

College of Natural and Computational

Science

Department of Computer Science

**COURSE TITLE:Introduction To Artificial Intelligence**

**COURSE CODE: CoSc3112**

**SUBMITTED TO INS.Lencho Desalegn**

**GROUP - 4**

***GROUP MEMBERS***

***NAME*** ***ID NO***

1. Adi Tulo …………...................................………..UGR/13195/14

2. Bikila Taye............................................... …..UGR/13491/14

3. Endalew Gonfa …………………………………………UGR/13790/14

4. Lemita Senbeto..........................................…UGA/0205/14

5. Wondimu Seyoum........................................UGA/0173/14

**Project Title:** Optical Character Recognition Using Customized Deep Learning

SUBMISION DATE: May 30/08/2016

JINKA ETHIOPIA 2016 E.C

**Table contents**

**Page**

**[Executive Summary I](#_Toc1262)**

**[1.1 Introduction 1](#_Toc9656)**

[1.2 Objectives 2](#_Toc792)

[1.3 PROJECT SCOPE 2](#_Toc19181)

**[2 Problem of the Statements 2](#_Toc13822)**

**[3.Methodology 3](#_Toc17872)**

[3.1 Data Loading and Preprocessing: 3](#_Toc6458)

[3.2 Model Architecture: 3](#_Toc29833)

[3.3 Model Training and Evaluation: 3](#_Toc1900)

[3.4 Conversion to ONNX: 3](#_Toc12288)

[3.5 ONNX Runtime Evaluation: 4](#_Toc25463)

**[4.Literature Review 4](#_Toc3758)**

**[5. Dataset 4](#_Toc5285)**

**[6 Expected Results/Significance 4](#_Toc8854)**

**[9 Discussion 5](#_Toc13617)**

**[10. Conclusion 6](#_Toc13384)**

**[11. Resources 6](#_Toc3515)**

**[.References 7](#_Toc19180)**

# Executive Summary

This report presents a study on the development of a customized deep learning model for Optical Character Recognition (OCR) tasks. The model combines convolutional neural networks (CNNs) to achieve state-of-the-art results in character recognition and text extraction. By leveraging the strengths of CNNs for feature extraction and RNNs for sequence modeling, the proposed model demonstrates high accuracy rates and robustness to noise and variations in text orientation.The customized deep learning model was evaluated on standard benchmark datasets, including MNIST and IAM Handwriting databases, showing superior performance compared to traditional OCR methods. The model's flexibility and adaptability make it suitable for a wide range of applications, such as document digitization, text analysis, and information retrieval.Optical Character Recognition is the area of Pattern Recognition that has a topic of studies over the past some decades. Optical character recognition is technique of automatically identifying of different character from a record picture additionally provide full alphanumeric recognition of printed or handwritten characters, text numerical, letters, and symbols in to a computer process able layout including ASCII, Unicode and so forth. Optical character recognition is the bottom for many distinct styles of programs in diverse fields, a lot of which we use in our daily lives. Cost effective and less time consuming, corporations, submit offices, banks, security systems, and even the field of robotics hire this system as the base in their Operations. These days, there are numerous portions of research and making use of OCR technology. These OCR technologies help to examine unique documents written in English, Chinese, Hindu, Arabic, Russian, and others languages. On This paper present review of some researches has been made in English, Arabic and Devanagaricharacters. And explained the methodology they use and challenge they face during development of Optical character recognition.

# 

# 1.1 Introduction

Optical Character Recognition (OCR) is a technology that converts scanned images or handwritten text into machine-readable text. Traditional OCR systems often rely on manual feature engineering and template matching techniques, which can be time-consuming and error-prone. In recent years, deep learning approaches have shown great promise in improving the accuracy and efficiency of OCR systems.Deep learning models, such as CNNs and , have revolutionized OCR by automating feature extraction and sequence modeling tasks. CNNs excel at extracting hierarchical features from input images, By combining these two architectures, customized deep learning models can learn to recognize characters and words accurately in various fonts and styles.The goal of this study is to develop a customized deep learning model for OCR tasks that outperforms traditional methods in terms of accuracy, robustness, and efficiency. By training the model on a large dataset of labeled text images, we aim to demonstrate the effectiveness of deep learning in character recognition and text extraction.The MNIST dataset consists of 60,000 training images and 10,000 test images of handwritten digits (0-9).

## 1.2 Objectives

### 1.2.1 General Objectives

* **The main goal** of this project is to develop a customized deep learning-based OCR system that achieves high accuracy across diverse datasets.

### 1.2.2 Specific objectives

**Specific objectives** include:

* Designing a deep learning model tailored for OCR tasks.
* Training the model on a comprehensive dataset to enhance generalization.
* Achieving a recognition accuracy of at least 95% on datasets.
* Training a CNN architecture to achieve high accuracy on the MNIST dataset.
* To collect data ,Preprocessing data ,Model Deployment to Achieve General Objectives

## 1.3 PROJECT SCOPE

The scope Of our product Optical Character Recognition on a grid infrastructure is to

provide an efiicient and enhanced model tool for the users to perform Document image Analysis. processing by reading and recognizing the characters in research,

Academic, governmental and business organizations that are having large pool Of

documented. scanned irnages. Irrespective Of the size of Document and the type Of

characters in documents, the product is recognizing them, searching them and processing them faster according to the needs Of the environrnent

### **2 Problem of the Statements**

### For better and high character recognition accuracy there are so many OCR techniques but still difficult to achieve 100% correct recognition especially for character that has similarity. The challenges I observe during review is many of them related to the data collection and preprocessing if we can identify and rid of those challenges we can get high correct recognition. The following issues created due to collecting input data using digitals camera. Instead of using camera to capture characters or scripts prefer to scan the document but let’s see what those challenges are.

# 3.Methodology

The methodology involves several key steps:

3.1 Data Loading and Preprocessing:

The MNIST dataset is loaded and split into training and testing sets. Images are reshaped to a format suitable for CNN input and normalized to facilitate model training.

3.2 Model Architecture:

A CNN architecture is constructed using Keras, comprising convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification.

## 3.3 Model Training and Evaluation:

The CNN model is trained on the training dataset using categorical cross-entropy loss and Adam optimizer. Performance is evaluated on the test set using metrics such as accuracy.

## 3.4 Conversion to ONNX:

ONNX inferences with 0.002775883674621582 second in average ONNX predicted value: 9

The ONNX's and keras' prediction are matching

The trained Keras model is converted into the ONNX format using keras2onnx. Debugging mode is enabled to ensure a successful conversion, preserving the model's architecture and weights.

## 3.5 ONNX Runtime Evaluation:

The ONNX model is loaded using ONNX Runtime, and its inference performance is benchmarked against the original Keras model. Inference times and prediction accuracies are compared to assess the efficacy of ONNX in a production or deployment scenario.

# 4.Literature Review

Character recognition technique has been completed through studies on different characters for example, English, Arabic, Chinese, Devanagari, Bangla, Farsi and Kannada and so on. Totally, the complete method is carried out in three phase Preprocessing, Feature extraction and recognition. In this paper only cover the study has been done on English, Images and Character 0-9

# Dataset

The MNIST dataset consists of 60,000 training images and 10,000 test images of handwritten digits (0-9). Each image is grayscale, 28x28 pixels.

*Shape of the dataset.*

Train: X=(60000, 28, 28), y=(60000,)

Test: X=(10000, 28, 28), y=(10000,)

* Source of data set is minist

# 6 Expected Results/Significance

* The project is expected to deliver an OCR system with improved accuracy and robustness compared to existing solutions. The contributions to the field include advancements in deep learning techniques for OCR and the development of a versatile tool that can be applied in various domains requiring text recognition.
* OCR Can be use for many things like Convert hard copy document to soft copy .
* This system useful for Photoshope,for online job worker,student and teacher can used it.
* To save the time of work.
* To reduce cost.
* To make work easy.
* Result we get as we seen from blow

### 

# 9.Discussion

* There are two possible reasons that contribute to the explanation of why networks trained on the same data would make different mistakes. For one,we know that the training method, back-propagation, is imperfect. It is subject to getting stuck in local minima. (There exist methods for escaping from local minima, but the computation involved in finding the absolute best point in weight space is usually prohibitive.) More funda mentally, though,there may be no unique "best" place in weight space, but rather many large regions that are equally consistent with the training data. The training set is extremely limited in size, and we believe that it leaves the network undetüetermined.
* We get accuracy 0.955% accuracy
* Other Get 0.966% accuracy

# 

# 10.Conclusion

The objective of the project to develop a web application which simulates an optical character recognition system by providing the functionality of text recognition is satisfied. This functionality is successfully implemented using deep learning by applying a CRNN architecture which is a combination of CNN and RNN. The development of this project using this architecture allowed the exploration of different neural network architectures leading to a hybrid structure of using CNNs and RNNs along with CTC loss. The dataset used for training this architecture is the “MJSynth” synthetic word dataset by the University of Oxford which contains 9 million images of 90k unique English words. The test accuracy when considering 10k images with condition that every character to be exactly correct is obtained to be

0.955%. The scope of this project can be further enhanced by extending it with certain advancements such as applying state-of-the-art text detection techniques. If there is a requirement to deploy the application for real-time usage, it should give state-of-the-art results. Hence in order to achieve good results application-wise, tools such as Google TensorFlow OCR engine, Google Cloud Vision API, AWS TensorFlow can be used. In the case of research-oriented work, existing pre-trained models such as the VGG architecture, ResNet can be explored, analyzed and enhanced along with efficient text detection techniques. These advanced approaches can enable the usage of the OCR system on various types of inputs such

as scanned documents, real-scene images, vehicle number plates and even QR code and barcodes. Artificial Intelligence is a field which is currently witnessing large-scale research and Optical Character Recognition is one of the applications experiencing high demand in terms of performance. Hence continuous development and enhancement both application-wise and research-wise can be carried out on this project.

# 11. Resources

We use many development software and hardware tools in the process of our project. The software tools that we used are such as:-

* **python** with Anaconda navigator and **TensorFlow**, **CNN** Deep learning library.
* we used CNN to evaluate the performance of the model by calculating

different performance metrics.

* **Colab** : Used to edit and run python code on Colab by connecting with Google Drive to Store our Data for the sake of Backup
* We use Graphical Processing Unit to run the code

The hardware tools that we used are processor Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz 3.19 GH, desktop installed RAM 16 GB and above,

# 

# References

**1.** J. Cowell and F. Hussain, “Amharic character recognition using a fast signature based algorithm,” Proc. Int. Conf. Inf. Vis., vol. 2003–Janua, pp. 384–389, 2003.

**2.** I. Stoianov, “Optical Character Recognition of Historical Documents,” Clover.Slavic.Pitt.Edu, 1995.

**3**.G. E. Hinton, S. Osindero, and Y.-W. Teh, “A fast learning algorithm for deep belief nets,” *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006, doi: 10.1162/neco.2006.18.7.1527.

**4 .**N. B. Karayiannis and G. W. Mi, “Growing radial basis neural networks: merging supervised and unsupervised learning with network growth techniques,” *IEEE Transactions on Neural Networks*, vol. 8, no. 6, pp. 1492–1506, Nov. 1997, doi: 10.1109/72.641471.

**APPENDIX** : Abbreviations

OCR :Optical Character Recognition

CNN:convolutional neural networks

ONNX:Open Neural Network Exchange

GPU:Graphical Processing unit

GHz:Giga herzs

RAM: Random Access Memory

GB:Giga byte