

ABSTRACT

The OCR project for Aadhaar cards is designed with the overarching goal of reducing reliance on external APIs such as Karza, thus enhancing data privacy and reducing operational costs. By integrating the YOLOv8 model at the frontend, the system gains advanced object detection capabilities, enabling it to accurately locate and extract Aadhaar card images from webcam input with minimal user intervention. The YOLOv8 model's real-time processing capabilities ensure efficient handling of image streams, further enhancing user experience and system responsiveness. Moreover, the project incorporates robust security measures to safeguard sensitive Aadhaar card information, including encryption protocols and secure data transmission methods. Through these features, the system not only streamlines the OCR process but also prioritises data privacy and security, aligning with regulatory requirements and user expectations. Additionally, the project emphasises scalability and extensibility, allowing for future integration with other identity documents and expansion into diverse use cases beyond KYC verification. In addition to its core functionality of Aadhaar card OCR, the project also masks the Aadhaar number in the provided Aadhaar image as the Aadhaar image must not be stored with the Aadhaar number visible in the company's server as per the RBI rules. The integration of machine learning algorithms not only enhances accuracy in text recognition but also enables adaptive learning capabilities, allowing the system to continuously improve its performance over time through feedback mechanisms and dataset enrichment. Furthermore, the project incorporates a modular architecture, facilitating easy integration with existing software systems and enabling customization to meet specific business requirements.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	i
	LIST OF FIGURES	iii
	LIST OF ABBREVIATIONS	iv
1	INTRODUCTION	1
1.1	INTRODUCTION ABOUT THE COMPANY PLACED	1
1.2	INTRODUCTION OF THE PROJECT	2
2	LITERATURE REVIEW	3
3	PROBLEM DEFINITION AND BACKGROUND	6
3.1	PROBLEM STATEMENT	6
3.2	OBJECTIVE	7
4	SOFTWARE REQUIREMENTS	9
4.1	SOFTWARE SPECIFICATIONS	9
5	PROPOSED SYSTEM	18
5.1	METHODOLOGY	18
6	IMPLEMENTATION	25
7	RESULT AND DISCUSSION	31
8	CONCLUSION AND FUTURE WORKS	40
	REFERENCES	
	PLAGIARISM REPORT	

LIST OF FIGURES

S.No	FIGURE	PAGE NO
1.1	LOGO	1
3.1	OBJECTIVES	8
4.1	FRONTEND TECHSTACK	10
4.2	BACKEND TECHSTACK	10
4.3	HTML,CSS,JAVASCRIPT	11
4.4	JAVA	12
4.5	SPRING BOOT	12
4.6	EASYOCR	13
4.7	TESS4J	14
4.8	OPENCV	15
4.9	YOLOV8	16
4.10	TENSORFLOW JS	17
5.1	DEVELOPMENT PHASE	24
6.1	FLOWCHART	30
7.1	CARD DETECTION PAGE	32
7.2	GRAYSCALE IMAGE	33
7.3	REDACTED OUTPUT	34
7.4	API RESPONSE 1	35
7.5	API RESPONSE 2	35
7.6	API RESPONSE 3	36
7.7	API RESPONSE 4	36
7.8	API RESPONSE 5	37

LIST OF ABBREVIATIONS

	ABBREVIATION
CIFCL	Cholamandalam Investment And Finance Company Limited
NBFCs	Non-Banking Financial Companies
SMEs	Small and Medium Businesses
CSEL	Consumer & Small Enterprise Loans
SBPL	Secured Business Personal Loans
OCR	Optical Character Recognition
Tess4J	Tesseract For Java
Tensorflow JS	Tensorflow JavaScript
OpenCV	Open Computer Vision
API	Application Programming Interface
SSPL	Server Side Public Licence
REST	Representational State Transfer