VISVESVARAYA TECHNOLOGICAL UNIVERSITY
JNANA SANGAMA, BELAGAVI

Project Work Phase-2 report on

Object Detection and Audio Generation Using Generative AI

Submitted in partial fulfillment of the Requirements of the 2nd Year in

MASTER OF COMPUTER APPLICATIONS

By

BINDU SHREE Y R

1RF23MC019

Under the Guidance of

Ms.K.Sangeetha
Assistant Professor
Department of MCA

RV Institute of Technology and Management®
Sy. No. 171/5, Kothanur, 8th Phase, JP Nagar, Bengaluru - 560076, Karnataka, India

2024-2025

RV Institute of Technology and Management®
Sy. No. 171/5, Kothanur, 8th Phase, JP Nagar,
Bengaluru - 560076, Karnataka, India

DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS

CERTIFICATE

This is to Certify that **Bindu Shree Y R** bearing USN: 1RF23MC019, student of **4th semester** MCA has satisfactorily completed the **Project Work Phase-2 (22MCA44)** on **"Object Detection and Audio Generation Using Generative AI"** in the academic year **2024-2025** as prescribed by VTU for IV Semester of Master of Computer Applications.

Internal Guide
Ms.K.Sangeetha
Assistant Professor
Department of MCA
RV Institute of Technology and Management
Bengaluru-76

Head of the Department
RV Institute of Technology and Management
Bengaluru-76
Head of the Department
Dept. of Master of Computer Applications
RV Institute of Technology and Management
Bengaluru - 560 076

Principal
RV Institute of Technology and Management
Bengaluru-76
Principal
RV Institute of Technology & Management

Examiner 1

Examiner 2

Signature:

DECLARATION

I, **BINDU SHREE Y R (USN:1RF23MC019)** student of **4th SEM**, Department of MCA, RV Institute of Technology and Management, Bengaluru–76, declare that the **Project work Phase-2** on **"Object Detection and Audio Generation Using Generative AI"** is a record of the original work done by me under the guidance and supervision of **Ms.K.Sangeetha, Assistant Professor**, Department of MCA, RVITM and this project work has not formed the basis of any Degree/Diploma/Fellowship or similar title to any candidate of any university.

Name: BINDU SHREE Y R
USN:1RF23MC019

Bindu Shree Y R
Signature of the candidate

ACKNOWLEDGEMENT

I would like to place on record, my gratitude to all those people who have helped me in making this **Project work Phase-2** a reality. Our Institution has played a prominent role in guiding in the right direction.

I express my sincere words of gratitude to the **Management of RV Institute of Technology and Management** for providing such a healthy environment for the successful completion of this project work.

I would like to thank our beloved Principal, **Dr. Nagashettappa Biradar**, for providing the necessary facilities to carry out this work.

I am extremely grateful to our beloved HoD, **Dr. M. Mrunalini,** Professor and Head, Department of MCA, RVITM for having accepted to patronize me in the right direction with all her wisdom.

I would also express my heartfelt thanks to my project guide **Ms.K.Sangeetha,** Assistant Professor, Department of MCA, RVITM for her continuous guidance and valuable suggestions for this project work.

I also express my heartfelt thanks to all the teaching and nonteaching staff members of MCA Department for their encouragement and support throughout this work.

BINDU SHREE Y R

1RF23MC019

TABLE OF CONTENTS

Page No.

1. INTRODUCTION ……...01

1.1 Aim of the Project Work……...01

1.2 Project Description……....01

1.3 Objectives of the Project……..…02

1.4 Scope and Limitations of the Project……...…03

2. LITERATURE SURVEY……05

2.1 Existing and Proposed System…….05

2.2 Feasibility Study……....06

2.3 Tools and Technologies Used……..…07

2.4 Hardware and Software Requirements……..07

3. SOFTWARE REQUIREMENTS SPECIFICATIONS…….....10

3.1 Users…….....10

3.2 Functional Requirements…….10

3.3 Non-Functional Requirements……11

4. SYSTEM DESIGN……....13

4.1 System Architecture…….13

4.2 System Perspective……....15

4.3 Context Diagram…….....16

5. DETAILEDDESIGN…….....18

5.1 Dataflow Diagram……18

5.2 Activity Diagram……...20

5.3 Use Case Diagram……21

5.4 Sequence Diagram……22

6. IMPLEMENTATION……...24

6.1 Coding Standard……....24

6.2 Snippet Code……24

6.3 Screenshot……26

6.4 Result and Discussion……...30

7. SOFTWARE TESTING…….....32

7.1 Unit Testing……...32

7.2 Automation Testing……..……32

7.3 Test Cases……..……...33

8. CONCLUSION……..……...34

9. FUTURE ENHANCEMENTS…….35

10. BIBLIOGRAPHY…….36

LIST OF FIGURES

Figure No.	Caption	Page No.
Figure 4.1	Layered System Design of the Image-to-Audio Description Application	13
Figure 4.2	Layered Architecture — Client, Middleware, Server, and Cloud	15
Figure 4.3	System Architecture of Image-to-Audio Description Web Application	16
Figure 5.1	Context Level DFD for Image to Audio System	18
Figure 5.2	Level-1 DFD-Detailed Process flow	19
Figure 5.3	Activity Diagram for Image-to-Audio System	20
Figure 5.4	Use Case Diagram for Image-to-Audio System	22
Figure 5.5	Sequence Diagram for Image-to-Audio System	23
Figure 6.1	Homepage design	26
Figure 6.2	Uploading image	26
Figure 6.3	Uploaded image is displaying	27
Figure 6.4	Audio Description Button visible	27
Figure 6.3	Generating image caption	28
Figure 6.6	Generated caption by BLIP	28
Figure 6.7	Refined Caption by Groq	29
Figure 6.8	Audio generating	29
Figure 6.9	Geneated Audio	30

LIST OF TABLES

Table No	Description	Page No
Table 6.1	Sample Results of the Image to Audio Description System	30
Table 7.1	Test Cases	33