CLASSIFICATION OF OBJECTS IN MULTIPLE LANGUAGES

# A minor project report submitted for the requirements for the Degree

***Of***

**B.Tech**

By

##### Anshuman Sekhar Dash-1501105478

##### Pruthiraj Marndi-1501105496

##### Swagat Dora-1501105509

##### Sanjay Tudu-1501105501

In

Under the Guidance of

**Prof. D K Swain**

Dept of CSEA, IGIT Sarang



Department of Computer Science Engineering & Application

INDIRA GANDHI INSTITUTE OF TECHNOLOGY,SARANG

DHENKANAL,ODISHA-759146

#### Abstract

**Keyword:** Object Detection, Object Classification, Computer Vision,

Natural Language

This project was an attempt at developing an object detection, classification and language translation system using modern computer vision technology and artificial intelligence. The project delivers an android application that is used to detects and classify objects in any speaking language.

In today’s era of Globalization, people travel to many different countries for many different purposes whether it is business or simply as a tourist. For an individual learning the names of different objects, cuisines and dishes found across the world is a tedious task. This project is to help users overcome this problem with just the use of a Smartphone. The project will use Object detection Technique using Tensorflow to classify objects and then google translate api to get the classification of objects in many different languages.

It can be used in Classification of Different Foods in User's Local Languages and Object Detection and Classification into different languages.

The Application will be developed using android studio and TensorFlow platform. The object will be detected using a smartphone camera and fed into the Tensorflow model which will process and give out the results.

**References**

1. <https://link.springer.com/chapter/10.1007/978-3-642-36124-1_23> Research on Computer Vision-Based Object Detection and Classification
2. Research on Computer Vision-Based Object Detection and Classification by-Juan Wu , Bo Peng , Zhenxiang Huang , Jietao Xie

#### ACKNOWLEDGEMENTS

It is a great pleasure and privilege to express my profound sense of gratitude to Prof. D. K. Swain, Dept. of CSEA, IGIT, Sarang, my guide for his suggestions, motivation and support during the minor project preparation and keen personal interest throughout the progress of my course work.

I would like to express my sincere thanks to Prof. (Dr.) S. N. Mishra, HOD, Dept. of CSEA, IGIT, Sarang, for his suggestion, motivation and support during the minor project preparation.

I have been fortunate enough to have all support motivation from my parents.

I sincerely thanks to the other faculties of the department of CSEA, IGIT, Sarang for their positive and constructive suggestion and kind cooperation during minor project preparation and presentation.

.

**Anshuman Sekhar Dash**

**Pruthiraj Marndi**

**Swagat Dora**

**Sanjay Tudu**

#### DECLARATION BY THE SCHOLAR

I hereby declare that the synopsis entitled **“Classification of Objects in multiple Languages”** under BPUT, ROURKELA, Odisha at **Indira Gandhi Institute Technology, Sarang, Dhenkanal, Odisha** is an authentic record of my work carried out under the supervision **Prof. D K Swain, Dept. of CSEA, IGIT, Sarang.** I have not submitted this minor project report elsewhere for any other degree or diploma.

Anshuman Sekhar Dash

Pruthiraj Marndi

Swagat Dora

Sanjay Tudu

Department of CSEA,

IGIT,Sarang

#### 

#### CERTIFICATE

This is to certify that the project report entitled “**Classification of Objects in multiple Languages** “ is a record of bonafied work carried out by **Anshuman Sekhar Dash , Pruthiraj Marndi , Swagat Dora , Sanjay Tudu** under our supervision and guidance .In our opinion , the report has fulfill requirements of the degree of B.Tech in CSE of the BPUT, ROURKELA, Odisha at **Indira Gandhi Institute Technology,Sarang,Dhenkanal,Odisha** . This work has not been submitted elsewhere for any other degree or diploma.

(Supervisor)

D K Swain

Asst. Professor IGIT Sarang

11-12-2018

#### LIST OF TABLE

**NAME OF THE TABLE TABLE TITLE PAGE NO.**

Table-1 Revision History 2

Table-2 Product Functions 3

Table-3 Glossary 8

#### LIST OF FIGURES

**FIGURE NO. NAME OF THE FIGURE PAGE NO.**

1. User Interfaces 5
2. Analysis Models 9

#### LIST OF ABBREVIATIONS

**OCMLS** Object Classification in Multiple Language System

**SRS** Software Requirement Specification

**YOLO** You Only Look Once

**OpenCV** Open Computer Vision

**EULA** End User License Agreement

**UI** User Interface

**UX** User Experience

**REQ** Requirement

Software Requirements Specification

for

OCMLS

Version 1.0

Prepared by

Anshuman Sekhar Dash

Pruthiraj Marndi

Swagat Dora

Sanjay Tudu

Department of

Computer Science and Engineering

IGIT, Sarang

11-12-2018

DECLARATION

It is hereby declared that this report is an authentic record of my work carried out under the supervision of D K Swain, S K Patra & S K Sahoo, DEPT. OF CSEA, IGIT, SARANG. I have not submitted this report elsewhere for any other degree or diploma.

Anshuman Sekhar Dash

Pruthiraj Marndi

Swagat Dora

Sanjay Tudu

Department of CSEA

IGIT, SARANG

Signature of Co-Guide

S K Patra S K Sahoo

Signature of Guide

D K Swain

Signature of H.O.D

Prof. Sarojananda Mishra

Table of Contents

Table of Contents 2

Revision History 2

1. Introduction 3

1.1 Purpose 3

1.2 Document Conventions 3

1.3 Intended Audience and Reading Suggestions 3

1.4 Product Scope 3

1.5 References 3

2. Overall Description 3

2.1 Product Perspective 3

2.2 Product Functions 3

2.3 User Classes and Characteristics 4

2.4 Operating Environment 4

2.5 Design and Implementation Constraints 4

2.6 User Documentation 4

2.7 Assumptions and Dependencies 4

3. External Interface Requirements 5

3.1 User Interfaces 5

3.2 Hardware Interfaces 5

3.3 Software Interfaces 5

3.4 Communications Interfaces 5

4. System Features 6

4.1 Scan for Objects 6

4.2 Process the Objects 6

4.3 Translation of The Image 6

**Chapter 5. Other Nonfunctional Requirements 7**

5.1 Performance Requirements 7

5.2 Safety Requirements 7

5.3 Security Requirements 7

5.4 Software Quality Attributes 7

Appendix A: Glossary 8

Revision History(If required)

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Date** | **Reason For Changes** | **Version** |
| Anshuman Sekhar Dash | 11-12-2018 | First Draft | 1.0 |
|  |  |  |  |

Table-1

# Introduction

## Purpose

This SRS describes the specifications of the Object Classification in multiple Languages System (OCMLS).This SRS describes the part of OCMLS related to Computer Vision, Deep Learning Model and Application Development.

## Document Conventions

The document focuses on the high priority requirements which will be implemented for the final deliverable.

## Intended Audience and Reading Suggestions

The document is intended for requirements engineer, domain expert, developer, QA engineers and project manager.

Before reading this document it is highly recommended to read the Project Abstract to get an overview of the product.

## Product Scope

1. Detection of objects of the Real World for Classification in many different Languages
2. User corrections feature to get better results in subsequent updates.
3. Search History System to track of the Detected Objects
4. Talk out Loud feature to get the correct pronunciation of words.

## References

1. <https://link.springer.com/chapter/10.1007/978-3-642-36124-1_23> Research on Computer Vision-Based Object Detection and Classification
2. Research on Computer Vision-Based Object Detection and Classification by-Juan Wu , Bo Peng , Zhenxiang Huang , Jietao Xie

# Overall Description

## Product Perspective

The product to specified belongs to the class of object detection applications but unlike other object detecting applications emphasis was put on classifying objects according to the user’s language choice in order to get more specific results.

## Product Functions

|  |  |
| --- | --- |
| **Product Functions** | **Description** |
| Scan for Objects | Look around for objects that can be classified when the user opens the application and draw a bounding box around the object. |
| Accept the object | Get the input from the user of the object that requires to be classified |
| Process and Classify | Process the object image using the feature map inside the inbuilt tensorflow model and return the result. |
| Translate | The Result is translated into the desired language and the result is shown to the user. |
| Word Pronunciation | The voice api is called to generate the pronunciation of the word produced. |
| Error Reporting | The User reports any error made in the classification for further improvement. |

Table-2

## User Classes and Characteristics

1. **Normal Users:**-Less frequency of use, may not use all the functions with minimal technical expertise.
2. **Test Users:-**Frequent use of the application in different operating conditions with high technical expertise, checks and report errors on all the product functions.

## Operating Environment

The OCMLS System application will be running on an android Smartphone with an Inbuilt camera. The Trained deep learning model will be kept inside the application and will internet connection of the Smartphone for translations and error reporting to the cloud servers. It will require ambient light surroundings to operate on.

## 

## …

## Design and Implementation Constraints

1. RTL language readers may face difficulty in reading translated texts.
2. The Trained Tensor Flow model may get very large, making the app size large.
3. The in-built camera may not be able to generate good quality pictures.

## User Documentation

1. Help and How-to-Use section will be provided inside the application with screenshots
2. Demo and Hands-On will be uploaded to YouTube.

## Assumptions and Dependencies

1. YOLO deep Learning Model
2. Tensor Flow Deep Learning Platform
3. Google Translate Api

# External Interface Requirements

## User Interfaces



Tiger

Fig-1

## Hardware Interfaces

1. Supported Devices include android based Smartphone running Android Lollipop or Later, which has an in built camera.
2. The device must contain at least 200MB of free space
3. The device must support internet connectivity either with Wi-Fi or Mobile Data.

## Software Interfaces

1. YOLO deepNet v3 object detection Model available with open source license which will be fed with images of processed objects that requires translation.
2. OpenCV v1 image processing library for java which will take input the raw image of the object and then scale it and process it.
3. Android Studio v3.5 for developing the android application

## Communications Interfaces

1. GET request to google translate api using http protocol to get translated texts.
2. POST request to cloud server to upload errors and corrections.

# System Features

## Scan For Objects

4.1.1 Description and Priority

The System looks for objects to feed to the deepNet. It is the very first step having highest priority.

4.1.2 Stimulus/Response Sequences

The User opens the application, Camera feed starts. The System looks for objects from the camera feed.

4.1.3 Functional Requirements

The application when started should call the inbuilt camera api to start the real time video feed displayed in the orientation on how the device is oriented with respect to the user. The video should be of maximum resolution possible by the device. The valid objects which can be classified should be enclosed by a bounding box augmented on the video feed.

REQ-1:Camera Feed

REQ-2:Bounding Box on Objects Detected

## Process The Objects

4.1.1 Description and Priority

The Image of the Object is sent to the deepNet Residing inside the application which looks for features and gives out the classification with a probability value. It is if of High priority.

4.1.2 Stimulus/Response Sequences

The User touch on a bounding box, a progress dialog starts. The System Processes the image and the application proceeds to another screen containing different languages.

4.1.3 Functional Requirements

The image should be properly resized and processed before sending to the deepNet. The deepNet should be able to process the image in least amount of time so that the application doesn’t seem to freeze. Upon completion a trigger is sent for the application to switch to different activity. YOLO tiny deepNet Model available open Source needs to be used .Image Processing through OpenCV library.

REQ-1:DeepNet for object detection and classification

REQ-2:Image Processing

REQ-3:Trigger

## Translation of The Image

4.1.1 Description and Priority

The System generates the object classification in multiple languages using google translate api.It is of medium priority.

4.1.2 Stimulus/Response Sequences

After the activity is switched the user is provided with a range of languages which on selecting will give out the result in that language along with an option for word pronunciation.

4.1.3 Functional Requirements

The language list should contain only those languages for which proper translation in available with google translate. The google translate api needs to be called for both language translation and voice for which internet connectivity is needed.

REQ-1:Internet Connectivity

REQ-2:Google Translate api

# Other Nonfunctional Requirements

## Performance Requirements

1. Ambient light conditions for the camera
2. Good contrast of the image of the object with the surroundings

## Safety Requirements

1. Correctness Check before accepting error corrections
2. Discarding Use of Images that may be termed Offensive in some Parts of World

## Security Requirements

1. EULA for storing of classified images, as it may hamper user’s privacy.

## Software Quality Attributes

1. The Deep Learning Model that will be used needs to be properly trained and only after satisfying the quality level of over 60% should be used inside the application.
2. The UI for the application needs to designed following the material design guidelines and conventional UX design Principles.

Appendix A: Glossary

**E.x:**

|  |  |
| --- | --- |
| **Term** | **Definition** |
| Application | The final deliverable ,an android application |
| Author | Person submitting an article to be reviewed. In case of multiple authors, this term refers to the *principal author*, with whom all communication is made. |
| Computer Vision | Branch of Artificial Intelligence dealing with intelligent processing of images and real time video feeds. |
| DeepNet | The deep learning model used for detecting and classification of objects in a image. |
| EULA | End user license agreement , a piece of agreement between the user and the application publisher. |
| Software Requirements Specification | A document that completely describes all of the functions of a proposed system and the constraints under which it must operate. For example, this document. |
| Stakeholder | Any person with an interest in the project who is not a developer. |
| User | The user of the Application. |

Table-3

Appendix B: Analysis Models

Backend

DeepNet Object Classification Model

Application

Camera Feed

Translation API

User

Fig- 2

Conclusion

In this SRS Document about the Requirements to develop the OCMLS (Object Classification in Multiple Language System).We studied various functional and non-functional requirements. We also studied various use cases and UI diagrams along with UML Diagrams. We studied in detail about the Project and its external interface requirements.

**References**

1. <https://link.springer.com/chapter/10.1007/978-3-642-36124-1_23> Research on Computer Vision-Based Object Detection and Classification
2. Research on Computer Vision-Based Object Detection and Classification by-Juan Wu , Bo Peng , Zhenxiang Huang , Jietao Xie

**TABLE OF CONTENTS**

**ABSTRACT ii**

**ACKNOWLEDGEMENTS iii**

**DECLARATION BY THE CANDIDATES iv**

**SUPERVISOR’S CERTIFICATE v**

**LIST OF TABLE vi**

**LIST OF FIGURES vii**

**LIST OF ABBREVIATIONS viii**

**Chapter 1. Introduction 3**

1.1 Purpose 3

1.2 Document Conventions 3

1.3 Intended Audience and Reading Suggestions 3

1.4 Product Scope 3

1.5 References 3

**Chapter 2. Overall Description 3**

2.1 Product Perspective 3

2.2 Product Functions 3

2.3 User Classes and Characteristics 4

2.4 Operating Environment 4

2.5 Design and Implementation Constraints 4

2.6 User Documentation 4

2.7 Assumptions and Dependencies 4

**Chapter 3. External Interface Requirements 5**

3.1 User Interfaces 5

3.2 Hardware Interfaces 5

3.3 Software Interfaces 5

3.4 Communications Interfaces 5

**Chapter 4. System Features 6**

4.1 Scan for Objects 6

4.2 Process the Objects 6

4.3 Translation of The Image 6

**Chapter 5. Other Nonfunctional Requirements 7**

5.1 Performance Requirements 7

5.2 Safety Requirements 7

5.3 Security Requirements 7

5.4 Software Quality Attributes 7

**Chapter 6. Conclusion 10**

**REFERENCES 11** 28