

REPORT ON PROJECT STAGE - I

**AI/ML BASED ROBOTICS - IMPLEMENTATION FOR VISION,
COMMUNICATION, AND ADVANCED MECHANICS**

SUBMITTED TO SAVITRIBAI PHULE PUNE UNIVERSITY
FOR PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF ENGINEERING
In
Electronics and Telecommunication Engineering

By
Vedant Dinesh Butala (42112)
Sumedh Ravindra Kasat (42133)

GUIDED BY :
PROF. GRISIH SHRIKISAN MUNDADA

DEPARTMENT OF ELECTRONICS AND TELECOMMUNICATION ENGINEERING
PUNE INSTITUTE OF COMPUTER TECHNOLOGY
PUNE – 43

OCTOBER 2024

Department of Electronics and Telecommunication Engineering
Pune Institute of Computer Technology, Pune – 43

CERTIFICATE

This is to certify that the Project Stage - I Report entitled
“AI/ML-Based Robotics - Implementation for Vision, Communication, and Advanced Mechanics”
has been successfully completed by
Vedant Dinesh Butala (42112)
Sumedh Ravindra Kasat (42133)

towards the partial fulfillment of the degree of **Bachelor of Engineering in Electronics and Telecommunication Engineering** as awarded by the **Savitribai Phule Pune University**, at **Pune Institute of Computer Technology** during the academic year 2024-25

INTERNAL GUIDE
(Prof. Girish S. Mundada)

HOED
(Dr. M. V. Munot)

ACKNOWLEDGEMENT

We would like to express our heartfelt gratitude to Prof. Girish Shrikisan Mundada, our esteemed guide, for his unwavering support, expert advice, and continuous encouragement throughout the development of this project. His profound knowledge in the field of electronics and telecommunications, combined with his patient mentorship, has been invaluable in steering our project toward successful completion. His guidance was instrumental in helping us overcome challenges and refine our approach, ensuring that we achieved our objectives.

We also extend our sincere thanks to the Pune Institute of Computer Technology (PICT) for fostering an environment of innovation and learning. The resources and facilities provided by the institute enabled us to conduct thorough research and experimentation. We are particularly grateful to the faculty and staff of the Department of Electronics and Telecommunication for their constant support and encouragement.

Thanking You,

Vedant Dinesh Butala (42112)
Sumedh Ravindra Kasat (42133)

CONTENTS

	Abstract	i
	List of Acronyms	ii
	List of Symbols	iii
	List of Figures	iv
	List of Tables	v
1	Introduction	1 - 7
	1.1 Background	1
	1.2 Relevance	2
	1.3 Motivation	3
	1.4 Problem Definition	3
	1.5 Scope and Objectives	4
	1.6 Technical Approach	4
	1.7 Organization of the Report	6
2	Literature Survey	8 - 16
	2.1 Introduction	8
	2.2 Large-Scale and Vision-Based Navigation for Autonomous Robots	11
	2.3 SLAM and Hybrid Navigation Techniques for Autonomous Robots	13
	2.4 Summary of Findings and Relevance	14
	2.5 Conclusion of Literature Survey	16
3	Methodology	17 - 27
	3.1 Introduction	17
	3.2 System Design	18
	3.3 Hardware Setup	19
	3.4 Software Development and Testing	22
	3.5 Datasets used in this research work	23
4	Results and Discussions	28 - 36
	4.1 Improved Navigation, Localization, and Communication for Autonomous Robots	28
	4.2 Optimized Control, Adaptability, and System Efficiency for Autonomous Robots	29
	4.3 Conclusion of Results and Discussions	30
5	Conclusions & Future Scope	37
	References	39 - 41
	Plagiarism Report	42

ABSTRACT

This project presents the design and implementation of an AI/ML-driven autonomous robotic system that integrates computer vision, intelligent communication, and adaptive mechanical control. The primary objective is to enable real-time navigation, decision-making, and interaction capabilities in dynamic and unstructured environments, thereby reducing reliance on manual control and enhancing operational efficiency.

The system architecture employs Convolutional Neural Networks (CNNs) for object detection and classification, along with Simultaneous Localization and Mapping (SLAM) for real-time environment mapping and localization. Natural Language Processing techniques are incorporated to facilitate voice-based human-robot interaction, while Reinforcement Learning (RL) is applied to optimize path planning and motion control. These AI models are interfaced with a robust hardware stack comprising LiDAR sensors, IMUs, STM32 microcontrollers, and power-regulated PCBs, orchestrated through the ROS2 middleware to ensure low-latency, modular, and scalable operation.

Extensive testing in indoor environments demonstrates high object recognition accuracy, improved localization precision, low-latency response to dynamic obstacles, and seamless human-robot communication. These outcomes validate the feasibility and efficiency of the proposed system for use in applications such as warehouse automation, smart surveillance, healthcare robotics, and collaborative industrial systems.

By merging the strengths of AI and embedded robotics, this project underscores the transformative potential of intelligent autonomous systems in shaping the future of human-centric and adaptive robotic solutions.

Abbreviations and Acronyms

AI	Artificial Intelligence
ML	Machine Learning
RL	Reinforcement Learning
CNN	Convolutional Neural Networks
ROS	Robot Operating System
SLAM	Simultaneous Localization and Mapping
RFID	Radio-Frequency IDentification
PID	Proportional, Integration, and Derivation

List of Symbols

% Percentage
\$ Dollar

List of Figures

Fig. 1	Proposed Block Diagram	19
Fig. 2	Connecting to the Controller	20
Fig. 3	Prototype Hardware	20
Fig. 4	2D Lidar Generated Map	32
Fig. 5	Commands Running on Ubuntu 22.04	33
Fig. 6	RQT Graph	33
Fig. 7	RViz2 Simulation Software	34
Fig. 8	3D Simulation on Gazebo	34
Fig. 9	SLAM Compatibility for ROS	35
Fig. 10	Hardware Compatibility for ROS	35
Fig. 11	Performance Parameter Evaluation Chart	36

List of Tables

1	Comprehensive Summary Table of Research Papers	9
2	Datasets for Computer Vision	24
3	Dataset for Simultaneous Localization and Mapping (SLAM)	26