samples for training and used the remaining 100 for testing set. First, the audio file was divided into segments of noise or words to be able to reveal the audio CAPTCHA. In the test files, the words include digits. Next, they managed to identify where the locations of the digits in the CAPTCHA file starts and finishes. Following the audio files are converted into Mel- Spectrogram (an acoustic time-frequency representation of a sound). Whenever the frequency of the digit pronounced matched with the frequency of the digit in the audio files then the digit was recognized at that particular interval of time.
According to their tests, 98% accuracy was obtained for individual digit recognition precision, and around 89% accuracy for the entire digit recognition.

In[4], Tam, Jennifer et.al described the security of audio CAPTCHAs used by many popular Web sites was analyzed by running machine learning experiments designed to break them. The techniques used were:

- AdaBoost - Using decision stumps as weak classifiers for AdaBoost, anywhere from 11 to 37 ensemble classifiers are built. The number of classifiers built depends on which type of CAPTCHA we are solving. Each classifier trains on all the segments associated with that type of CAPTCHA, and for the purpose of building a single classifier, segments are labeled by either -1(negative example) or +1 (positive example). A segment can then be classified as a particular letter, digit, or noise according to the ensemble classifier that outputs the number closest to 1.
- SVM - Support vector machine- The scale parameters are stored so that test samples can be scaled accordingly. Then, a single multiclass classifier is created for each set of features using all the segments for a particular type of CAPTCHA. We use cross-validation and grid search to discover the optimal slack penalty (C=32) and kernel parameter (γ=0.011).
- k-NN - k-nearest neighbor(k -NN): k-NN was used as final method for classifying digits. For each type of CAPTCHA, five different classifiers are created by using all of the training data and the five sets of features associated with that particular type of CAPTCHA. Again cross-validation was used to discover the optimal parameter. Euclidian distance was used distance metric.

They achieved correct solutions for test samples with accuracy up to 71%. Such accuracy is enough to consider these CAPTCHAs broken. Training several different machine learning algorithms on different types of audio CAPTCHAs allowed to analyze the strengths and weaknesses of the algorithms so that a design for a more robust audio CAPTCHA could be suggested.

iii. **Audio-Recognition techniques**

In [8], Boyang Zhang et.al has taken advantage of the robust machine learning techniques developed for image classification and applied them on the sound recognition problem. They proposed the use of the Mel spectrogram, a transformation that details the frequency composition of the signal over time. They used a Free-sound Dataset (FSD) which is a collection of crowdsourced annotations of 297,144 audio clips. A subset (4,970) of these audio clips comprise the competition's curated dataset, which have been cleaned and validated to remove label noise. The second dataset is the Yahoo Flickr Creative Commons 100M dataset (YFCC). The YFCC dataset contains 99,206,564 photos and 793,436 videos. The soundtracks of a subset (19,800) of YFCC videos comprise the competition's noisy dataset. All audio data were sampled at 44.1 kHz and range from 0.2 - 30 s in length.

Each raw audio waveform was first processed by trimming the silent sections of the clip and then either further trimmed or zero-padded to equal a length of 2 seconds. Next, each processed clip was transformed into its Mel spectrogram representation. A spectrogram is a visual depiction of a signal's frequency composition over time.
The Mel spectrogram is used to provide our models with sound information similar to what a human would perceive. The raw audio waveforms are passed through filter banks to obtain the Mel spectrogram. After this process, each sample has a shape of 128 x 128, indicating 128 filter banks used and 128 time steps per clip. Then they introduced a Deep CNN Model to classify the audio. Using this self-developed CNN architecture, they achieved a LWLRAP score of 0.813 and a top-5 accuracy of 88.9%, when predicting 80 sound classes on the validation set.

iv. **Character-Recognition techniques**

In [6] Md Fazuel Kader et.al, an artificial neural network based color and size invariant character recognition system was proposed which was able to recognize English characters (A~Z) and numbers (0~9) successfully. The feed-forward network has two layers: one is input layer and another is output layer. No hidden layer is used. A supervised manner was used to train the neural network.

The whole recognition process consists of four basic steps: preprocessing, normalized character matrix creation, network establishment and recognition. Preprocessing consists of digitization, noise removal and boundary detection of the digitized character matrix.

- Input Character Image - Our system is able to recognize any colored printed character image with white background and font size is between 18 and 96.
- Digitization and Matrix Creation from Character Image - In order to able to recognize characters by computer the character image is first digitized into a matrix i.e. transformed into a binary form for the ease of handling by the computer.

- Boundary Detection - After creating the digitized binary matrix from the input character image, the detection of boundary is very much important to recognize character correctly.
  - For top boundary detection, scan the character matrix starts at the top-left corner and remove all rows from top having only 0's.
  - For bottom boundary detection, scan the character matrix starts at the bottom-left corner and remove all rows from bottom having only 0's.
  - For left boundary detection, scan the character matrix starts at the top-left corner and remove all columns from left having only 0's.
  - For right boundary detection, scan the character matrix starts at the top-right corner and remove all columns from right having only 0's.
- Normalization - The process of equating the size of all extracted character bitmaps (binary array).For size invariant character recognition, we have converted the boundary detected input character matrix into 12×8 normalized matrix.

Finally, we have tested our network by more than 20 samples per character on average and give 99.99% accuracy only for numeric digits (0~9), 98% accuracy only for letters (A~Z) and more than 94% accuracy for alphanumeric characters by considering inter-class similarity measurement.

IV. PROPOSED METHODOLOGY

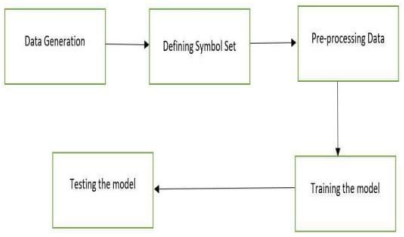

Fig 3. Process of character recognition

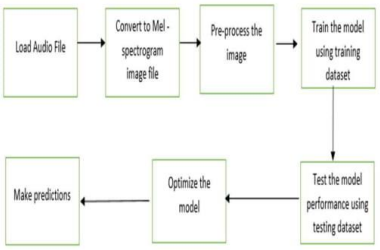

Fig 4. Audio captcha recognition

V. ANALYSIS

- Construction of **Standard Test Database** for Text-Based CAPTCHA - A rich and high quality text-based CAPTCHA image database is the necessary foundation for the research of text-based CAPTCHA breaking. At present, the researchers get CAPTCHA images mainly by web access and software generation. However, due to the diversity and timeliness of text-based CAPTCHA, it has not been possible to construct a common image database in the field of text-based CAPTCHA recognition. It is necessary to collect, classify, organize, and establish the text-based CAPTCHA images database. The database can provide the reliable training and testing data for research work and also provide the premise and basis of unified evaluation for various methods in this field.
- **Multi-type CAPTCHA Recognition**- At present, only when training set and test set belong to the same type, the classifier can effectively recognize CAPTCHAs. In fact, there are a variety of character changes in a CAPTCHA. Therefore, it is an arduous and important task to design a reasonable classifier to recognize various types of CAPTCHAs.
- **Segmentation-Free** CAPTCHA Recognition- After more than ten years of development, the text-based CAPTCHA breaking has achieved a high success rate in individual character. However, the breaking success rate of the CAPTCHA string is generally low, and the results are less.
- **Application of Deep Learning Model** - At present, in the field of CAPTCHA recognition, deep learning model can achieve better results than traditional methods. Furthermore, the study of the interrelationships and fusion applications between the various deep learning models is not thorough. We hope that newer and better deep learning models are proposed to make a breakthrough in CAPTCHA recognition, which will certainly promote the development in this field.
- **Misrecognition of Confusable Characters**- When using the deep learning network to extract character features automatically, the characters with similar features are easily confused. It has practical significance to improve the precision of feature extraction and the training methods in the deep learning network.

ACKNOWLEDGMENT

We would like to convey our gratitude and regards to our Professor Sougandhika Narayan for her knowledge, guidance, and constant supervision as well as for providing necessary information regarding the project & also for her support in planning and development of the project.

We would also like to acknowledge the help, support and feedback extended by all the professors of KSIT.

VI. REFERENCES

[1] Cakmak, Ahmet & Balcilar, Muhammet. (2019). Audio Captcha Recognition Using RastaPLP Features by SVM.

[2] Guixin Ye, Zhanyong Tang, Dingyi Fang, Zhanxing Zhu, Yansong Feng, Pengfei Xu, Xiaojiang Chen, and Zheng Wang. 2018. Yet Another Text Captcha Solver: A Generative Adversarial Network Based Approach. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security* (*CCS '18*). Association for Computing Machinery, New York, NY, USA, 332–348.

[3] Mori, G. & Malik, J.. (2003). Recognizing objects in adversarial clutter: breaking a visual CAPTCHA. IEEE Conf Comput Vision Pattern Recogn. 1. I-134. 10.1109/CVPR.2003.1211347.

[4] Tam, Jennifer & Simsa, Jirí & Hyde, Sean & Ahn, Luis. (2008). Breaking audio CAPTCHAs. Advances in Neural Information Processing Systems. 1625-1632.

[5] Chen, Jun & Luo, Xiangyang & Guo, Yanqing & Zhang, Yi & Gong, Daofu. (2017). A Survey on Breaking Technique of Text-Based CAPTCHA. Security and Communication Networks. 2017. 1-15. 10.1155/2017/6898617.

[6] Kader, Md Fazlul & Kaushik, Deb. (2012). Neural Network-Based English Alphanumeric Character Recognition. International Journal of Computer Science, Engineering and Applications. 2. 10.5121/ijcsea.2012.2401.

[7] Hasan, Walid. (2016). A Survey of Current Research on CAPTCHA. International Journal of Computer Science & Engineering Survey. 7. 1-21. 10.5121/ijcses.2016.7301.

[8] Zhang, B., Leitner, J. and Thornton, S., n.d. Audio Recognition using Mel Spectrograms and Convolution Neural Networks. San Diego: Boyang Zhang, p.5.

[9] Hermansky, Hynek & Morgan, Nathaniel & Bayya, A. & Kohn, P.. (1992). RASTA-PLP speech analysis technique. 1. 121 - 124 vol.1. 10.1109/ICASSP.1992.225957.

[10] Sinha, Anvesh & Tarar, Sandhya. (2016). Review Paper on Different CAPTCHA Techniques. www.ijcst.com. 7. 174-176.

[11] Shinde, Vishal and Prof. Vijay Rathi. "DIFFERENT TYPES OF CAPTCHA : A LITERATURE SURVEY." (2018).

Abstract of Paper in CSI Proceedings:

Department of Computer Science & Engg., K.S Institute of Technology (www.ksit.edu.in)

ISBN:978-81-929425-1-3

CSIAMSPID007 :SURVEY ON ALPHA-NUMERIC CHARACTER RECOGNITION IN AUDIO/ TEXT-BASED CAPTCHA

Anoop P S, Akshitha B S, Aafreen Hussain, Prof.Sougandhika Narayan
K. S. Institute of Technology, Bengaluru

Abstract: CAPTCHAs are computer-generated tests that humans can pass but current computer systems cannot. CAPTCHAs provide a method for automatically distinguishing a human from a computer program, and therefore can protect Web services from abuse by so-called "bots." Most CAPTCHAs consist of distorted images, usually text, for which a user must provide some description. Unfortunately, visual CAPTCHAs limit access to the millions of visually impaired people using the Web. Audio CAPTCHAs were created to solve this accessibility issue. Briefly, audio CAPTCHAs are sound files which consist of human sound under heavy noise where the speaker pronounces a bunch of digits consecutively.

Generally, these sound files are composed of a set of words to be identified, layered on top of noise and some periodic and non-periodic noises to get difficult to recognize them with a program but not for a human listener. However, with the advancements in deep learning, it becomes easier to build deep learning models that can efficiently recognize text, image, and audio-based CAPTCHAs.

So, we gather numerous randomly generated captcha files to train our neural network model and test the model's ability to recognize characters from audio and image captcha files. The objective of this project is to identify alpha-numeric characters with the use of neural networks. We construct suitable neural network and train it suitably. Our model will be able to extract the characters one by one and map the target output for training purpose. Further, the performance of our model will be evaluated based on various performance metrics like accuracy, sensitivity, specificity, precision, recall.

Proceedings of "34th CSI Karnataka State Student Convention (Online)", on theme "**Self-Reliance & Automation**", organized on 22nd and 23rd December-2020

53

APPENDIX II

CSI Certificates:


Kammavari Sangham (R) - 1952

K.S. INSTITUTE OF TECHNOLOGY

(Affiliated to VTU , Belagavi, Approved by AICTE, New Delhi & Accredited by NAAC & IEI)
No. 14, Raghuvanahalli, Kanakapura Main Road, Bengaluru – 560109

Department of Computer Science and Engineering

34TH CSI KARNATAKA STATE STUDENT CONVENTION

THEME: SELF-RELIANCE AND AUTOMATION

CERTIFICATE OF PARTICIPATION

This is to certify that Mr./Ms.

Aafreen Hussain

from K S Institute of Technology college

has participated and presented a paper entitled

Survey On Alpha-Numeric Character Recognition In Audio/Text-Based Captcha

in the domain Data Analytics, Artificial Intelligence, Machine Leaning and Software Engineering in 34th CSI Karnataka State Student Convention Jointly organized by Computer Society of India, Region V, Bangalore Chapter and Department of Computer Science and Engineering at K. S. Institute of Technology, Bangalore on 22nd and 23rd December, 2020.

Dr. Nalini N, Chairperson, CSI-BC
Mrs. Anitha Venkatesh, State Student Coordinator, CSI
Prof. K. Venkata Rao, Event Convener, Dept. of CSE
Dr. Rekha B Venkatapur, Head of the Department, Dept. of CSE
Dr. K. V. A. Balaji, Principal/CEO, K. S. I. T

Sponsored by:


Kammavari Sangham (R) - 1952

K.S. INSTITUTE OF TECHNOLOGY

(Affiliated to VTU , Belagavi, Approved by AICTE, New Delhi & Accredited by NAAC & IEI)
No. 14, Raghuvanahalli, Kanakapura Main Road, Bengaluru – 560109

Department of Computer Science and Engineering

34TH CSI KARNATAKA STATE STUDENT CONVENTION

THEME: SELF-RELIANCE AND AUTOMATION

CERTIFICATE OF PARTICIPATION

This is to certify that Mr./Ms.

Akshitha B. S

from K S Institute of Technology college

has participated and presented a paper entitled

Survey On Alpha-Numeric Character Recognition In Audio/Text-Based Captcha

in the domain Data Analytics, Artificial Intelligence, Machine Leaning and Software Engineering in 34th CSI Karnataka State Student Convention Jointly organized by Computer Society of India, Region V, Bangalore Chapter and Department of Computer Science and Engineering at K. S. Institute of Technology, Bangalore on 22nd and 23rd December, 2020.

Dr. Nalini N, Chairperson, CSI-BC
Mrs. Anitha Venkatesh, State Student Coordinator, CSI
Prof. K. Venkata Rao, Event Convener, Dept. of CSE
Dr. Rekha B Venkatapur, Head of the Department, Dept. of CSE
Dr. K. V. A. Balaji, Principal/CEO, K. S. I. T

Sponsored by:


Kammavari Sangham (R) - 1952

K.S. INSTITUTE OF TECHNOLOGY

(Affiliated to VTU , Belagavi, Approved by AICTE, New Delhi & Accredited by NAAC & IEI)
No. 14, Raghuvanahalli, Kanakapura Main Road, Bengaluru – 560109

Department of Computer Science and Engineering

34TH CSI KARNATAKA STATE STUDENT CONVENTION

THEME: SELF-RELIANCE AND AUTOMATION

CERTIFICATE OF PARTICIPATION

This is to certify that Mr./Ms.

Anoop P.S

from K S Institute of Technology college

has participated and presented a paper entitled

Survey On Alpha-Numeric Character Recognition In Audio/Text-Based Captcha

in the domain Data Analytics, Artificial Intelligence, Machine Leaning and Software Engineering in 34th CSI Karnataka State Student Convention Jointly organized by Computer Society of India, Region V, Bangalore Chapter and Department of Computer Science and Engineering at K. S. Institute of Technology, Bangalore on 22nd and 23rd December, 2020.

Dr. Nalini N, Chairperson, CSI-BC

Mrs. Anitha Venkatesh, State Student Coordinator, CSI

Prof. K. Venkata Rao, Event Convener, Dept. of CSE

Dr. Rekha B Venkatapur, Head of the Department, Dept. of CSE

Dr. K. V. A. Balaji, Principal/CEO, K. S. I. T

Sponsored by: