GAN Assisted Map Reconstruction for first responders using Sensor Fusion

Christian Jeardoe, Quinn Leboeuf,

Kyle Smejkal, Rufus Tadpatri

**Concept of Operations**

Revision - 1

15 September 2024

Concept of Operations

for

GAN Assisted Map Reconstruction for first responders using Sensor Fusion

TEAM <27>

APPROVED BY:

*Rufus Tadpatri 09/15/2024*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Project Leader Date

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Dr. Lusher Date

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

T/A Date

**CHANGE RECORD**

| **Rev.** | **Date** | **Originator** | **Approvals** | **Descriptions** |
| --- | --- | --- | --- | --- |
| **1** | **09/15/2024** | **GAN Assisted Map Reconstruction** |  | **Draft Release** |

**Table of Contents**

[**1. Executive Summary 6**](#_fo8crz2a5r4p)

[**2. Introduction 6**](#_ay0re09lce0j)

[2.1 Purpose 6](#_ry3r8nluj3ep)

[2.2 Overview 7](#_la17ffvpzhdb)

[2.3 Referenced Documents and Citations 8](#_j72ooiyjksn)

[**3. Operating Concept 8**](#_tlpsuknj755b)

[3.1 Scope 8](#_fjuiqjnb9ehm)

[3.2 Operational Description 8](#_xzjflwd1uwpw)

[3.3 System Description 9](#_iemr8onfn589)

[3.3.1 RGB 9](#_am0hkju883d5)

[3.3.2 Infrared 9](#_v2ysl1pmtfzl)

[3.3.3 Realsense/Kinect 9](#_qbkdug1u09is)

[3.3.4 Ultrasonic 9](#_eed0zzi11o0v)

[3.4 Modes of Operations 10](#_5k29n1antmx7)

[3.4.1 Reconnaissance 10](#_jhhb9zkq1dni)

[3.5 Users 10](#_e04oqhmr59iy)

[**4. Scenarios 10**](#_xguqyauxz45d)

[4.1. Search and Rescue 10](#_y4ie509hqq3l)

[4.2. Cave Exploration 10](#_rq6bx3ds232u)

[4.3. Sea-floor Mapping 11](#_q4hh61lwrguh)

[4.4. Healthcare 11](#_mm9348bvkllc)

[**5. Analysis 11**](#_y6hfln55dfv4)

[5.1 Summary of Proposed Improvements 11](#_l605py9jy28j)

[5.2 Constraints 11](#_e62vs8lrxsha)

[5.3 Impact 11](#_wknyr1xgxgxg)

[5.4 Risks and Mitigation 12](#_j6bw5t5pqvmp)

[5.5 Alternatives 12](#_43kl3i72i8yr)

[**Conclusion 12**](#_fjeri5w5np7c)

**List of Tables**

Table 1: List of Design Constraints 12

**List of Figures**

Figure 1: System Operations Overview 8

Figure 2: Example of Digital 3D Reconstruction of an Object. 9

# 1. Executive Summary

Emergency situations commonly occur spontaneously with little to no knowledge provided to rescuers and/or first responders. Fires, chemical incidents, or dangerous shootings are relatively common emergencies where these first responders must put themselves in danger to get an assessment of the environment in order to carry out important tasks such as retrieving someone from a dangerous area. To help aid in these types of situations and reduce physical human involvement during dangerous scenarios, we aim to give first responders the capability to receive a digital map of any enclosed environment safely.

With AI technology rapidly growing, we can use generative AI to help us create digitally mapped environments based on sensor data input. Multiple sensors will work in unison to provide very precise data for the software to build a 3D digital map. Using multiple sensors also allows the device to create a mapped environment even if visibility, heat, or other environmental factors create technical issues. A mobile robotic machine will be used to transport the sensors into wanted environments to scan data, which can then be retrieved after completion for analysis.

# 2. Introduction

This document explains the future operation for GAN-assisted map reconstruction using sensor fusion. The system will utilize RGB, infrared (IR), Realsense/Kinect, and ultrasonic sensor data. Each subsystem will produce individual point clouds, which will be integrated into a cohesive 3D map during the project's second stage. The final goal is a fully integrated 3D model of an enclosed environment

## 2.1 Purpose

The goal is to develop a system that generates 3D point clouds from various sensor inputs using Generative Adversarial Networks (GANs). By utilizing the data from RGB, infrared, Realsense/Kinect, and ultrasonic sensors, the system will create accurate 3D reconstructions of environments, which can be used in robotics, surveillance, and reconnaissance applications. The project aims to overcome the challenges of sparse, noisy, or low-resolution sensor data by applying multiple types of data in collaboration with generative AI. Ultimately, we aim to integrate point cloud-type data from different sensors into a unified 3D model. The system will be designed so it can be utilized in multiple scenarios with various technical systems.

## 2.2 Overview

For our provided goals, our system consists of 4 subsystems which are RGB, infrared, Realsense/Kinect, and ultrasonic sensors, which each will collect and provide unique data for the system. For each sensor type, GAN models will be developed to convert raw sensor data into 3D point clouds. These individual GAN models will be designed individually for each sensor, in which the generated point clouds from all sensor data will be integrated into a 3D model. This model will be visualized in digital twin environments such as Gazebo or Unity. The system will be designed to be flexible in required technological resources to account for deployment on resource-constrained platforms like a Rpi5 or Jetson Nano. Through this integration of sensor data and advanced machine learning, the project will allow users to see unknown yet accurate 3D environments digitally reconstructed for various purposes.

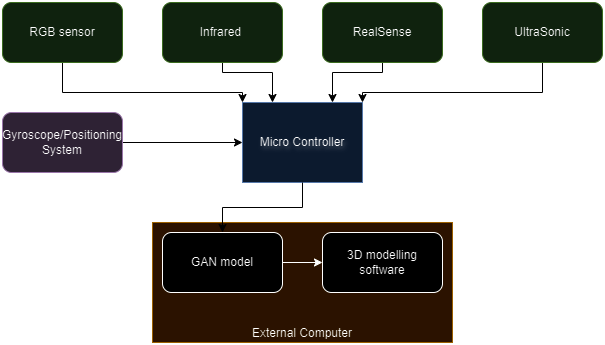


Figure 1: System Operations Overview

## 2.3 Referenced Documents and Citations

1. Figure 2: Wei, Y., Vosselman, G., Yang, M. Y., & University of Twente. (n.d.). Flow-based GAN for 3D Point Cloud generation. In University of Twente [Journal-article]. https://weiyao1996.github.io/files/publications/BMVC\_2022.pdf
2. Versatile and Scalable 3D RGB Point Cloud Generation from 2D Images in Unsupervised Reconstruction.
3. Point Cloud Segmentation Using RGB Drone Imagery.

# 3. Operating Concept

## 3.1 Scope

This ConOps covers the development and integration of GAN models for different sensor types: RGB, Infrared (IR), Realsense/Kinect (RGB-D), and Ultrasonic. It includes details on responsibilities, tasks, deliverables, and the timeline for the project.

## 3.2 Operational Description

Tasks: Combine point clouds from various sensors into a cohesive 3D model. Visualize and optimize the integrated 3D model in digital twin environments (Gazebo/Unity). Figure 2 provides an example that relates to our task goals, but we will be visualizing an environment instead of an object.

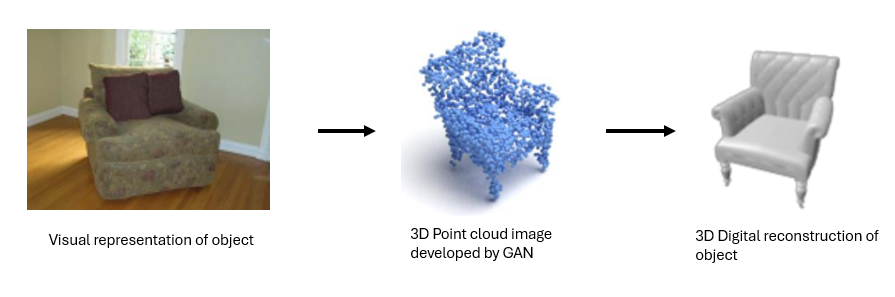


Figure 2: Example of Digital 3D Reconstruction of an Object.

## 3.3 System Description

### 3.3.1 RGB

A GAN model will be developed that converts RGB images to 3D point clouds. RGB data will be preprocessed to optimize input for the GAN. The RGB sensor allows the ability to train the GAN model to convert from RGB to point cloud in order to integrate into the final 3D model.

### 3.3.2 Infrared

The Infrared sensor will capture depth information and thermal imaging of the environment in order to train the GAN model to create a 3D point cloud of said environment. This sensor will be able to accurately capture data in areas that the other sensors may not be able to perform as well, such as low light environments.

### 3.3.3 Realsense/Kinect

The RealSense/Kinect camera plays a pivotal role in capturing depth information, and when combined with its RGB imaging capabilities, it can independently generate a detailed 3D map of a room. This dataset is essential for the GAN model, as it enhances the depth data, enabling the creation of high-resolution point clouds for accurate and refined 3D reconstructions

### 3.3.4 Ultrasonic

The ultrasonic sensor collects data that provides information on its distance between objects using sound waves. While there will be extreme challenges getting detailed information from this type of sensor in complex environments, it can be used extensively to confirm the accuracy of the developed information from other sensors such as a room/objects size, length, etc. ultrasonic sensors do not need visibility, providing a solution to environments covered in smoke or dark.

## 3.4 Modes of Operations

### 3.4.1 Reconnaissance

The only mode of operation for this system is to perform reconnaissance in an emergency situation.

## 3.5 Users

Primary users would be first responders. The intended purpose of this system is to assist in an emergency situation, so the user would not need prior knowledge other than basic robotic control. The user would maneuver the robot around the room, capturing imagery using the sensors of the robot.

# 4. Scenarios

## 4.1. Search and Rescue

The primary scenario for which the robot is being designed is search and rescue missions. These operations often occur in hazardous or inaccessible areas, such as collapsed buildings, dense forests, or areas affected by natural disasters. The GAN robot, equipped with various sensors, can navigate these dangerous environments and generate real-time, high-resolution maps. The robot can use its GAN model to reconstruct partially occluded regions and enhance imaging data, allowing rescue teams to detect and locate victims faster without putting themselves at risk.

## 4.2. Cave Exploration

Exploring cave systems is challenging due to the lack of natural light and complex, often unstable, terrain. The GAN robot can utilize depth sensors, infrared cameras, and LiDAR to generate detailed 3D maps of the cave environment. Its GAN model can enhance low-resolution or incomplete sensor data to create a comprehensive view of the cave, helping explorers or researchers navigate safely while minimizing the risk of getting lost or encountering dangerous areas.

## 4.3. Sea-floor Mapping

Mapping the sea floor is crucial for marine research, underwater construction, and locating shipwrecks. Traditional methods using sonar or cameras can result in noisy or incomplete data due to water conditions. The GAN robot, equipped with underwater sensors, can enhance this data in real time, providing clearer and more accurate images of the sea floor.

## 4.4. Healthcare

While our GAN model is being developed to be used on robots, it is not limited to that use. For example, in healthcare, the GAN model can be used to improve medical imaging techniques. By generating high-resolution images from noisy or low-quality sensor data (such as MRIs, CT scans, or ultrasounds), GANs can help doctors make more accurate diagnoses.

# 5. Analysis

## 5.1 Summary of Proposed Improvements

A robot with these sensors can capture a 3D model of the environment more accurately with sensors backed by machine learning to create the 3D environment, while being able to fill the gap caused by environmental factors for one or more sensors.

## 5.2 Constraints

Many different factors can lead to constraints in the operation of the system, as outlined in the table below.

| **Constraints** | **Reasoning** |
| --- | --- |
| 1) Microcontroller | The GAN model will be running on a microcontroller which limits the computation power for the model. |
| 2) Sensor Resolution | Sensor resolution can greatly affect the quality of the 3D model and different sensors can capture data in different ways. |
| 3) Time | The GAN model will be limited due to the limited training data and training times due to deadlines. |

Table 1: List of Design Constraints

## 5.3 Impact

This technology could have a great impact for mankind, as it would allow us to more accurately map areas before humans need to enter. The best example of this would be for first responders getting a 3D map of an environment and its hazards before people have to go into said environment. Another positive impact would be mapping and charting areas that are unknown or unsafe for humans such as caves or the ocean floor.

## 5.4 Risks and Mitigation

1. **Model Accuracy**: GAN models may not achieve the desired accuracy.
   1. *Mitigation*: Iterative testing and validation; explore alternative architectures.
2. **Integration Issues**: Difficulty in merging point clouds from different sensors.
   1. *Mitigation*: Early and continuous integration testing; effective communication within the team.
3. **Visualization Challenges**: Inaccurate or inefficient visualization of the 3D model.
   1. *Mitigation*: Regular feedback sessions; thorough testing in Gazebo/Unity.

## 5.5 Alternatives

There are other projects that are using sensor fusion and machine learning to create 3D point cloud maps, like Nvidia. However, these projects mainly focus on making a model based on RGB and Lidar data rather than the 4 sensors used here.

# Conclusion

This ConOps provides a comprehensive framework for the successful development and integration of GAN models for generating 3D point clouds from various sensor types. By adhering to this plan, the team will produce a unified 3D model and demonstrate its capabilities in a digital twin environment, contributing valuable insights and tools for 3D reconstruction and visualization.