NOVA: An Edge AI Autonomous Robotic Assistant for Intelligent Construction Site Inspection

Shreem Dixit, Undergraduate Researcher, Department of Computer Science
Dr. Gongfan Chen, Assistant Professor, Albert School of Construction

UNIVERSITY OF NORTH CAROLINA CHARLOTTE

Introduction

Sites never stop changing — NOVA is built to keep watch.

- Construction sites are large, dynamic environments that change throughout a project's lifecycle.
- Routine inspections are essential for tracking progress and identifying potential issues early.
- Manual inspections can be time-consuming and difficult to perform consistently.
- Some areas may be inconvenient or hazardous for workers to access repeatedly.

An intelligent robot could help address these challenges by observing, documenting, and reporting on changing site conditions.

*This project introduces **NOVA**, a robotic prototype built to bring these capabilities to real-world construction sites.*

Objectives

Main objective

To develop NOVA, an intelligent mobile robot that can assist with construction-site inspections through simple spoken commands.

Key objectives

- Travel to selected inspection zones.
- Capture images or video of the surrounding area.
- Describe site conditions in clear, everyday language.
- Reduce the need for repeated manual inspection.
- Test tasks safely in simulation before transferring them to the physical robot.

The goal is not to replace construction professionals, but to provide them with an additional tool for safer and more consistent site observation.

Meet NOVA

What is NOVA

NOVA is a prototype mobile robot designed to assist with routine construction-site observation. Instead of requiring complex controls, users can interact with the robot using simple spoken instructions.

NOVA at a Glance

Feature	Purpose
Voice interaction	Lets users give simple spoken commands
Mobile movement	Travels to selected inspection areas
Camera-based observation	Captures photos or videos of site conditions
Clear reporting	Explains what it observes in everyday language
Local operation	Can work even when internet access is limited

The NOVA prototype combines mobility, visual observation, and voice interaction in one platform.

How NOVA Works

1. LISTEN
The user gives NOVA a spoken instruction. *Example: "NOVA, go to Zone 2 and take a picture."(Automatic Speech Recognition from EMEET 360 degree microphone)*
2. UNDERSTAND
NOVA identifies the requested location and task. (Using AI Model – Moondream)
3. MOVE
The robot travels to the selected inspection zone. (Virtual environment – Gazebo, using sensors of GPU LiDAR)
4. OBSERVE
NOVA captures images or video of the surrounding area. (Using the OpenCV Camera feature + Gazebo Camera Sensor)
5. REPORT
The robot reviews what it sees and provides a clear summary. (Gemma organizes findings + Moondream analyzes the image)

Progress and Early Results

Capabilities Demonstrated:

- The robot can navigate through its environment while avoiding obstacles along its path.
- Users can give combined instructions such as: "NOVA, go to Zone 1 and take a picture."
- After reaching the selected zone, NOVA can automatically capture an image or record video.
- The collected visual data is sent to Moondream, an AI vision model that examines the scene.
- Moondream generates a detailed description of visible objects, activities, surroundings, and possible areas of concern.
- The robot can complete the full sequence—from receiving a command to reporting what it observed—with minimal manual control.

Simple Workflow

Voice Command → Understand Task → Move to Zone → Capture the Area → Provide a Summary

```
Audio captured.
Understanding speech...
You said: "what is bashing for nova take a picture"
Nova command: "take a picture"

Nova: Taking a picture now.
Searching for camera...
Camera found at index 0
Camera warming up...
Archived picture saved to: /home/shreem/robot/captured_images
e_2026-07-08_12-41-24.jpg
Latest picture saved to: /home/shreem/robot/robot_view.jpg
Reading captured image...
Moondream is analyzing the new picture...

========== NEW IMAGE DESCRIPTION ==========
A woman with long brown hair wearing a necklace stands agains
te wall, her gaze directed to the left side of the image. The
eatures a red stripe running along the top edge of the wall,
s illuminated by red lights and appears to be part of an indo
ting system.
==========================================
Analysis time: 89.6 seconds

Nova: The picture has been captured and described.
Nova: I am ready for another command.
```

**Successful voice-command test showing NOVA capturing an image, analyzing it with Moondream, and generating a detailed scene description.*

Applications and Future Plans

Applications and Limitations

Track Progress: Compare new and previous images; accuracy depends on image quality.

Improve Safety: Identify visible hazards; does not replace human inspection.

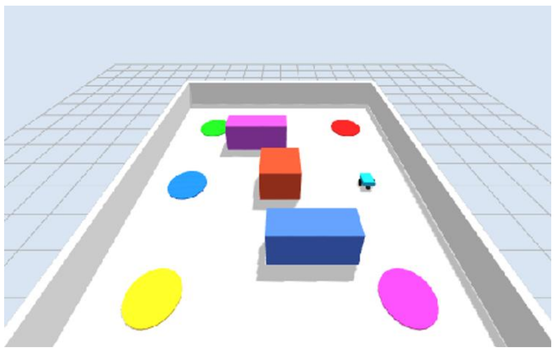
Detect Faults: Highlight visible cracks or damage; hidden defects may be missed.

Future Plans

Connect NOVA with more additional devices for hands-free control.

Provide deeper image analysis instead of only describing the scene.

Generate clear reports showing what changed, what is wrong, and what needs attention.



**Gazebo simulation environment showing NOVA navigating between predefined inspection zones and obstacles.*