International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Ai-Powered Autonomous Drone Navigation in Complex Environments

[1]N. V. S. Vijaya Lakshmi, [2]V. Rambabu, [3]Ch S K Chaitanya, [4]V Anil Santhosh

Submitted: 10/05/2023 **Revised:** 18/07/2023 **Accepted:**08/08/2023

Abstract— Autonomous navigation in intricate surroundings is essential in time-critical situations like as disaster response or search and rescue operations. Complex settings provide substantial problems for autonomous platforms to travel owing to their difficult characteristics: limited tight corridors, unstable pathways with trash and impediments, uneven geological formations, and inadequate illumination conditions.

This study presents a multimodal fusion methodology to tackle the challenge of autonomous navigation in intricate landscapes, including collapsed cities and natural caverns. Initially, we replicate intricate landscapes using a physics-based simulation engine and gather an extensive dataset for training purposes.

We present a Navigation Multimodal Fusion Network (NMFNet) with three branches to efficiently process three visual modalities: laser, RGB pictures, and point cloud data.

The comprehensive experimental findings demonstrate that our NMFNet significantly surpasses previous state-of-the-art methods while attaining real-time performance. We furthermore demonstrate that the utilization of several senses is crucial for autonomous navigation in intricate situations. Ultimately, we effectively implement our network on both simulated and actual mobile robots.

Keywords: *NMFNet, experimental, modalities*

INTRODUCTION

Autonomous navigation is a well-established domain within robotics research, required for mobile robots to accomplish a sequence of activities in locations typically navigated by humans daily. The primary objective of autonomous navigation is to direct a robot to traverse its surroundings while avoiding collisions with impediments. Navigation is a fundamental ability for intelligent entities, necessitating decision-making across several temporal and spatial domains. In actuality, autonomous navigation is a complex operation, since the robot must complete the perception-control loop amongst uncertainty to achieve autonomy.

Recently, learning-based methodologies, such as deep learning models, have shown the capability to directly formulate end-to-end policies that translate raw sensor data into control instructions [1], [2]. This comprehensive strategy diminishes implementation complexity and efficiently leverages input data from many sensors (e.g., depth camera, laser), consequently decreasing costs, power consumption, and computing time. Another advantage is that the end-to-end relationship between input data and control outputs can produce an arbitrarily nonlinear complex model (i.e., from sensor to actuation), which has yielded unexpectedly positive results in various control challenges, including lane following, autonomous driving, and Unmanned Aerial Vehicle (UAV) control. Nevertheless,

1,2,3,4International School Of Technology And Sciences For Women, A.P, India.

METHODOLOGY Inspired by the recent advancements in autonomous driving [28], [29], [33], our objective is to develop a framework that directly correlates the input sensory data X = (D; P; I) with the output steering directives Y. Consequently, we develop NMFNet with three branches to address three visual modalities.

A. Two-Dimensional Features

Extracting significant characteristics from 2D pictures is crucial for success in several vision tasks. This study use ResNet8 to extract deep features from the input RGB picture and laser distance map. ResNet8 comprises three residual blocks, with each block including a convolutional layer, ReLU activation, skip connections, and batch normalization procedures. A comprehensive illustration of the ResNet8 architecture is shown in Fig. 3.

Similar to [29], we use ResNet8 to extract deep features from the 2D pictures due to its lightweight design, which delivers competitive performance while being resilient to disappearing and expanding gradient issues during training.

RELATED WORK

Sensor fusion for autonomous robot navigation is a prominent study area in robotics [9]. Conventional approaches address this issue with algorithms derived from the Kalman Filter [10]. This technique enables the integration of data from several sensors and sensor types, including visual, inertial, GPS, and pressure sensors. Lynen et al. [11] developed a methodology using the Extended Kalman Filter (EKF) for the navigation of Micro Aerial Vehicles (MAVs). In [12], the authors devised a method using the Extended Kalman Filter (EKF) to assess the status of an unmanned aerial vehicle (UAV) in diverse situations in real-time.

Mascaro et al. [13] introduced a graph-optimization technique for integrating input from several sensors to estimate UAV posture.

In addition to the conventional localization and navigation tasks, multimodal fusion is used in various applications, including object identification [14] and semantic segmentation [15], [16] in dynamic settings. In both [14] and [16], multimodal data from visual sensors are integrated and analyzed inside a deep learning framework to address difficult lighting circumstances.

Recently, several approaches have been introduced to directly derive control rules from unprocessed sensory input. The approaches may be classified into two primary categories: reinforcement learning and supervised learning.

With the advent of deep learning, Convolutional Neural Networks (CNN) have been extensively used to develop end-to-end perception systems [21]–[26]. In [27], Bojarski et al. introduced the first end-to-end navigation system for autonomous vehicles using 2D pictures. Smolyanskiy et al. [28] expanded this concept for aerial robots using three cameras as input. Likewise, the designers of DroNet [29] used CNN to ascertain the steering angle and forecast the collision probability based on the RGB picture input. Gandhi et al. [30] proposed a navigation technique for UAVs by analyzing both unsuccessful and successful crash attempts. In [31] [32], a combination of CNN and Variational Autoencoder was used to estimate the steering control signal. Monajjemi et al. [5] introduced an innovative approach for agile UAV control.

Recently, the authors in [33] suggested integrating the navigation map with visual input to derive a deterministic control signal.

Reinforcement learning algorithms have been extensively used to derive general rules from robotic experiences [17], [34], [35]. In [36], the authors presented a continuous control system using deep reinforcement learning. Zhu et al. [37] tackled the target-driven navigation issue using an input image of a target item. Wortsman et al. [38] presented a self-adaptive visual navigation system using meta-learning.

The authors in [39] used semantic information and geographical linkages to enable a robot to travel to target items. An end-to-end regression system for UAV racing in simulation was presented in [40]. The authors in [41] [42] suggested training the reinforcement policy in simulated settings and then transferring the acquired policy to real-world applications. The authors integrated deep reinforcement learning with convolutional neural networks to capitalize on the benefits of both methodologies.

Piotr et al. [7] introduced a methodology of augmented memory to train autonomous agents for navigation in extensive and visually intricate settings (complex 3D mazes).

Although reinforcement learning techniques provide general control policies with robust mathematical frameworks, they need many trial-and-error experiments, which are perilous and impractical in actual safety-critical robotic systems. Conversely, supervised learning techniques use pre-existing data to acquire control rules.

Supervision data may be acquired via actual human expert trajectories [29], [30] or conventional controllers [45].

This process is labor-intensive and expensive, although feasible with actual robots. Consequently, the supervised learning strategy is often preferred over the reinforcement learning method when using actual robotic systems. Nonetheless, managing the domain transition between expert advice and actual robot paths in supervised learning approaches is not straightforward.

In this work, we choose the end-to-end supervised learning approach for the ease of deploying and testing in real robot systems. We first simulate the complex environments in physics-based simulation engine and collect a large-scale for supervised learning. We then proposed NMFNet, an effective deep learning framework to fuse visual input and allow the robots to navigate autonomously in complex environments.

Artificial Intelligence Applications Of Drones

Drones in military applications

Drones will significantly augment military capabilities worldwide and will persist in altering the nature of warfare in various ways that affect ground forces. There are substantial and warranted concerns that command and control drones offer critical intelligence on enemy development zones and primary targets. Drone technology enable commanders to make more informed judgments and operate more efficiently in the field due to these insights. While military defense temporary personnel are integrating drones with cutting-edge computer vision and image recognition, relatively young military engineers have swiftly begun to merge drones with artificial intelligence to develop systems that occasionally rival human observation teams in addressing military challenges and advancements.

The MQ-9 Reaper is a military drone currently operational, capable of a range of about 1852 kilometers. The length is around 36 feet, with an elevation reaching 50,000 feet. It can execute endurance and significant altitude, providing surveillance and an aerial assault. The drones support armed forces by facilitating data collection and enabling strategic operations. Drones in these models exemplify state-of-the-art innovation, rendering them unparalleled assets in contemporary military strategy and facilitating significant advancements in modern warfare. Businesses like Shield AI, AeroVironment, and Lockheed Martin illustrate that Shield AI drones can navigate uncharted areas without GPS, allowing military personnel to collect information that fosters rapid advancement and influences operations through drone utilization. They can ascertain whether they are being followed during mapping efforts. tactical surveillance and engagement evaluations Administrators can make decisions without the concern of being constrained by machine vision navigational aids and mechanical components, which collectively generate misleading insights. The AI must undergo a structured learning process, which I believe will be more intriguing.

Drones in Obstacle detection and avoidance

A significant use of AI in drone technology is obstacle identification and avoidance. Unlike traditional drones that rely on human navigation to avoid obstacles, AI-powered drones are designed to operate autonomously, even in intricate and unpredictable environments. Equipped with contemporary sensors and computer vision frameworks, these advanced drones can identify obstacles in their immediate surroundings, such as buildings, trees, or other aerial objects, and autonomously navigate around them. AI computations serve a fundamental role by analyzing sensor data in real-time, enabling drones to Identify possible hazards and generate alternative flight paths that avoid collisions while keeping aligned with mission objectives. This autonomous navigation capacity is crucial in high-stakes circumstances, such as reconnaissance and rescue operations, when drones must traverse intricate urban environments or dense natural settings to locate and assist those in need. In these instances, AI-driven obstacle avoidance enhances both the security and efficiency of operations, enabling drones to access remote areas swiftly and securely without necessitating continuous human oversight. By integrating AI-driven obstacle detection and route optimization functionalities, drones can execute intricate tasks with enhanced reliability, becoming them an essential tool for mission-critical applications across several industries. This innovation not only expands the operating spectrum of drones but also underscores the critical role AI plays in unlocking new possibilities for autonomous systems.

Drones in agriculture

The use of AI in drones has significantly revolutionized agriculture by enhancing efficiency, sustainability, and accuracy in agricultural practices. Drones equipped with modern sensors and AI technologies provide a comprehensive overview of agricultural areas, enabling farmers to monitor crops in real-time and make data-driven choices to enhance output.

This application is directly linked to crop health monitoring, wherein drones equipped with multispectral and hyperspectral cameras capture high-resolution images of crops. Subsequently, AI algorithms analyze these images to identify indicators of stress, such as diseases, pests, and nutrient deficiencies, thereby allowing farmers to intervene promptly. Drones have the potential to transform every aspect of agriculture, from monitoring climatic conditions and humidity levels to determining optimal crop growth timing. Drones may also assess soil conditions, determining the required quantity of fertilizer, among other factors. It is possible to identify damaged plants and, theoretically, to grow trees by dispersing seed pods of the species into the soil at regular intervals, with expenditures decreasing from USD 2.8 billion in 2018 as they become more affordable and adaptable. Similarly, the construction industry was anticipated to be the second-largest economic market for drones in 2020, behind the agricultural sector.

Drones in autonomous navigation

AI-driven autonomous navigation in drones signifies a significant technical progress, offering substantial adaptability and flexibility for diverse applications. In contrast to conventional pre-programmed fixed-wing drones, AI-driven drones may autonomously devise and modify paths based on real-time conditions to adapt to changing settings. The capacity for real-time decision-making is facilitated by sophisticated algorithms in AI that evaluate data from onboard sensors in the drone, enabling the optimization of flight routes in response to weather conditions, obstructions, or changes in mission parameters. Drones equipped with AI can autonomously traverse agricultural fields, continually improving their flight paths by taking into account wind conditions, crop orientation, and topography in real-time. This will facilitate tasks such as crop inspection, land surveying, and the application of fertilizers and pesticides with little human interaction. In search and rescue missions, AI-equipped drones will navigate complex and isolated terrains, adjusting their trajectories to avoid obstacles and enhance the likelihood of locating missing individuals.

Safety is enhanced as AI enables drones to detect and avoid obstructions during flight, preventing collisions. Utilizing autonomous navigation, such devices can perform inspections of infrastructure and facilitate deliveries with little human interaction in intricate areas. This is due to the constant processing of environmental data and the adjustment of flight paths to align with environmental conditions, resulting in activities being performed with greater efficiency and precision.

Radar positioning and returning home

Marine radar systems, mostly using X-band (9 GHz) and optionally S-band (3 GHz) radars, are essential for navigation and collision avoidance, particularly in low-visibility situations. To improve the location capabilities of these radars, many approaches have been devised, including E-RACON positioning, image matching, radar conspicuity mapping, terrain matching, and Simultaneous Location and Mapping (SLAM). E-RACON positioning employs radar beacons that transmit Morse code, enabling the vessel to ascertain its location by measuring range and bearing to numerous beacons; however, this technique requires many beacons and is constrained by range limitations. Image matching entails comparing current radar scans with reference images to ascertain the vessel's location; however, its precision diminishes as the vessel distances itself from the reference point. Radar conspicuity mapping generates a comprehensive map of radar targets in a specified region, enabling boats to correlate current radar data with the map to ascertain their location within around 30 meters. Terrain matching, which emulates radar returns using terrain data, has promise but is constrained by imperfections in terrain data and radar settings. SLAM allows warships to create their own radar maps while navigating, even in undiscovered territories, by combining radar data with conventional dead reckoning, resulting in positioning precision of 20–25 meters. These approaches provide potential answers, with SLAM being the most precise, delivering dependable position fixing that may ultimately satisfy the IMO's accuracy standards of 10 meters [4].

EXPERIMENTS

A. Dataset

Data Acquisition In contrast to the conventional autonomous navigation challenges faced by self-driving vehicles or UAVs that can gather data in real-world contexts, constructing intricate settings like collapsed towns or buildings in reality is a formidable effort. Consequently, we develop simulation models of these settings in Gazebo and gather visual data from the simulation. Specifically, we gather data from three categories of intricate environments:

The home collapsed due to an accident or calamity, such as an earthquake, resulting in many things scattered across the ground.

Collapsed city: Analogous to a collapsed dwelling, but pertaining to the external environment. In this situation, the roadway is strewn with rubble from the fallen structure.

A natural cave: An elongated natural tube characterized by low luminosity and uneven geological formations.

To construct the simulated environments, we first develop the 3D models of commonplace things seen in both indoor and outdoor settings (e.g., beds, tables, lights, computers, tools, trees, automobiles, rocks, etc.), as well as damaged items (e.g., shattered vases, broken dishes, and detritus). The items are then selected and positioned manually throughout each area to construct the whole simulated setting.

In each location, we use a mobile robot model outfitted with a laser sensor and a depth camera positioned atop the robot to gather visual data. The robot is operated manually to traverse each location.

We get visual data throughout the robot's movement. All visual data (D; P; I) are synced with the robot's current steering signal at each timestamp.

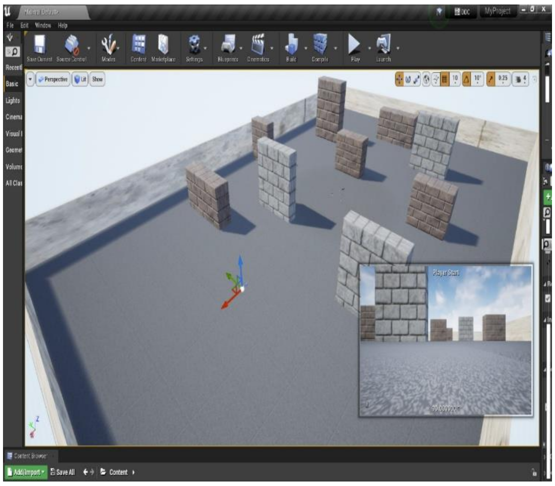
Statistical Data We specifically develop 539 three-dimensional object models to construct intricate worlds. These items are used to construct a total of 30 environments (i.e., 10 instances for each environment). On average, the collapsed home settings consist of around 130 items within an area of 400 m². The collapsed city has 275 artifacts distributed over an area of 3,000 m², and the natural cave habitats include 60 artifacts across an estimated area of 4,000 m². We manually operate the robot for 40 hours to get the data.

We capture around 40,000 visual data triples (D; P; I) for each kind of environment, culminating in a large-scale collection including 120,000 records of synchronized RGB images, point clouds, laser distance maps, and ground-truth steering angles. Approximately 45% of the information is acquired by domain randomization by applying random textures to the surroundings (Fig. 5). For each setting, we allocate 70% of the data for training and 30% for testing. All 3D settings and our dataset will be made freely accessible to promote future study.

Implementation

We use the TensorFlow framework to create our network. The network is tuned with stochastic gradient descent with a fixed learning rate of 0.01 and a momentum of 0.9.

The dimensions of the input RGB image and distance map are (480, 640) and (320, 640), respectively, while the point cloud data is sampled to 20,480 points. The network is trained with a batch size of 8, using around 30 hours on an NVIDIA 2080 GPU.



Results

The regression outcomes using Root Mean Square Error (RMSE) for our NMFNet and additional cutting-edge methodologies. The table indicates that our NMFNet far surpasses other approaches. Specifically, our NMFNet, trained on domain randomization data, attains an RMSE of 0.389, demonstrating a significant improvement compared to existing approaches that use just RGB pictures, such as DroNet [29].

This further supports that using several visual modalities as input in our fusion network is essential for effectively navigating complicated situations.

In three intricate environmental types, we see that the RMSE of the collapsed home outcomes exceeds that of the collapsed city and the natural cave. A potential explanation is that the collapsed home setting is much smaller than others, but has a greater number of things. Consequently, the robot would have more difficulty navigating the fallen home without colliding with the things. Table I indicates that the implementation of domain randomization significantly enhances the performance of our NMFNet with DR compared to the configuration without domain randomization (NMFNet without DR). Conversely, the VariationNet technique [31] has the most inaccuracy across all three complicated settings, although Inception-V3 demonstrates satisfactory performance.

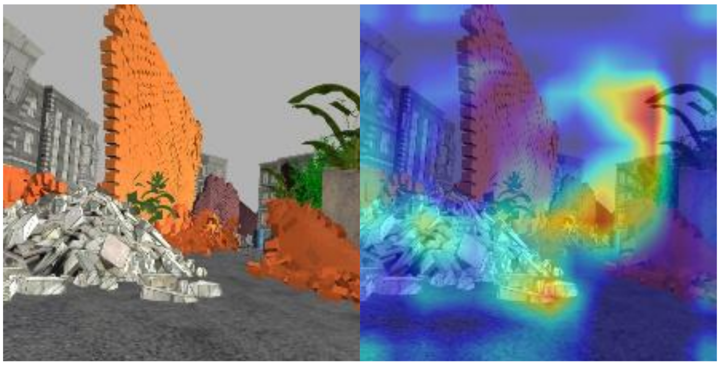


Fig. 4. The activation map when different modalities are used to train the network.

Conclusion

This research has looked at how artificial intelligence (AI) is being implemented into drone technology, illustrating how it is transforming a variety of sectors, including logistics, agriculture, military, and environmental monitoring. Drones equipped with artificial intelligence (AI) have significantly enhanced their autonomy, operational efficiency, and decision-making capabilities, therefore enabling novel applications in areas such as resource optimization, intelligent surveillance, precision agriculture, and disaster response. Despite the considerable promise of AI-powered drones, concerns about data security, regulatory constraints, and ethical difficulties remain paramount.

Policymakers, corporate executives, criteria. Drones equipped with AI may autonomously traverse agricultural fields while continually improving their flight paths; thus, researchers must collaborate to successfully tackle these issues and promote the safe and responsible use of AI in drone systems. This research underscores the substantial potential of AI to revolutionize drone-based applications, while also emphasizing the need of ongoing development and regulatory measures. To optimize the advantages of AI-integrated drones while mitigating associated risks, future advancements should concentrate on enhancing AI algorithms, increasing computing efficiency, and ensuring ethical AI implementation.

REFERENCES

[1] S. D. Panjaitan et al. "A Drone Technology Implementation Approach to Conventional Paddy Fields Application". In: *IEEE Access* 10 (2022), pp. 120650–120658. doi: 10.1109/ACCESS.2022.3221188.

[2] S. Sanz-Martos et al. "Drone Applications for Emergency and Urgent Care: A Systematic Review". In: *Prehospital and Disaster Medicine* 37.4 (2022), pp. 502–508. doi: 10.1017/S1049023X22000887.

[3] C. A. S. Lelis et al. "Drone-Based AI System for Wildfire Monitoring and Risk Prediction". In: *IEEE Access* 12 (2024), pp. 139865–139882. doi: 10.1109/ACCESS. 2024.3462436.

[4] T. Ko¸c. *Drone Technologies and Applications*. IntechOpen. 2023. doi: 10.5772/ intechopen.1001987.

[5] Aleksandar Petrovski and Marko Radovanovi´c. "Application of Drones with Artificial Intelligence for Military Purposes". In: *Journal of Advanced Military Technolo- gies* 4.3 (2022), pp. 45–56.

[6] Osim Kumar Pal et al. "A Comprehensive Review of AI-enabled Unmanned Aerial Vehicle: Trends, Vision, and Challenges". In: *arXiv preprint arXiv:2310.16360* (2023). url: https://arxiv.org/abs/2310.16360.

[7] Fadi AlMahamid and Katarina Grolinger. "Autonomous Unmanned Aerial Vehicle Navigation using Reinforcement Learning: A Systematic Review". In: *arXiv preprint arXiv:2208.12328* (2022). url: https://arxiv.org/abs/2208.12328.

[8] Enkhtogtokh Togootogtokh et al. "An Efficient UAV-based Artificial Intelligence Framework for Real-Time Visual Tasks". In: *arXiv preprint arXiv:2004.06154* (2020). url: https://arxiv.org/abs/2004.06154.

[9] Mohamed-Amine Lahmeri, Mustafa A. Kishk, and Mohamed-Slim Alouini. "Artificial Intelligence for UAV-enabled Wireless Networks: A Survey". In: *arXiv preprint arXiv:2009.11522* (2020). url: https://arxiv.org/abs/2009.11522.

[10] Montaser N. A. Ramadan et al. "Towards Early Forest Fire Detection and Prevention Using AI-powered Drones and the IoT". In: *Journal of IoT and AI Systems* (2024).

[11] S. D. Panjaitan et al. "A Drone Technology Implementation Approach to Conventional Paddy Fields Application". In: *IEEE Access* 10 (2022), pp. 120650–120658. doi: 10.1109/ACCESS.2022.3221188.

[12] S. Sanz-Martos et al. "Drone Applications for Emergency and Urgent Care: A Systematic Review". In: *Prehospital and Disaster Medicine* 37.4 (2022), pp. 502–508. doi: 10.1017/S1049023X22000887.

[13] C. A. S. Lelis et al. "Drone-Based AI System for Wildfire Monitoring and Risk Prediction". In: *IEEE Access* 12 (2024), pp. 139865–139882. doi: 10.1109/ACCESS. 2024.3462436.

[14] T. Ko¸c. *Drone Technologies and Applications*. IntechOpen. 2023. doi: 10.5772/ intechopen.1001987.

[15] Aleksandar Petrovski and Marko Radovanovi´c. "Application of Drones with Artificial Intelligence for Military Purposes". In: *Journal of Advanced Military Technolo- gies* 4.3 (2022), pp. 45–56.