

Article

Road Condition Detection and Emergency Rescue Recognition Using On-Board UAV in the Wildness

Chang Liu ^{1,2} and Tamás Szirányi ^{1,2,*} 

¹ Department of Networked Systems and Services, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary

² Machine Perception Research Laboratory of Institute for Computer Science and Control (SZTAKI), Kende u. 13–17, H-1111 Budapest, Hungary

* Correspondence: sziranyi.tamas@sztaki.hu; Tel.: +36-20-999-3254

Abstract: Unmanned aerial vehicle (UAV) vision technology is becoming increasingly important, especially in wilderness rescue. For humans in the wilderness with poor network conditions and bad weather, this paper proposes a technique for road extraction and road condition detection from video captured by UAV multispectral cameras in real-time or pre-downloaded multispectral images from satellites, which in turn provides humans with optimal route planning. Additionally, depending on the flight altitude of the UAV, humans can interact with the UAV through dynamic gesture recognition to identify emergency situations and potential dangers for emergency rescue or re-routing. The purpose of this work is to detect the road condition and identify emergency situations in order to provide necessary and timely assistance to humans in the wild. By obtaining a normalized difference vegetation index (NDVI), the UAV can effectively distinguish between bare soil roads and gravel roads, refining the results of our previous route planning data. In the low-altitude human-machine interaction part, based on media-pipe hand landmarks, we combined machine learning methods to build a dataset of four basic hand gestures for sign for help dynamic gesture recognition. We tested the dataset on different classifiers, and the best results show that the model can achieve 99.99% accuracy on the testing set. In this proof-of-concept paper, the above experimental results confirm that our proposed scheme can achieve our expected tasks of UAV rescue and route planning.



Citation: Liu, C.; Szirányi, T. Road Condition Detection and Emergency Rescue Recognition Using On-Board UAV in the Wildness. *Remote Sens.* **2022**, *14*, 4355. <https://doi.org/10.3390/rs14174355>

Academic Editors: Valerio Baiocchi, Alessandro Mei and Xianfeng Zhang

Received: 29 July 2022

Accepted: 30 August 2022

Published: 2 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the development of artificial intelligence technology, machine vision technology is being applied to many aspects of human life. For example, in precision agriculture, it can be used for crop farming [1]; in transportation, the use of machine vision technology enables the detection and counting of vehicles in highway scenes [2]; in human daily life, it can also detect pedestrians in surveillance videos [3] and human falls [4], etc. Due to the flexibility and portability of drones and other features, unmanned aerial vehicle (UAV) vision has also been developed to a certain extent. Drone vision technology [5] is also increasingly being noticed and applied by the research community as well as industry. Especially in the field of wilderness rescue, drones will be even more advantageous, as they can reach places that are inaccessible to humans, which makes them extremely convenient for wilderness rescue [6,7]. In wilderness environments, where the GPS, mobile, radio, or road path is not fully covered and where a few people are present or congregate nearby, there is no proper infrastructure. Indeed, most of the wilderness is forested, mountainous, or unexplored, so when one or more people are in these environments for whatever reason, there is a certain potential danger. For example, for people who love hiking [8] and for

those who are lost in the wilderness, a timely drone rescue and possible path monitoring are necessary when they are in danger.

Drones flying at low latitudes can detect in real-time what is happening on the ground, and to some extent enable search and rescue in the event of a natural disaster [9,10]. Similarly, remote sensing data from satellites [11–13] at high latitudes can also provide multiple channels of information to humans. Satellite data can be downloaded in advance and is constantly updated if the satellite is working. It is usually of a lower resolution than the image information captured by low altitude drones, but satellite data can provide information in multiple bands [14], which is not possible with most drones. Satellite remote sensing images play an irreplaceable role in the field of earth observation, and the different combinations of multiple channels can provide mankind with a wealth of ground-based information. Satellite remote sensing has been used in many fields, such as food security in agriculture [15], water body detection [16–19], segmentation of remote sensing images [20], land-use land-cover classification [21,22], change detection [23–25], environmental monitoring [26], etc. Based on geographical location information, we can fuse satellite information with real-time information captured by drones at low latitudes, both of which complement each other to provide greater assistance to humans. Of course, with the development of drone technology, a part of the drone is designed to carry multispectral camera to capture information of different bands, but such sensors are far from the bands captured by satellites.

The research background of this work is for a field environment with a poor network and bad weather conditions. When humans are in the wild without the network, owing to a lack of active antenna towers and/or cable infrastructure, they can encounter some unexpected problems and must seek help from the outside world; at this time, the drone rescue is particularly important. Drones will be more efficient compared to rescue workers. In addition to the above-mentioned network problems, wilderness environments also feature poor road conditions [27]; due to the sparse population, many roads in the wilderness are not well built, are usually composed of bare soil, or there may even be no roads that can allow people to walk. At this time, if there is a rainstorm and other bad weather, then the soil roads will become very muddy, which is extremely unfavorable to human use. Considering the poor road conditions in most of the wilderness, coupled with the effects of bad weather, it is particularly important that drones can detect road conditions and plan optimal routes for humans in distress. Even if humans have maps downloaded in advance on their phones in the above environment, there is no guarantee that the downloaded maps will match 100% with the actual roads encountered, especially in the wild, as many roads are not included in the maps [28].

Based on the above research background, the main goals of this work are as follows:

- The UAV plans an optimal route for human navigation in real-time in wilderness environments with poor mobile networks and bad weather;
- With the local road quality ensuring the probable walking or driving speed, multispectral images are combined with weather information to estimate the walking speed of pedestrians, and the throughput of the road should also be evaluated;
- Different roads are assigned weights by geometric features, such as road length and width on the road extraction map, as well as estimated pedestrian speed and road throughput;
- Human–UAV interaction in low latitudes by adjusting the flight altitude of the UAV to facilitate the recognition of emergency hand gesture signals and potential hazards. e.g., injured or tired persons, or vehicles with technical defects.

The main challenges to be solved to achieve the above goals are as follows:

- The drones use extracted road network maps with different weights to navigate the user in real-time and plan an optimal route that meets the needs by weighted searching algorithms;

- The classification of normalized difference vegetation index (NDVI) values by multispectral images allows the differentiation of different road types. The walking speed of pedestrians on the road and the road throughput can be obtained by certain analyses;
- Road extraction allows us to obtain the length, width, and connectivity characteristics of a road. Pedestrian detection and tracking techniques allow for a theoretical evaluation of the road's throughput. Combined with weather information and road surface materials, the walking speed of pedestrians can be estimated. In short, these are the main basis for road priority assignment;
- The model for human–UAV interaction needs to be accurate, stable, and robust enough when there is a problem with navigation or when a human is in danger and needs to be rescued. Different models need to be compared in order to select the best and most reliable one.

In this proof-of-concept paper, we propose a technique to detect road conditions and plan the best route for humans in a field environment using UAV vision and multispectral imagery based on road extraction [29]. We use the road extraction techniques and modify the shortest path searching algorithms that have been proposed and widely used. The type of road material is detected based on multispectral information from satellites or UAVs, and the walking speed of pedestrians on different roads and the throughput of the road are evaluated in combination with weather information to plan an optimal route for humans. Based on the different flight heights of drones, drones at high altitudes can plan the optimal route for humans, and drones at low altitudes can perform human–UAV interaction. Users can interact with the drones through dynamic “Attention” and “Cancel” gestures [30] when humans encounter potential dangers or unexpected situations on the ground in the wild, such as when the best route planned according to the drones is wrong, the actual situation is that there is no more road to walk on ahead, or the user is threatened by terrorists, etc. At this point the drone can recognize emergency rescue signals and re-route the user or send rescuers directly. However, this work is limited by the accuracy of road extraction techniques, which is the main reason why we establish low-altitude human–drone interaction to assist humans in real-time. The road throughput and the assignment of road weights have not been carried out automatically.

The main innovations and contributions of this paper are as follows:

1. The surface material of the roads in the wilderness environment is classified using an immediate UAV-boarded multispectral camera or, additionally, the latest available satellite-based multispectral camera images. The main distinction was made between bare soil roads and concrete or gravel roads in good condition;
2. When flying at high altitude, the UAV analyzed the road conditions based on the detected road surface materials and weather information, evaluated the throughput of different roads, and assigned different weights to the roads extracted from the map for optimal route planning;
3. When the UAV is flying at low altitudes, human–UAV interaction is possible. Based on the media-pipe hand landmarks, the “ok, good, v-sign, sign for help” dataset is created to identify emergency distress signals, so that corresponding measures can be taken to re-route or send rescuers;
4. The fusion of low-altitude drone imagery and high-altitude satellite imagery allows for a wider range of search and rescue information. Drones use their flexibility for high-altitude road condition detection and extraction, low-altitude human–UAV interaction, and emergency hazard identification to maximize their ability to help humans in distress.

In the subsequent sections, Section 2 focuses on the related work, including a description of the technical background and the relevant datasets used, as well as the process of data processing and the parts that have been published and completed. Section 3 is the main implementation approach, introducing the whole system flowchart, detection of road surface materials, analysis of road conditions, and evaluation of road throughput, followed by the collection of emergency rescue dataset and training of the models. Section 4 shows

the experimental results of different parts of the whole system, including the results of the detection and analysis of road conditions in the wilderness environment when the UAV is flying at high altitudes, and the results of the recognition of emergency rescue signals by the UAV at low altitudes, as well as the evaluation of the model. Sections 5 and 6 contain a discussion section, the conclusions of this work, and an outlook for future work.

2. Related Work

2.1. Technical Background

The technical background in which this work is implemented is a GPU-equipped UAV with a NVIDIA Jetson AGX Xavier developer kit [31], along with a Parrot Sequoia+ Multispectral Camera [32]. A standalone on-board system is very important, since in the wilderness we do not have network to rely on. From Sabir Hossain's experiments [33] on different GPU systems, it was evident that the Jetson AGX Xavier was powerful enough for this proposed system. As shown in Figure 1, the UAV can process the video information captured by the camera in real-time. The Parrot Sequoia+ multispectral camera is chosen because it is multispectral and can capture four channels of information. The details of the specifications are shown in Table 1.



Figure 1. Drone with on-board GPU and multispectral camera.

Table 1. Jetson AGX Xavier and Parrot Sequoia+ multispectral camera specification.

Jetson AGX Xavier [31]		Parrot Sequoia + Multispectral Camera [32]	
GPU	512-core Volta GPU with Tensor Cores	Sensor	Multispectral sensor + RGB camera
CPU	8-core ARM v8.2 64-bit CPU, 8MB L2 + 4MB L3	Multispectral sensor	4-band
Memory	32GB 256-Bit LPDDR4x 137GB/s	RGB resolution	16 MP, 4608 × 3456 px
Storage	32GB eMMC 5.1	Single-band resolution	1.2 MP, 1280 × 960 px Green (550 ± 40 nm) Red (660 ± 40 nm) Red-edge (735 ± 10 nm) Near-infrared (790 ± 40 nm)
DL Accelerator	(2×) NVDLA Engines	Multispectral bands	Global
Vision Accelerator	7-way VLIW Vision Processor (2×) 4Kp60 HEVC/(2×) 4Kp60 12-Bit Support	Single-band shutter	Rolling
Encoder/Decoder		RGB shutter	
Size	105 mm × 105 mm × 65 mm	Size	59mm × 41mm × 28mm
Deployment	Module (Jetson AGX Xavier)	Weight	72 g (2.5 oz)

In addition to information captured by UAV cameras with multispectral sensors, satellite information is also involved. Currently, the European Space Agency (ESA) [34] or United States Geological Survey (USGS) [35] has many publicly available datasets that users can download and use for free, and the channel information provided by satellites is more and more extensive. As such, UAV equipped with GPU can operate in real-time without relying on the local network. Although its flight time is affected by battery life, etc., humans are working on solving this problem, and the battery life can be predicted [36], meaning that stable energy can be provided during the disaster [37]. Thus, in special cases for emergency rescue, the flight time can be controlled according to demand. In our proposed drone system, the endurance of a GPU-equipped UAV with an NVIDIA Jetson AGX Xavier developer kit, along with a Parrot Sequoia+ multispectral camera, is around half an hour. The endurance of the drones will be improved in our system if the battery capacity is increased and if the drones are flown with rotary wings. Another solution for the constraint of the drone endurance is that drones can also work together in the wilderness, with different drones performing different tasks and collaborating with each other through communication to carry out rescue work. Indeed, UAV can revisit the wilderness after returning from the charging station, or a replacement team is activated, where the returning unit checks the proposed roads, thus, taking advantage of the return time. The training of the model for the human–computer interaction part is carried out on the ground station computer, which is equipped with an NVIDIA GeForce GTX Titan GPU and an Intel(R) Core (TM) I7-5930k CPU.

2.2. Related Datasets and Data Processing

The dataset presented in Figure 2 was captured by a drone equipped with a Parrot Sequoia+ multispectral camera, located in Biatorbágy, Hungary. The latitude of Biatorbágy is 47.470682, and the longitude is 18.820559. Biatorbágy (Wiehall-Kleinturwall, in German) is a town in Pest County, Budapest Metropolitan Area, Hungary [38]. From Figure 2, we can see that the image captured by the drone contains four bands, which are green, near-infrared, red-edge, and red. This UAV dataset was collected in 2017, and current satellite imagery from Google Maps [39] is also shown in Figure 2. Some changes in houses and vegetation can be clearly seen, but several of the main roads are unchanged.

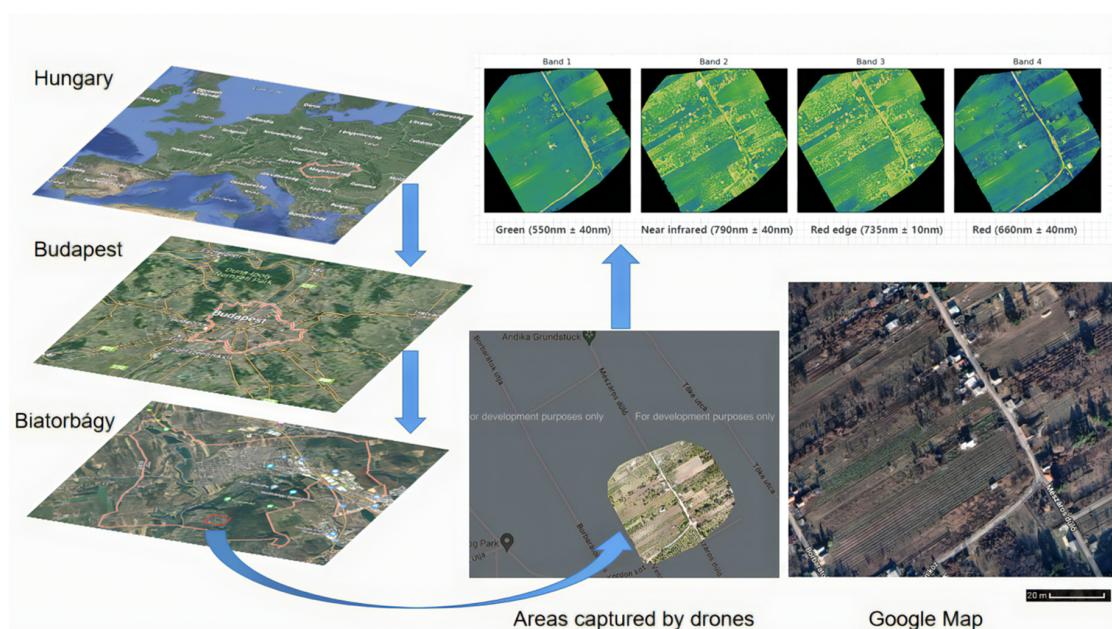


Figure 2. UAV with Parrot Sequoia+ multispectral camera dataset (The location is Biatorbágy, Budapest, Hungary).

GeoEye's OrbView-3 satellite (2003 to 2007) was among the world's first commercial satellites to provide high-resolution imagery from space. OrbView-3 collected 1 m panchromatic (black and white) or 4 m multispectral (color) imagery at a swath width of 8 km for both sensors. One meter imagery enables more accurate viewing and mapping of houses, automobiles, and aircraft, and makes it possible to create precise digital products. Four meter multispectral imagery provides color and near-infrared (NIR) information to further characterize cities, rural areas, and undeveloped land from space. Imagery from the OrbView-3 satellite complements existing geographic information system (GIS) data for commercial, environmental, and national security customers [40]. The data downloaded in this paper are from USGS (<https://earthexplorer.usgs.gov/>, accessed on 1 May 2022) [35], located in Australia, and the coordinates are $-35.038926, 138.966428$. Specific satellite information and information on the satellite datasets used in this paper are shown in Table 2.

Table 2. Orbview-3 Specifications and GeoEye's OrbView-3 satellite Dataset Attribute.

Orbview-3 Specifications [37]			Dataset (BirdWood, Adelaide, Australia) Attribute [35]	
Imaging Mode	Panchromatic	Multispectral	Entity ID	3V070304M0001619071A52 0001900252M_001655941
Spatial Resolution	1 m	4 m	Acquisition Date	2007/03/04
Imaging Channels	1 channel	4 channels 450–520 nm (blue) 520–600 nm (green) 625–695 nm (red) 760–900 nm (NIR)	Map Projection	GEOGRAPHIC
Spectral Range	450–900 nm		Date Entered	2011/11/10

In this work, we mainly used data from the Birdwood neighborhood. Birdwood is a town near Adelaide, South Australia. It is in the local government areas of the Adelaide Hills Council and the Mid Murray Council [41]. Figure 3 shows the specific information of the OrbView-3 satellite dataset used in this work, containing information in four different bands, namely blue, green, near-infrared, and red, as well as satellite information from the OrbView-3 satellite near Birdwood and satellite image information from Google Maps. Combining the information from the satellite dataset in Table 2, the area covered by the satellite is in Australia, and our study site is in the town of Birdwood in the city of Adelaide. Specific scale information and the location of the geographical coordinates of the study site, the town of Birdwood, can also be found in Figure 3. We selected the entire satellite image of Birdwood because the location is near a forested area, far from the urban center, and is mostly unoccupied in the wilderness, with both good and bad road conditions, which is consistent with the background of this work. The arrows in Figure 3 show the data selection process, and, finally, we captured the same areas on Google Maps for comparative display. Figure 3 also shows the four bands of the entire satellite map.

In the data processing phase, the input to the proposed system is the video sequence captured by the camera of the airborne UAV in real-time. When the available UAV does not have a camera with multispectral sensors, then the multispectral satellite data downloaded in advance will play a crucial role; when the UAV's camera has multispectral sensors, then the multispectral satellite information acquired in advance is available to be fused and used, and the two complement each other. Thus, whether through the drone data source or satellite data source, multispectral information can be accessed and used by humans. Since the UAV is equipped with an NVIDIA Jetson AGX Xavier, it can perform the processing of the video sequences captured by the camera in real-time into image sequences, which can be RGB images or multispectral images. Here, RGB three-channel images will be used as input for the road extraction part, and multispectral images will be used as input when detecting road conditions. Figures 2 and 3 show the experimental data areas from UAVs and satellites used in this task, respectively. The data types and features of the specific UAV and satellite referenced experimental areas are shown in Table 3.

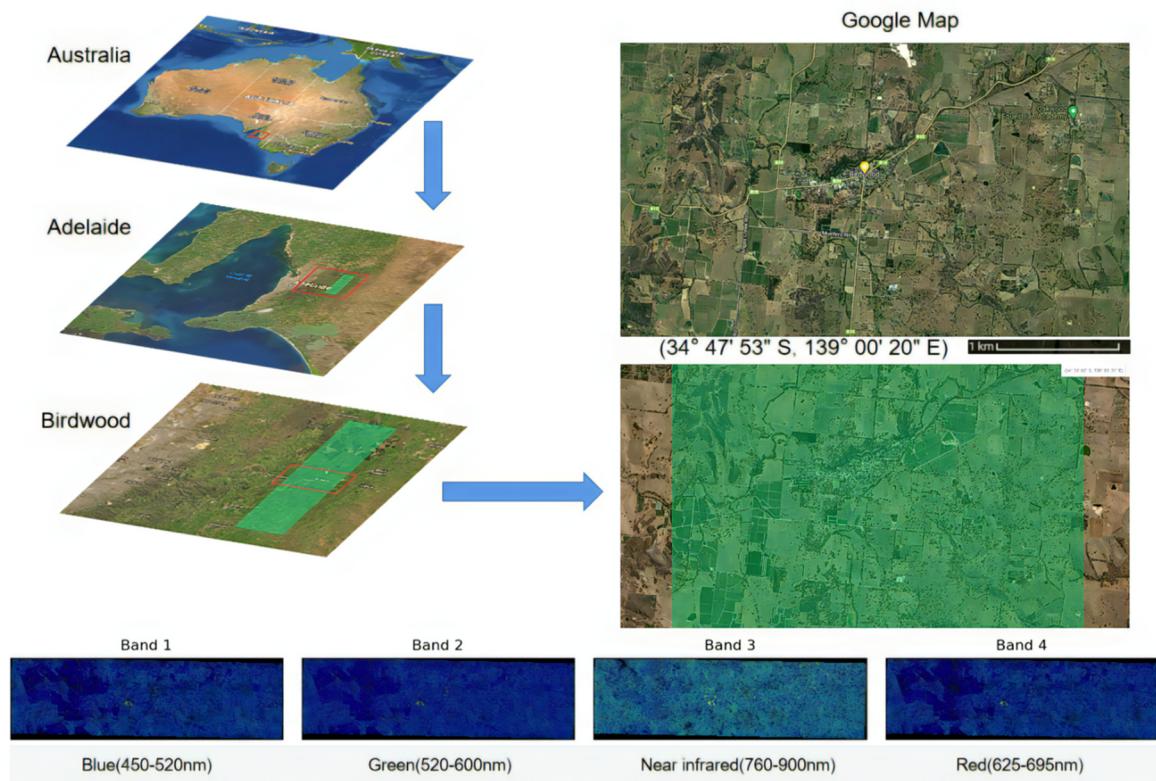


Figure 3. GeoEye's OrbView-3 satellite Dataset (The location is Birdwood, Adelaide, Australia).

Table 3. The data types and features of the UAV and satellite referenced experimental areas.

Data Types and Features	UAV Dataset	Satellite Dataset
RGB Images Resolution	$15,735 \times 14,355$	7202×2151
Multispectral Images	Red ($660 \pm 40 \text{ nm}$) Near-infrared ($790 \pm 40 \text{ nm}$)	Red (625–695 nm) Near-infrared (760–900 nm)
Cloud Cover	0	0
Coverage area	Small	large

2.3. Road Extraction and Optimal Route Planning

This subsection describes the work that has been carried out in road extraction and optimal route planning. For the task of road extraction, there are already many publicly available datasets collected by satellite that can be used, such as the Massachusetts Road Dataset [42] and DeepGlobe Dataset [43], and different road extraction networks have been proposed for these datasets, such as U-Net [44], Seg-Net [45], LinkNet [46], etc. In this work, considering our research background and the problem to be solved, we chose D-LinkNet [47], which won the first place in the Deep Globe Road Extraction Challenge. In addition to its better precision, recall, and F1 than other networks on the same dataset, D-LinkNet can handle road characteristics, such as narrowness, connectivity, complexity, and long span to some extent by expanding the sensory field and assembling multi-scale features in the central part while preserving detailed information, which is needed in our research work. For optimal route planning, road connectivity must be considered, because the routes are continuous rather than intermittent. D-LinkNet is a LinkNet with a pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. D-LinkNet can perform road extraction for the input three-channel RGB images, regardless of whether they are UAV images or satellite images. It can perform road segmentation well, labeling the roads as foreground and the other parts of the image as background. The DeepGlobe dataset also fits the context of our work, which is the field environment.

D-LinkNet can solve the road connectivity problem well, and the trained model can be used directly on the on-boarded UAV.

In the optimal route planning part, we have proposed and published the corresponding solution [48]. According to the map of road extraction, different weights are assigned to different roads. The weights are assigned by the road condition information, mainly including the road surface material, the length and width of the road, the walking speed of pedestrians on that road, and the estimated road throughput combined with the weather information, etc. Finally, the road map with weights is used for route planning by the A^* algorithm [49], and the best route can be planned for the user based on the road network map with different road priorities. Of course, users are required to report their starting location and the location of the destination they want to go to. Some optimal route planning results with complex road networks can be found in [48].

3. Methodology

3.1. Proposed System

The main workflow of the proposed system is to extract the road network by processing and analyzing the video streams captured in real-time using a UAV. The road surface condition is then detected in combination with pre-downloaded weather information and satellite information to assess the pedestrian speed and the road throughput. Next, the evaluation of pedestrian speed and road throughput is combined with the length and width of the road from the road extraction section to prioritize the different roads in the extracted road network. The final optimal route planning is based on the start location and destination provided by the user. When the user encounters a problem or potential danger, the user can attract the drone's attention through dynamic emergency gestures, and the drone can communicate with the user at low altitudes by adjusting its flight altitude, thus, providing maximum assistance to humans in distress.

The flow chart of the entire system is shown in Figure 4, where satellite data and weather data can be downloaded and analyzed and processed in advance. With the continuous development of weather forecasting, the information obtained from weather forecasts is becoming more and more accurate and can be relied upon [50]. As described in the introduction to the first section of this paper, the background of this work is that of a human being in a poorly networked and weathered field environment. The input to the system is a video sequence captured by a UAV flying in real-time, and the GPU-equipped UAV can process the video sequence to obtain a resolution of $1024 \times 1024 \times 3$ RGB images. If the UAV camera has a multispectral sensor, then the system will get four different channels of image information, namely green, near-infrared, red-edge, and red. The above steps prepared the way for the next road extraction and road condition analysis. The UAV at this stage was flown at an altitude of 50 m or more, and the RGB map obtained through data processing had a range of $512 \text{ m} \times 512 \text{ m}$.

The RGB image with a resolution of $1024 \times 1024 \times 3$ is used as the input of the road extraction network. With the road extraction by D-LinkNet [47], we can obtain information on the length and width of the road. Based on the satellite image information corresponding to the area where the UAV flies and the weather information downloaded in advance, we can determine whether there is bad weather, such as heavy rain, dusty weather, or a snowstorm, etc. This weather can lead to problems, such as a muddy bare soil road surface, low air visibility, or smooth road surfaces becoming icy, snowy, sandy, or water-covered, which can make them difficult for pedestrians to walk on. At this point, the material of the road surface is particularly important. Having a better condition road material is beneficial for human walking in rainy weather if the low visibility of the road caused by nearby bare soil is not considered in windy weather. On the contrary, if the road surface is composed of bare soil, then it will be difficult for humans to walk after heavy rain. Roads made of concrete or gravel are not so muddy, so the weather has a great influence depending on the different road materials used in an area. For the detection of road material information, we use multispectral information from satellites or drones to assess the normalized difference

vegetation index [51,52] and, depending on the classification of the normalized difference vegetation index, the road material is detected, since the normalized difference vegetation index values are different between bare soil and roads with concrete or gravel cover.

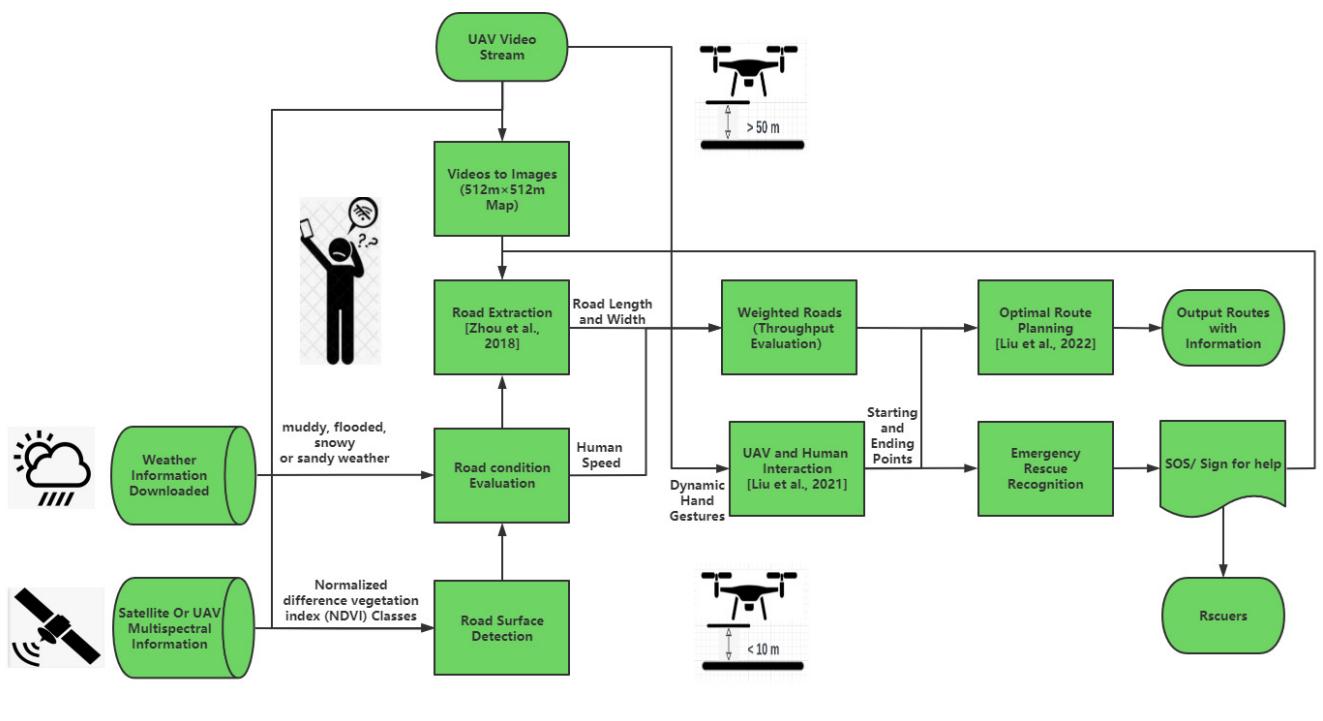


Figure 4. Flowchart of the proposed whole system [30,47,48].

For a road, we define two basic measures over the geometrical parameters of width and slope, namely walkability (the speed of pedestrians) and transferability (the number of pedestrians able to enter or exit the road during the analysis period). Combining weather information and road surface material information, we can estimate the speed of pedestrians. With pedestrian detection and tracking we can evaluate the throughput of the road [53,54] and finally combine the road length and width information to assign priority to the road. The road map with the weights can then be used to plan the best route using the A^* algorithm [48]. The system will eventually output an optimal route for the current situation to the user, thus, enabling the UAV to help humans navigate in the field.

Since, in the real world, there may be emergency situations, such as a human being encountering a threat or a route going wrong, and other special situations, communication between the drone and the human being is crucial. The drone can adjust its flight altitude, and when it lowers its flight altitude to less than 10 m, it can complete the interaction process with the human being on the ground through dynamic attention and cancellation gestures [30]. The drone will accurately identify potential dangers and emergencies, re-route humans in distress, or call SOS directly to request the intervention of rescuers.

3.2. Road Surface Detection and Road Throughput Evaluation

The normalized difference vegetation index (NDVI) is a simple graphical indicator that is often used to analyze RS measurements and assess whether or not the target being observed contains green healthy vegetation. This index will be referred to by the abbreviation NDVI throughout the rest of the paper. The NDVI [51,52,55] is derived from the red: near-infrared reflectance ratio, where NIR and RED are the amounts of near-infrared and red light, respectively, reflected by the vegetation and captured by the sensor of the satellite. The formula is based on the fact that chlorophyll absorbs RED, whereas the mesophyll leaf structure scatters NIR. Thus, NDVI values range from -1 to $+1$, where negative values correspond to an absence of vegetation [52]. The NDVI quantifies vegetation by measuring

the difference between near-infrared (NIR) (which the vegetation strongly reflects) and red light (which the vegetation absorbs/has a low reflectance) [56]. In our Section 2.2, both the UAV dataset and the satellite dataset are present in the near-infrared band and the red band. The formula for NDVI is given as follows in Equation (1), where NIR is near-infrared light and Red is visible red light:

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red}) \quad (1)$$

The value of the NDVI will always fall between -1 and $+1$. Values between -1 and 0 indicate dead plants, or inorganic objects, such as stones, roads, and houses. On the other hand, NDVI values for live plants range between 0 to 1 , with 1 being the healthiest and 0 being the least healthy [57]. Bare soils range from about 0.2 – 0.3 [58]. Based on the above presentation of the different NDVI values, we can determine that the value of NDVI for a road not covered by vegetation should be negative. However, if the road consists of bare soil only, then the NDVI values will be distributed in the range of 0.2 – 0.3 [58]. This effectively distinguishes the roads covered by concrete or gravel from those covered by bare soil.

Before evaluation of road throughput, we can estimate the human walking speed on different kinds of roads based on weather information and road surface material information. For rainy weather or just after a heavy rain, the speed of a human walking in the bare soil road is about 1.07 m/s [59]. The asphalt road or stone road will be more friendly to humans and, thus, human walking speed in rainy weather on the asphalt or stone road is about 1.2 m/s [60]. For dusty weather, which brings a reduction in air visibility, walking on asphalt-covered roads will be less dusty and faster than on bare soil-covered roads. For the snowy weather, a smooth road surface will speed up the human walking speed, so the road material is no longer the main factor affecting the human walking speed. For no adverse weather conditions, the human walking speed on different roads is the same. The speed can also be effectively estimated from other parts of the remotely sensed areas, where the situation is more definite.

After evaluating the speed of humans walking on different types of roads, the drones can see that humans are moving within and between different roads. In a limited time, for the 10 min duration of the drone analysis, the drone can evaluate the throughput of different roads and, depending on the results of the road extraction, different throughputs are obtained for different roads. Equation (2) is the definition of the road throughput. In our previous work, the UAV can count the number of pedestrians in real-time [61]. Throughput is defined as the number of distinct people able to enter or exit the road during the analysis period. To aid its interpretation, the throughput should be divided into five components according to whether or not the people entered, exited, never entered, or never exited the road during the analysis period [54]. The five categories are described as follows and shown in Figure 5. In Figure 5 we can see that a total of five categories of people are included, and each person in each category is given a specific ID. The ID of the same person is kept constant when entering and leaving the road, and the technique of assigning IDs to different people and tracking them can be found in our previous work in [61].

$$\text{Through-put} = N/T \quad (2)$$

where N is the number of distinct people able to enter or exit the road during the analysis period, and T is the time during the analysis period (e.g., 10 min). The five categories are as follows:

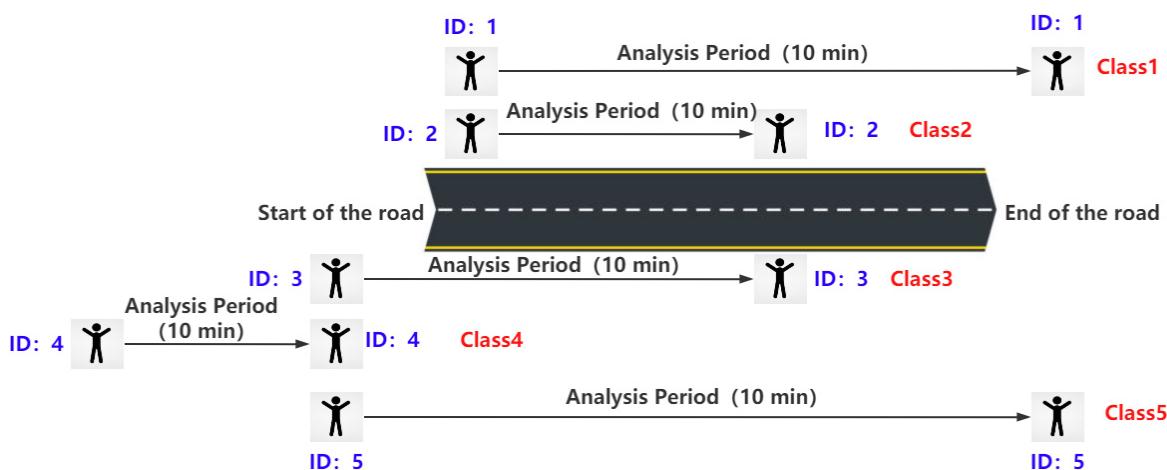


Figure 5. The five categories used to assess the road throughput.

1. Human Class 1—people that were present at the start of the analysis period and were able to successfully exit the road before the end of the analysis period;
2. Human Class 2—people that were present at the start of the analysis period but were unable to successfully exit the road before the end of the analysis period;
3. Human Class 3—people that were able to enter the road during the analysis period but were unable to successfully exit the road before the end of the analysis period;
4. Human Class 4—people that tried to enter the road during the analysis period but were completely unsuccessful;
5. Human Class 5—people that entered during the analysis period and were able to successfully exit the road before the end of the analysis period.

The percent of incomplete trips will be the sum of human classes 1, 2, 3, and 4, divided by the sum of all human classes ($1 + 2 + 3 + 4 + 5$). Generally, higher throughputs and lower percentages of incomplete trips are desired since they reflect the productivity of the road [54]. The throughput of the road and the length and width geometrical characteristics of the road directly affect the road weights that are assigned, that is, the priority assignment problem.

The task of optimal route planning based on roads with different weights has been solved in [48]. The road weight assignment in [48] is manual and based only on the road connectivity and length-width geometric properties, and does not make an analysis and evaluation of pedestrian walking speed, road surface material, road throughput, etc. This work complements the prioritization evaluation and weight assignment part of different roads in the previous road map networks. It is important to emphasize that in the route planning and navigation section, we are assuming that humans can report the starting location and destination for the UAV.

3.3. Emergency Rescue Recognition

The main flaw of drones flying at altitudes above 50 m, as is responsible for navigation, is that they are not able to see the situation on the ground in real-time and cannot meet the needs of the user in the current situation, so it is necessary to adjust the flight height to carry out human–UAV interaction. There are many uncertainties in the wilderness environment, so the drone flying at low altitudes below 10 m can communicate with the user through dynamic gestures [30] and, once there is an emergency, it is possible for the drone to make a timely response and take measures.

The Signal for Help [62] is a single-handed gesture that can be used to alert others that people feel threatened and need help. Originally, the signal was created as a tool to combat the rise in domestic violence cases around the world linked to self-isolation measures that were related to the COVID-19 pandemic. The signal is performed by holding one hand up with the thumb tucked into the palm, then folding the four other fingers down,

symbolically trapping the thumb by the rest of the fingers. It was designed intentionally as a single continuous hand movement, rather than a sign held in one position, so it could be made easily visible. As this gesture is widely spread and popularized, it is increasingly known and used as a signal for potential hazard identification in this work, and we created a dataset of the gesture by mixing the gesture into some common human gestures, and the detail of the dataset is shown in Table 4.

Table 4. Emergency rescue recognition dataset.

Name	Number of Data	Hand Gestures
ok	4232	
v-sign	3754	
Good	4457	
SignForHelp (dynamic)	3504	 1. Palm to camera and tuck thumb 2. Trap thumb

The datasets were collected using a 1080P 160° fisheye surveillance camera module for Raspberry Pi on the 3DR SOLO UAV system. Three people from our lab participated in UAV emergency rescue gesture dataset collection, the genders were two males and one female, aged between 25 and 30 years old. We collected data for each gesture in different orientations to cover as many situations as possible, making our dataset more generalizable. Table 4 shows the details of the UAV emergency rescue recognition dataset. The acquisition of this dataset was based on media-pipe hand landmarks, where we extracted the positional information of 21 key points on the human hand and saved them in a CSV file. The data extracted for each hand gesture was taken from a different person separately, and the final amount of data for each gesture is given in Table 4.

Figure 6 shows the main process of emergency rescue recognition, which is carried out in advance at the ground base station, and the final obtained model will be deployed directly on the UAV for use. The dataset collection is based on media-pipe hand landmarks [63,64], and the first process is the extraction of 21 hand key points. We collected four gestures from different people in our lab, which are ok, v-sign, good and sign for help. This extracted hand key-point data were stored in a CSV file, after which 70% of the entire dataset was used as the training set and 30% was used as the test set to be tested on different classifiers. We tested on the following five different classifiers: logistic regression [65], ridge classifier [66], random forest classifier [67], gradient boosting classifier [68], and deep neural network [69]. The classifier with the highest model accuracy at the end will win and be deployed for use on the UAV.

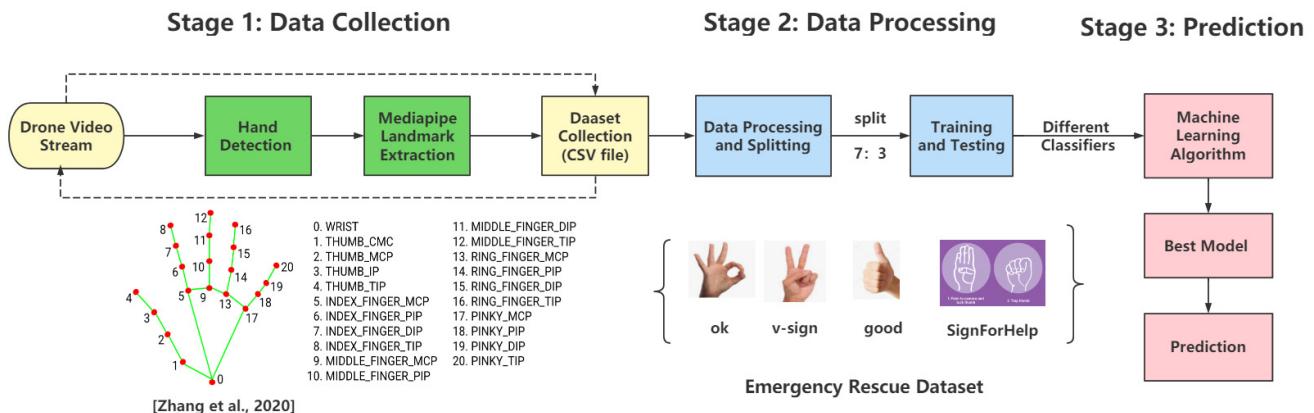


Figure 6. Architecture of the emergency rescue recognition method [64].

The architecture of the emergency rescue recognition method is divided into three main stages. The first stage is data collection, where the 21 key points of the hand are continuously presented from the video stream of the UAV input and stored in a CSV file for later model training. The second stage is the processing of the collected data and the separation of the data set into a training set and a test set. The third stage is the training of the different classifiers and the prediction of the results.

4. Experiments and Results

This section presents the experimental results of this work and contains two main subsections. The first, Section 4.1, presents the results of the detection of road surface materials on different datasets and shows the results of optimal route planning on different types of roads as well. The second part, Section 4.2, is the evaluation of a human–machine interaction model for UAVs in low altitude flight and a demonstration of simulated emergency rescue recognition in wildness.

4.1. Road Surface Condition Detection and Optimal Route Planning Results

This subsection shows the results of the road surface material detection and demonstrates the navigation task of optimal path planning. For the UAV dataset presented in Section 2.2, we obtained the results shown in Figures 7 and 8 with the proposed system. Figure 7 contains the RGB images taken by the UAV and the resulting NDVI images obtained by multi-channel image calculation. Normally, the range of NDVI is -1 to $+1$, but in this UAV dataset, we show the NDVI results as -0.19 to 0.94 , in order to clearly distinguish some road surface materials. The NDVI image is well illustrated in Figure 7, and by comparing the analysis with the RGB image, we can see that the red area from -0.19 to 0.25 is largely uncovered by vegetation and mostly inorganic material. The orange area from 0.25 to 0.45 is mainly bare soil, while the yellow or green part above 0.45 is covered by vegetation. We zoom in on the part containing the road to show it at the bottom of Figure 7, and we can see that the main road in the middle is red, which means that the road surface is composed of non-soil gravel material, while the narrow road on the right is close to orange in color, and we can see that the surface material is mainly composed of bare soil. In this way, the different road composition materials can be effectively distinguished by the NDVI.

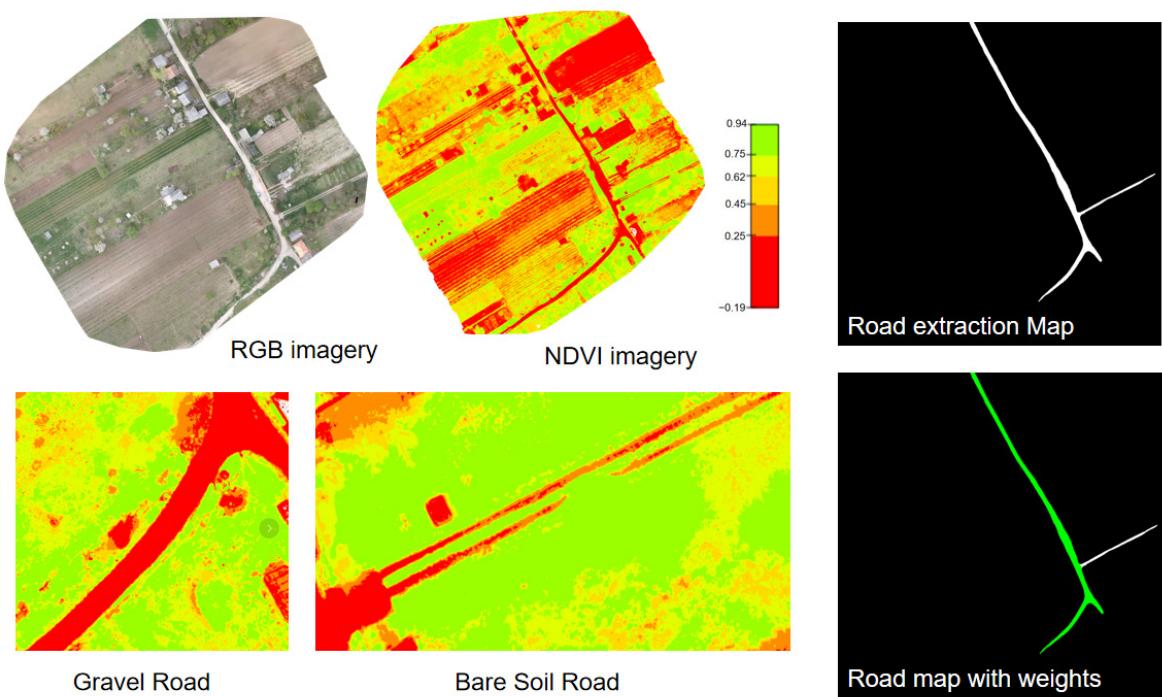


Figure 7. Road surface condition detection and road extraction with weights results for the UAV dataset.

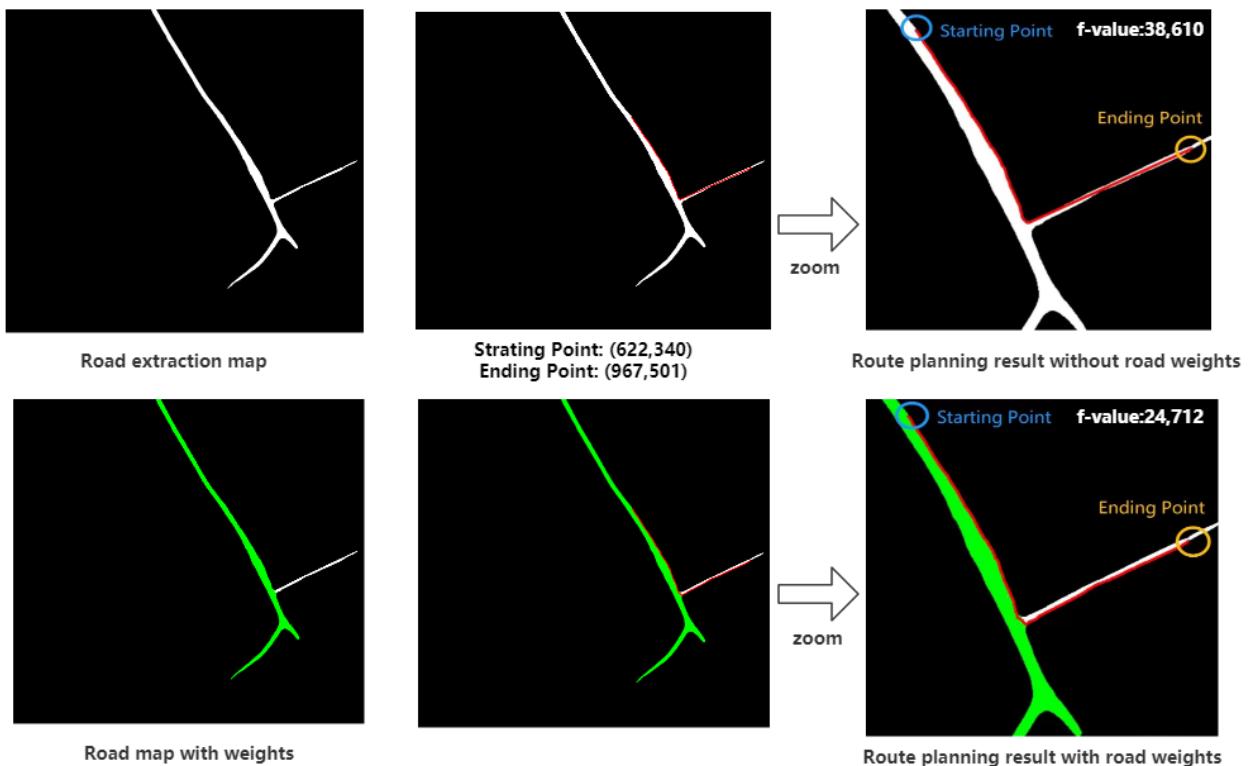


Figure 8. Optimal route planning results for the road maps with and without weight assignment in the UAV dataset.

The material composition of the road surface can have an impact on the walking speed of pedestrians under different weather conditions. According to the analysis and description inside the methodology in Section 3, we can evaluate the pedestrian speed by walking on different road surfaces under different weather conditions. In this UAV dataset, if the user encounters rainy weather, then the road consisting of narrow bare soil is very muddy and the walking speed of pedestrians is slower, but the walking speed of pedestrians on the main road consisting of gravel is faster. After road extraction, we can obtain the geometric features of the length and width of different roads. Applying the previously deployed yolo3-tiny model [61], we can evaluate the transferability of the road and, thus, obtain the throughput of that road. It is important to note here that the throughput of the road in the real case is not tested and, therefore, is not shown accordingly in the results section. The road surface material, the length and width of the road, the walking speed of the pedestrians, and the throughput of the road are combined and evaluated to give priority to the extracted road network. Of these, road throughput according to Equation (2) is used as the main criterion and basis. The results of road extraction and road priority assignment for the UAV dataset can also be found in Figure 7. The road network included in this dataset is relatively simple and not very complex due to the relatively small area included in the UAV dataset. More complex cases and optimal route planning results can be found in our published paper [48].

Figure 8 shows the best intelligent navigation route provided by the UAV based on the starting and ending position coordinates reported by the user. It can be clearly seen that in the two path planning comparison plots on the right side of Figure 8, the value of f in the plot with road priority assignment is reduced, where f indicates the cost of the shortest path at the pixel scale. Each pixel is considered a node, except that these nodes are given different weights. Here, $f(n)$ is the sum of $g(n)$ and $h(n)$, where n is the next node on the path, $g(n)$ is the cost of the path from the start node to the n node, and $h(n)$ is a heuristic function that estimates the cost of the cheapest path from n to the goal. Details of how the modified A^* algorithm finds the minimum value of f can be found in our previously published work [48]. The minimum value of f corresponding to the best route, found using the search algorithm, is 38,610. On the contrary, if the road is given priority, the minimum value of f corresponding to the best route found by the search algorithm is reduced to 24,712. This greatly improves efficiency for the user. As can be seen in the path planning in Figure 8, the optimal route search algorithm makes maximum use of the roads marked in green, because green roads represent roads in good condition, while white roads represent roads in poor condition. To show that the algorithm for optimal path planning works well, we used the results of a more complex road network from our previous work [48] in Figure 9 to show that the UAV is able to plan the best route when there are multiple paths to choose from. The f -value of the optimal path obtained by the best route search for the road map without priority assignment in Figure 9 is 4733, and the f -value of the best path planned for the road map with the priority road, on the contrary, is 1044. It should be noted that the image used in Figure 9 is from the DeepGlobe Road Extraction dataset [40], and the images are RGB three-channel.

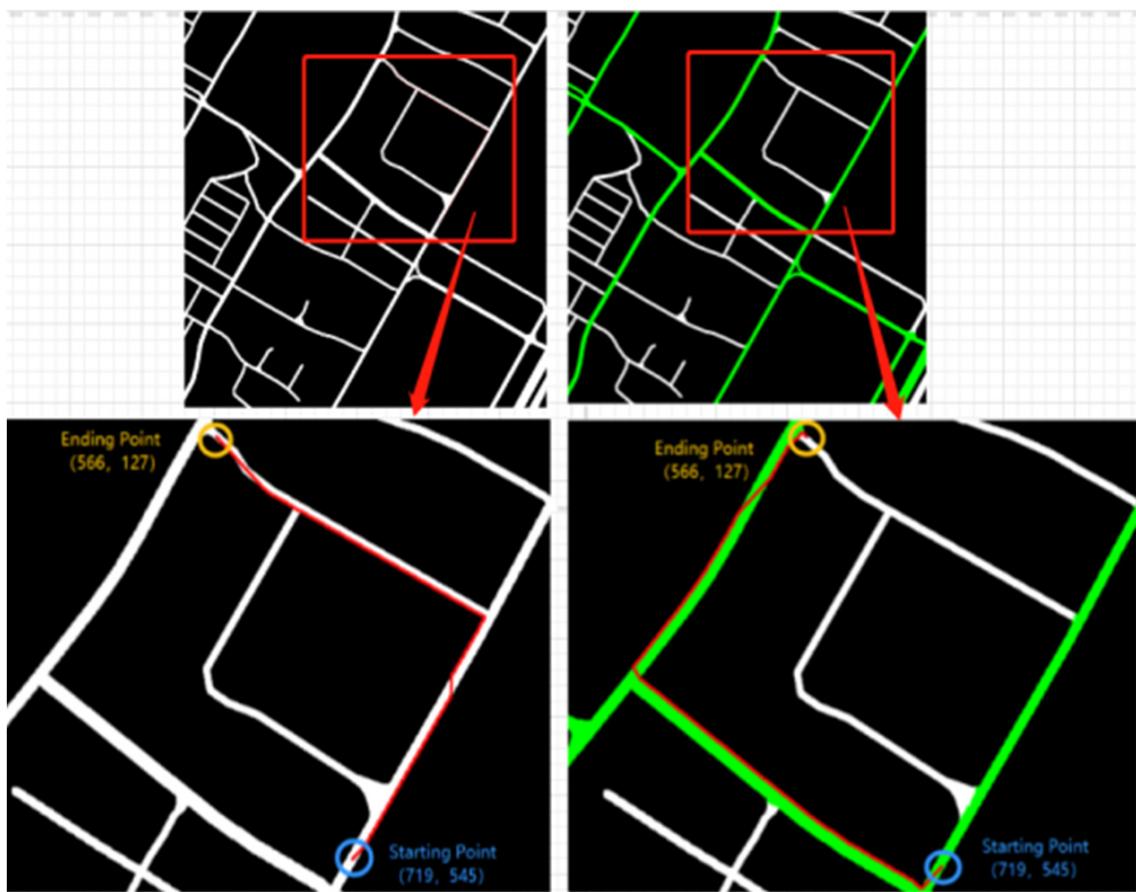


Figure 9. Optimal route planning results for the complex road maps with and without weight assignment in the previous work [48].

For the satellite dataset presented in Section 2.2, we also performed the corresponding tests. Figure 10 shows the NDVI index results for the whole satellite image and the results after classifying the NDVI. Here, similar to the processing of the UAV dataset described above, we shifted the NDVI value results from -1.0 to $+1.0$ to -0.2 to $+1.0$, displaying them with the purpose of distinguishing more clearly the material of the road surface. Based on the different values of NDVI we classified them into three categories. Those with NDVI values less than 0.2 are classified as being in good condition, which is generally inorganic in this case, and for road areas, where the surface material is composed of non-soil. On the other hand, NDVI values in the range of 0.2 to 0.4 are classified as being in a bare soil condition, and NDVI values greater than 0.4 are classified as vegetation. These three categories are shown in gray, tomato red, and green, respectively. These results can be found in Figure 10. We can see that, for roads, those with pavements consisting of bare soil are shown in tomato color, while those with pavements consisting of concrete and in good condition are shown in gray. In summary, roads with different road materials are well distinguished, which also provides the basis for the subsequent road priority assignment.

Figure 11 zooms in on the NDVI classification results of Figure 10, and we can see that the entire satellite image undergoes constant regional zooming, which allows a good distinction between the road conditions in different conditions. The entire satellite image was selected to show the results for the suburban area near Birdwood, but the rest of the road information is also well differentiated and can be compared with the actual situation in Google Maps. In fact, the classification of the whole satellite image is consistent with the information observed by Google Maps satellites as far as the places covered by vegetation are concerned. If you zoom in on the area, the information for the road is also the same as that observed by Google Maps, so the different road surface materials are effectively

distinguished from each other. The last zoomed area in Figure 11 shows that the road network is clear; some roads are shown in tomato color which means that the roads are composed of bare soil, while some roads are shown in gray color which means that the surface of these roads is composed of non-bare soil, such as gravel or asphalt, and they are not very muddy after bad weather, especially after heavy rain.

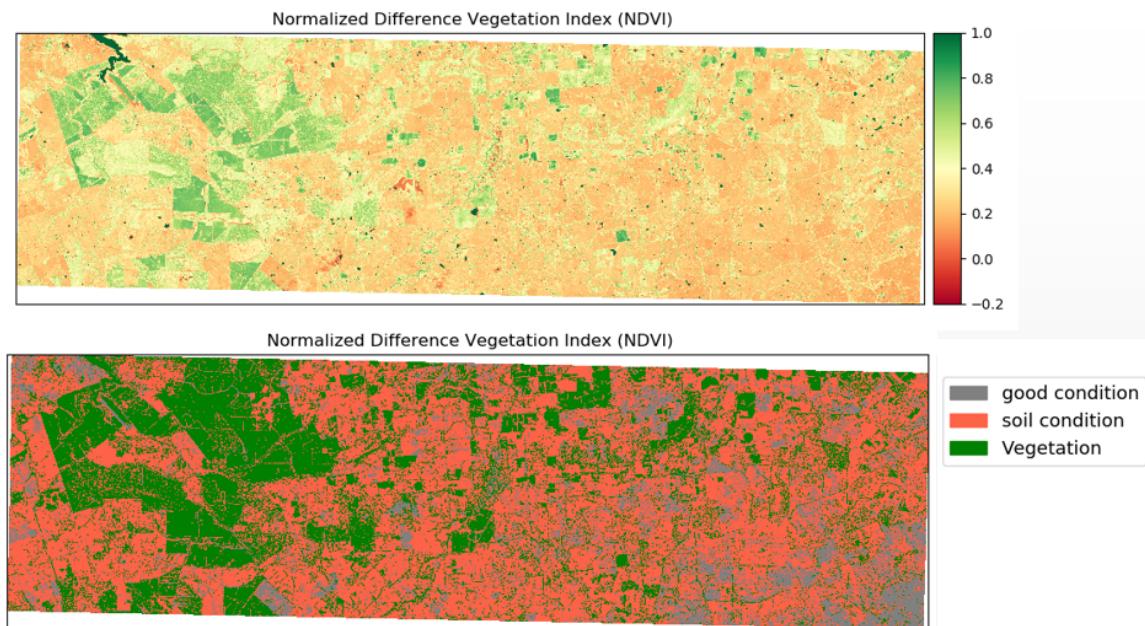


Figure 10. NDVI index and NDVI classes results for GeoEye’s OrbView–3 satellite dataset.

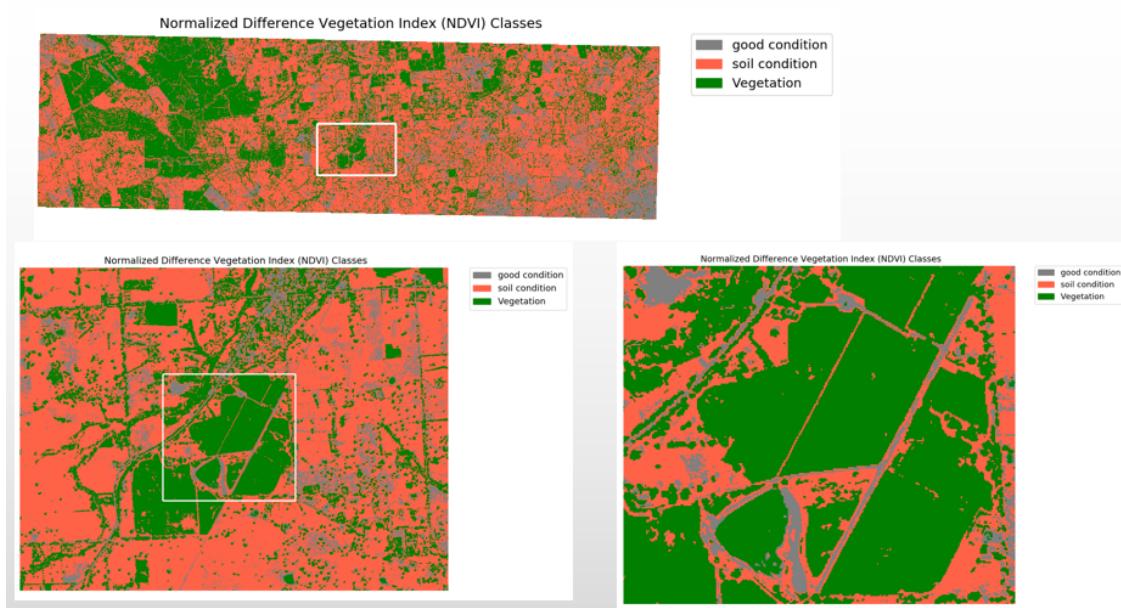


Figure 11. Road surface condition detection and classification results for GeoEye’s OrbView–3 satellite dataset by zooming in on white boxes.

Near the Birdwood area, we performed the detection and classification of the road surface material in a small area, as well as the extraction and prioritization of the roads in the area, and the related results are shown in Figure 12. According to the classification of road surface materials of the different roads shown in Figure 11, we can evaluate the walking speed of pedestrians under the influence of bad weather. After heavy rain, the

road surface composed of bare soil is muddy, which is unfavorable for pedestrians to walk on. On the contrary, if the road composition is non-muddy, such as concrete or gravel, then the road is less muddy and the walking speed of pedestrians is about 1.2 m/s, which is faster than the walking speed on the muddy road. The walking speed of pedestrians also affects the throughput of that road, so it is necessary to evaluate the speed. The results of road extraction allow us to obtain the two geometric characteristics of road length and width. Similarly, by deploying the yolo3-tiny model we can analyze the throughput of each road in 10 min. Here, the corresponding results are not shown, due to the lack of real cases to test. Finally, the road weights are assigned by considering the composition of the road material, the walking speed of pedestrians, the length and width of the road, and the throughput of the road. Here, green represents roads with priority and white represents roads without priority, consistent with our previous work [48]. Theoretically, as shown in Figure 5, road throughput is obtained by counting and tracking the number and location of pedestrians on each road in real-time at low altitude by a GPU-equipped UAV [61]. The road surface material, the estimation of pedestrian speed, the road length and road width, and the road throughput assessment are the main basis for assigning road weights.

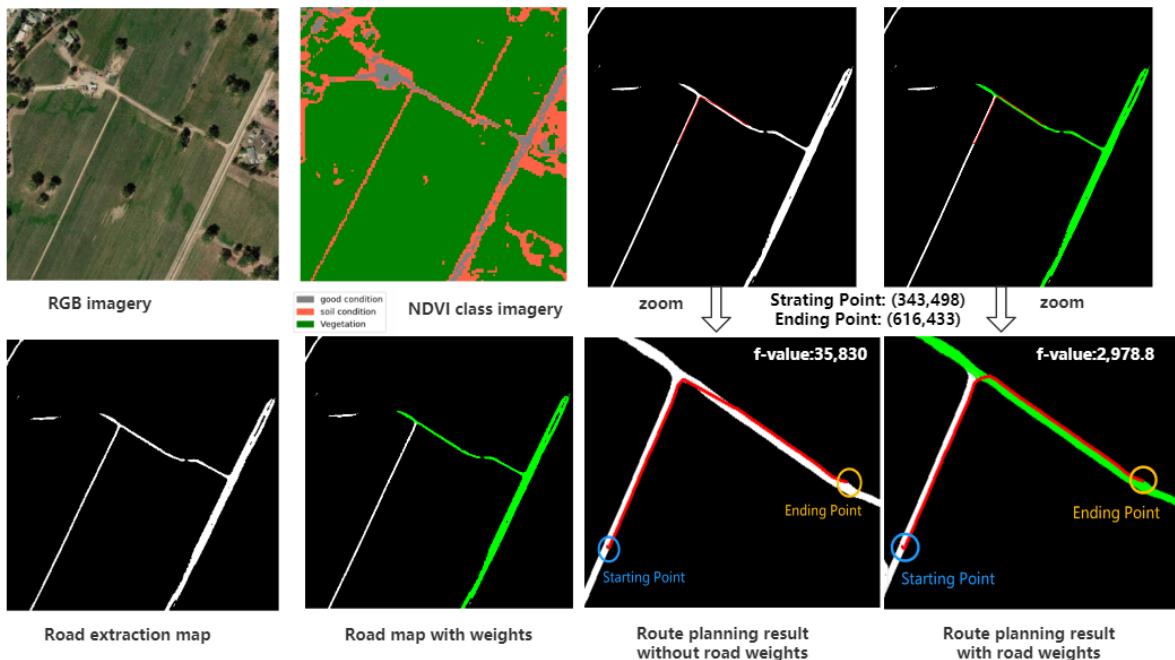


Figure 12. Results of road extraction and optimal path planning on the GeoEye’s OrbView–3 satellite dataset.

The results of the best route planning are also shown in Figure 12, where we can see that the minimum value f from the final route search is different based on the same starting position and ending position coordinates reported by the user. The blue circle in the figure represents the starting location, the yellow circle represents the destination, and the red line is the best route searched by the algorithm. We can see that the value of f in the best route search result is 35,830 when no priority is assigned to the roads and, on the contrary, that the value of f is reduced to 2978.8 when roads have priority, which means that the route is faster and more efficient than the previous one. The maps shown here for road extraction are simple. The relatively complex road extraction results and the corresponding optimal route planning results can be found in [48], which also shows that the A^* algorithm with weights is effective. Since the extraction result of D-LinkNet does not match the real situation 100%, we can see that, in the road extracted image, a section of the road with bare soil on the surface is missing compared to the actual RGB image, which may cause the user to encounter some emergency situations, such as the route planned by the UAV appearing

to have no road in front of it, or the route planned by the UAV is best in the model, but the user actually found a better road to walk on. These situations require the drone to communicate with the user in real-time, and emergency rescue situations are possible to re-plan the routes for the users.

4.2. Human–UAV Interaction for Emergency Rescue Recognition

Based on the dataset created in Table 4, we split the entire dataset into 70% and 30% portions, with the 30% serving as the testing set. The testing set for validating the accuracy of the model contains 4784 data points, i.e., 30% of the entire data set was randomly selected to validate the accuracy of the model. Compared to other hand gesture recognition methods, such as using 3D convolutional neural networks [70], we finally chose the hand key points as the basic feature for emergency rescue recognition. The reason is that the features of the hand key points are concise, intuitive, and easy to distinguish between different hand gestures. In contrast, 3DCNN is both time-consuming and struggles to train large neural networks. As for the classifiers, we tested on five different classifiers, namely logistic regression [65], ridge classifier [66], random forest classifier [67], gradient boosting classifier [68], and deep neural network [69]. The accuracies obtained for these classical and commonly used classifiers on the testing set of our dataset are shown in Table 5. We have retained six decimal digits for the accuracy results. We can see that the DNN results are the highest, which is why we chose to train the model with DNN at the base station and finally deploy it on the drone. The model accuracy of the DNN can reach 99.9% on the testing set, and this accuracy is crucial for the research part of this work, as the drones need to be ready to accurately identify when the user is in an emergency or when there is potential danger around them.

Table 5. Accuracy of emergency rescue recognition dataset on different classifiers.

Classifiers	Accuracy on Testing Dataset
Logistic Regression [65]	99.5193%
Ridge Classifier [66]	98.5789%
Random Forest Classifier [67]	99.6865%
Gradient Boosting Classifier [68]	99.5402%
Deep Neural Network [69]	99.9164%

The DNN model has been programmed using Keras Sequential API in Python and compiled using Keras with a TensorFlow backend. There are four layers with batch normalization behind each one and 128, 64, 16, and 4 units in each dense layer sequentially. The total number of parameters included is 18,260, of which 17,844 are trainable parameters and 416 are non-trainable parameters. The last layer of the model is with Softmax activation and 4 outputs. The categorical cross-entropy loss function is utilized because of its suitability to measure the performance of the fully connected layer's output with Softmax activation. Adam optimizer with an initial learning rate of 0.0001 is utilized to control the learning rate. Figure 13 shows the changes in accuracy and loss of the DNN model throughout the training process. We can see that the model stabilizes in accuracy and loss after 20 epochs of training and, after 100 epochs of training, the model achieves an accuracy of 99.99% on the training set and 99.92% on the testing set. Figure 14 shows the confusion matrix of the model on the testing set, which is an evaluation of the DNN model, and we can see that the predictions are concentrated on the diagonal, meaning that most of them are accurately predicted. The processing time of the DNN model was computed using the start a timer clock function in Python code. The real running time of the emergency rescue recognition framework is around 20 ms. In human interaction, the FPS value is maintained at around 5, which is in accordance with the real motion.

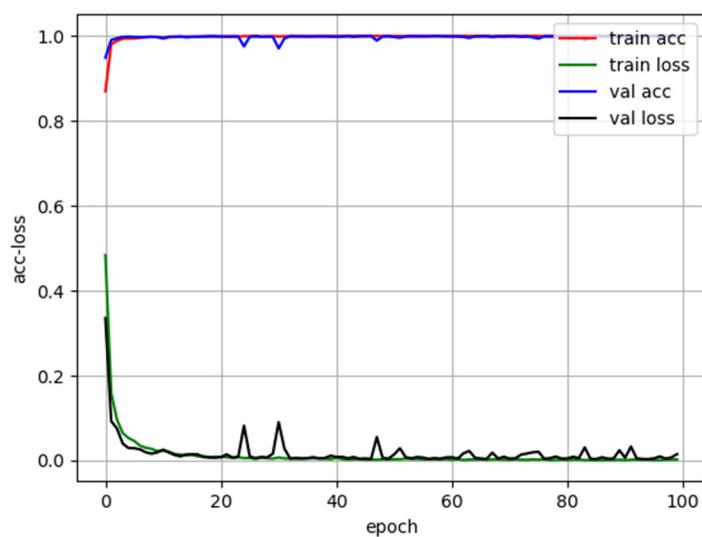


Figure 13. Evaluation of emergency rescue recognition DNN model.

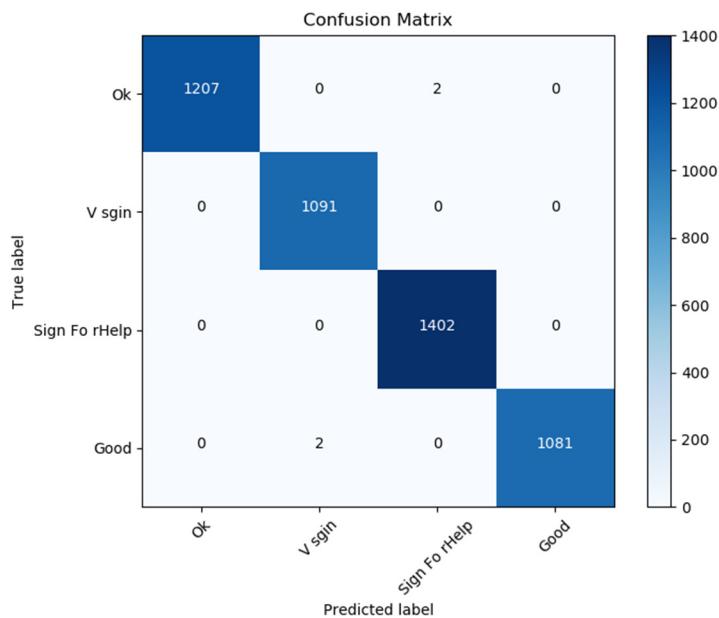


Figure 14. Confusion matrix with predicted labels on X-axis and true labels on the Y-axis tested in the testing set.

Figure 15 shows a demonstration of emergency rescue recognition. While the UAV was flying at an altitude of less than 10 m, we tried to simulate a field environment and performed a demonstration of emergency rescue recognition. It can be seen that different hand gestures are well predicted and, due to the high accuracy of the model, the model is very sensitive to the recognition of switching between different gestures, even dynamic sign for help gestures. This part of the function also assists the drone in its navigation tasks. When a human is following the route planned by the drone and suddenly encounters a situation where there is no road ahead or a better alternative route is available, then the human can interact with the drone in real-time and the drone will re-route for the human. The background environment in Figure 15 is a simulated field environment in the laboratory, where you can see the presence of some plants of about the same height as a human in the background, and object detection [61] was implemented in our previous work to give an early warning when a human is present. The first row of three images in Figure 15 shows the following normal static gestures: good, v-sign, and ok, which means

that everything is normal from the user's point of view, and we can see that the categories and probabilities of the gestures are well presented in the top right corner. The second row shows the dynamic emergency gesture (sign for help), where a warning is immediately given so that the drone can take the appropriate rescue measures, re-route for humans, or directly call a rescuer. It is important to note that the main reason we did not choose voice interaction here was due to the background of the paper, where the field environment combined with the bad weather and external noise was not conducive to verbal interaction. Therefore, gesture recognition is one of the best possible solutions. Later, UAV may interact with sharp sounds or projected phrase/signals on the ground.

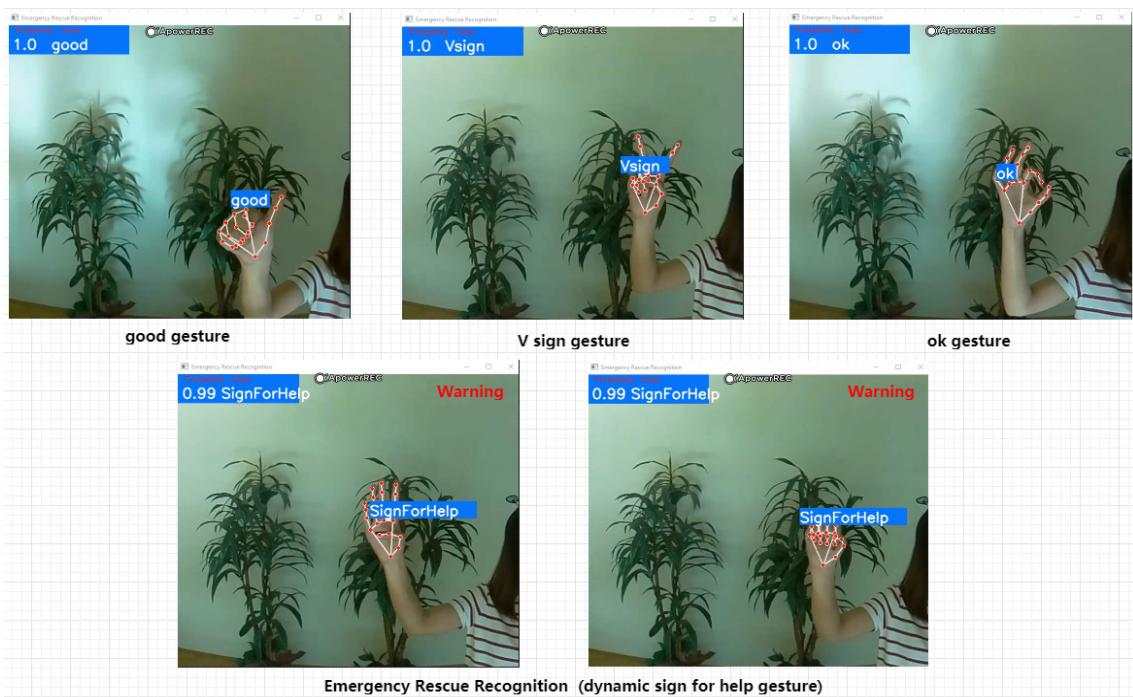


Figure 15. Demonstration of the results of simulated wilderness environments for emergency rescue recognition.

5. Discussion

The main concept of this work is to propose a road condition detection and emergency rescue recognition system for people in wilderness environments with poor network and bad weather conditions. The system utilizes the flexibility of UAVs and the collaboration between low altitude UAVs and high altitude satellites to provide the most timely assistance and monitor the best routes for humans in distress. Firstly, the system's upfront input is satellite data and weather data, and the real-time input is a video sequence captured by the drone camera. For drones equipped with multispectral camera sensors, in addition to the real-time video sequences, we can also access four channels of multispectral images. If the UAV does not have a multispectral sensor, then the detection of road material is completely dependent on the multispectral image information from the satellite. Next comes the data processing of the captured video sequences by the UAV. The UAV equipped with a GPU can segment the video sequences into image sequences in real-time, where RGB images are used as input for road extraction, and hyperspectral images are used as input for road material detection. The main backbone technologies of this work are as follows: the NDVI values are classified by multispectral images for the purpose of detecting the material composition of the road, i.e., bare soil road and non-bare soil gravel road can be well differentiated. The information on the different material compositions of the road surface is combined with the weather information entered in advance to estimate the speed of pedestrians walking on the different materials of the road surface. Based on the

yolo3-tiny model previously deployed on the UAV, the throughput of each road can be evaluated. D-LinkNet is the backbone technical support for road extraction and, based on the extracted roads, we can obtain geometric characteristics, such as the length and width of the different roads. Combining all that has been described, the pedestrians walking speed, the throughput of the road, the length and width of the road, and the connectivity are all used as the main basis for giving priority to the road. Finally, roads in the road network are given different priorities and the A^* algorithm with weights is applied to the starting and ending positions reported by the user for optimal route planning and ultimately for the purpose of navigation. As the optimal route planning is limited by the accuracy of the road extraction results, it is necessary for the UAV to adjust the flight altitude and have real-time human–UAV interaction with the user. Thus, a dataset for emergency rescue recognition was built and a reliable DNN model with high accuracy was trained. The model is primarily used for emergency rescue or to adjust navigation routes for the user.

The road condition detection in this work is different from the traditional road condition detection methods [71–74]. In this work, road condition detection is achieved by fusing a pre-trained road extraction model from a relevant dataset with NDVI information from different bands obtained from remote sensing information. Road extraction using D-LinkNet is then combined with advance weather information downloads to obtain the width of the road, the length of the road, the speed of pedestrians, the surface material of the road, and the throughput of the road, ultimately enabling the detection of road conditions. Most of the existing methods rely on various sensors, such as Wi-Fi, GPS, accelerometers, microphones, GSM antennae, in-vehicle standard sensors, laser sensors, stereo cameras, RGB cameras etc., which are usually deployed on ground vehicles, ground robots, or smartphones; however, the platform used as the data source for this paper is a drone with multispectral camera. The UAV platform is well-matched for aerial reconnaissance, and it has an unhindered large field of view. It allows for navigation through difficult terrains and facilitates safe and quick inspections. Some of the above-mentioned sensors will not work well in the wilderness without a network, and it is important to consider that these sensors do not affect the endurance of the drone once it is deployed. The endurance of a GPU-equipped UAV with an NVIDIA Jetson AGX Xavier developer kit, along with a Parrot Sequoia+ multispectral camera is around half an hour in our proposed system. With the development of artificial intelligence technology, the data sets collected by different sensors on different platforms are becoming more and more available, and it is in the area of autonomous driving that most applications in road detection are being made. Classification of road surface materials and the prediction of the damage to the road surface are being achieved with a high accuracy [75–79], but the above platforms are limited by a smaller field of view compared to drones, and the use cases are different.

The innovation of using drones for emergency rescue recognition at low altitude is the use of the latest dynamic gesture recognition technology for potential hazards (such as sign for help), avoiding interference from the external environment in the field. There are many options for using UAVs for emergency rescue, and vision technology is one of the best solutions to avoid interference from the drone rotors or external environmental noise to the voice recognition technology for search and rescue [80]. The emergency rescue recognition system proposed in this paper is a pre-trained model deployed on a UAV that runs in real-time which can then return the results to the ground base station as a projected signal or a sharp sound. There are many special and specific applications for drones for emergency rescue, such as searching for people [81], for man overboard rescue scenarios [82], etc. Most of these emergency rescue approaches incorporate machine learning methods to propose some new models with high accuracy from already existing datasets or use the collected datasets to train specified models to be deployed for use on UAVs. In this work, the core of the emergency rescue approach is to gain the attention of the drone through local dynamic gestures in order to help the user, such as in cases of re-routing. This dataset was collected in our lab and divided 7 to 3 into a training set and a testing set, respectively. The models

were trained on each of the four most-used classifiers, and the winning DNN model was finally deployed for use on the drone.

The limitations of this paper are discussed based on the experimental results of the whole system as follows. In the optimal path planning section, the user is required to actively report their starting position and destination to the UAV, a process that requires the UAV to fly at low altitude, and an interaction which must be completed through dynamic interactive gestures [30]. Due to the lack of real wilderness tests, the evaluation of the walking speed of pedestrians on different roads and the road throughput is not shown in the results section, but according to our previous work proposal [61], it is theoretically possible to count and track the number and location of people in real-time by deploying a yolo3-tiny model on the UAV. The road network in the UAV dataset used in this work is relatively simple. Of course, many publicly available UAV datasets with complex road networks already exist, but these are collected with cameras that do not use multispectral sensors on board, so they do not meet the needs of road surface material detection and classification. For satellite datasets, there are many and mostly free human-accessible databases, such as the sentinel-1 dataset and sentinel-2 dataset, which are publicly available and have many bands. It is worth noting that when selecting these datasets, we need to avoid datasets with cloud cover. The system proposed in this work is limited by the accuracy of road extraction techniques, which is the reason for creating the emergency rescue recognition dataset and training the highest accuracy model. Real-time interaction at low altitude is necessary when a human is following the route planned by the UAV and there is a special situation, such as there being no road ahead or a better alternative route available where the UAV can re-route for the user. Finally, because the endurance of the UAV proposed in this paper is limited (around half an hour), the battery life of the UAV should be well predicted and controlled and, if necessary, swarms of UAVs can also be added, with different UAVs performing different tasks to ensure the successful completion of road condition detection and emergency rescue recognition through cooperation.

The cooperation between drones and satellites is particularly important when extreme situations are encountered in the field, for example, when roads are flooded. Even if the drone is equipped with a multispectral camera, the satellite data can be downloaded in advance and used together. By downloading the satellite information in advance, we can obtain geographical information about the area, such as the terrain, road networks, rivers, etc. At the same time, research in related fields, such as water level detection by drones during floods [83] and flood detection [84], can be considered for deployment and application to drones. Similarly, reference can be made to water segmentation in very high-resolution aerial and satellite imagery [85], which indicates a problem at the same location if the results do not match the road and river information downloaded in advance from the satellite data. Solutions for river detection and flood detection can be found in our previous work [20] using the fusion MRF method. Humans in such extreme conditions will need to wait for rescue workers to be transported by boat.

Considering people from different cultures, we use hand gestures which have general meanings, so that in each context the human can communicate receiving the feedback (e.g., audio message or projected text on the surface) from the UAV board. In short, the results that can be achieved and applied in this work are the detection of road conditions and the accurate recognition of emergency signals by UAVs for human-machine interaction.

6. Conclusions

The main contributions and novelties of this paper are as following. Firstly, the information provided by the UAV itself or by the combination of both UAV and satellite enables the detection of road surface materials and use multispectral information to differentiate between different road types (bare soil roads and non-bare soil roads). Secondly, the walking speed of pedestrians on different roads was estimated in combination with the constituent materials of the road surface and weather information, and the throughput of the road was also assessed, which complements and theoretically supports our previous road weights

assignment work [48]. Thirdly, drones flying at high altitudes can plan the best route for humans in distress; drones flying at low altitudes can communicate freely with humans on the ground in real-time and provide accurate recognition of emergency rescue information. Drones maximize the multi-faceted ways in which they can help users in distress. Finally, the information provided by low-altitude drones is fused with hyperspectral image information provided by high-altitude satellites to provide humans with a wider range of search and rescue information. The main challenges have been solved, as follows:

- Drones can plan the optimal routes in real-time for humans in the wilderness with poor network conditions and bad weather;
- Different road types are well differentiated, the walking speed of pedestrians is estimated, and the throughput of the road can theoretically be evaluated;
- In the road extraction network map, the priority of the different roads is given on a sufficient basis, which refines our previous work;
- Here, UAVs flying at low latitudes can perform human–machine interaction tasks very accurately.

In short, the novel features of this proof-of-concept paper are as follows:

- Estimating the soil type from multispectral information;
- Estimating the road quality from the weather and soil information;
- Finding the paths/roads on the terrain;
- Weighting the paths by walking/driving quality measure;
- Finding the optimal routes based on the weighted route-map.

This paper has shown a tool-set for a future emergency related situation. In future work, if emergency rescue situation simulation is allowed, we will test this system in a field environment. We will also try to collect UAV multispectral datasets with complex road networks. Finally, for the A^* algorithm of path planning, we will try to use and compare other related path planning search algorithms. Different algorithms for route search with weights will be further compared and optimized.

Author Contributions: C.L. designed and implemented the whole system and wrote and edited the paper. She also created the emergency rescue dataset with her colleagues. T.S. had the initial idea of this project and worked on the paper, who supervised and guided the whole project. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Ministry of Innovation and Technology NTDI Office within the framework of the Autonomous Systems National Laboratory Program and by the Hungarian National Science Foundation (NKFIH OTKA) No. K139485. This research was also funded by Stipendium Hungaricum scholarship and China Scholarship Council.

Data Availability Statement: The Satellite dataset can be downloaded from this link <https://earthexplorer.usgs.gov/> (accessed on 1 May 2022) and the UAV dataset presented in this paper are available on request from the authors.

Acknowledgments: The work is carried out at Institute for Computer Science and Control (SZTAKI), Hungary and the authors would like to thank their colleague László Spórás for the technical support and for the helpful scientific community of Machine Perception Research Laboratory.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mavridou, E.; Vrochidou, E.; Papakostas, G.A.; Pachidis, T.; Kaburlasos, V.G. Machine vision systems in precision agriculture for crop farming. *J. Imaging* **2019**, *5*, 89. [[CrossRef](#)] [[PubMed](#)]
2. Song, H.; Liang, H.; Li, H.; Dai, Z.; Yun, X. Vision-based vehicle detection and counting system using deep learning in highway scenes. *Eur. Transp. Res. Rev.* **2019**, *11*, 51. [[CrossRef](#)]
3. Varga, D.; Szirányi, T. Robust real-time pedestrian detection in surveillance videos. *J. Ambient Intell. Humaniz. Comput.* **2017**, *8*, 79–85. [[CrossRef](#)]
4. Panahi, L.; Ghods, V. Human fall detection using machine vision techniques on RGB–D images. *Biomed. Signal Process. Control* **2018**, *44*, 146–153. [[CrossRef](#)]

5. Kanellakis, C.; Nikolakopoulos, G. Survey on computer vision for UAVs: Current developments and trends. *J. Intell. Robot. Syst.* **2017**, *87*, 141–168. [[CrossRef](#)]
6. Kashino, Z.; Nejat, G.; Benhabib, B. Aerial wilderness search and rescue with ground support. *J. Intell. Robot. Syst.* **2020**, *99*, 147–163. [[CrossRef](#)]
7. Alsamhi, S.H.; Almalki, F.A.; AL-Dois, H.; Shvetsov, A.V.; Ansari, M.S.; Hawbani, A.; Gupta, S.K.; Lee, B. Multi-drone edge intelligence and SAR smart wearable devices for emergency communication. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 6710074. [[CrossRef](#)]
8. Heggie, T.W.; Heggie, T.M. Dead men hiking: Case studies from the American wilderness. *Med. Sport.* **2012**, *16*, 118–121. [[CrossRef](#)]
9. Mishra, B.; Garg, D.; Narang, P.; Mishra, V. Drone-surveillance for search and rescue in natural disaster. *Comput. Commun.* **2020**, *156*, 1–10. [[CrossRef](#)]
10. Alsamhi, S.H.; Shvetsov, A.V.; Kumar, S.; Shvetsova, S.V.; Alhartomi, M.A.; Hawbani, A.; Rajput, N.S.; Srivastava, S.; Saif, A.; Nyangaresi, V.O. UAV Computing-Assisted Search and Rescue Mission Framework for Disaster and Harsh Environment Mitigation. *Drones* **2022**, *6*, 154. [[CrossRef](#)]
11. Harris, R. *Satellite Remote Sensing—An Introduction*; Routledge Kegan & Paul: London, UK, 1987.
12. Patino, J.E.; Duque, J.C. A review of regional science applications of satellite remote sensing in urban settings. *Comput. Environ. Urban Syst.* **2013**, *37*, 1–17. [[CrossRef](#)]
13. Lo, C. *Applied Remote Sensing*; Taylor & Francis: Abingdon, UK, 1986.
14. Zhu, L.; Suomalainen, J.; Liu, J.; Hyppä, J.; Kaartinen, H.; Haggren, H. *A Review: Remote Sensing Sensors—Multi-Purposeful Application of Geospatial Data*; IntechOpen: London, UK, 2018; pp. 19–42.
15. Karthikeyan, L.; Chawla, I.; Mishra, A.K. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *J. Hydrol.* **2020**, *586*, 124905. [[CrossRef](#)]
16. Yang, X.; Qin, Q.; Grussenmeyer, P.; Koehl, M. Urban surface water body detection with suppressed built-up noise based on water indices from Sentinel-2 MSI imagery. *Remote Sens. Environ.* **2018**, *219*, 259–270. [[CrossRef](#)]
17. Özelen, E. Water body detection analysis using NDWI indices derived from landsat-8 OLI. *Polish J. Environ. Stud.* **2020**, *29*, 1759–1769. [[CrossRef](#)]
18. Yuan, K.; Zhuang, X.; Schaefer, G.; Feng, J.; Guan, L.; Fang, H. Deep-learning-based multispectral satellite image segmentation for water body detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7422–7434. [[CrossRef](#)]
19. Dang, B.; Li, Y. MSResNet: Multiscale residual network via self-supervised learning for water-body detection in remote sensing imagery. *Remote Sens.* **2021**, *13*, 3122. [[CrossRef](#)]
20. Sziranyi, T.; Shadaydeh, M. Segmentation of remote sensing images using similarity-measure-based fusion-MRF model. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1544–1548. [[CrossRef](#)]
21. Talukdar, S.; Singha, P.; Mahato, S.; Pal, S.; Liou, Y.A.; Rahman, A. Land-use land-cover classification by machine learning classifiers for satellite observations—A review. *Remote Sens.* **2020**, *12*, 1135. [[CrossRef](#)]
22. Castelluccio, M.; Poggi, G.; Sansone, C.; Verdoliva, L. Land use classification in remote sensing images by convolutional neural networks. *arXiv* **2015**, arXiv:1508.00092.
23. Peng, D.; Zhang, Y.; Guan, H. End-to-end change detection for high resolution satellite images using improved UNet++. *Remote Sens.* **2019**, *11*, 1382. [[CrossRef](#)]
24. Asokan, A.; Anitha, J.J.E.S.I. Change detection techniques for remote sensing applications: A survey. *Earth Sci. Inform.* **2019**, *12*, 143–160. [[CrossRef](#)]
25. Szirányi, T.; Zerubia, J. Multilayer Markov Random Field Models for Change Detection in Optical Remote Sensing Images. *ISPRS J. Photogramm. Remote Sens.* **2015**, *107*, 22–37.
26. Li, J.; Pei, Y.; Zhao, S.; Xiao, R.; Sang, X.; Zhang, C. A review of remote sensing for environmental monitoring in China. *Remote Sens.* **2020**, *12*, 1130. [[CrossRef](#)]
27. Laurance, W.F.; Clements, G.R.; Sloan, S.; O'connell, C.S.; Mueller, N.D.; Goosem, M.; Venter, O.; Edwards, D.; Phalan, B.; Balmford, A.; et al. A global strategy for road building. *Nature* **2014**, *513*, 229–232. [[CrossRef](#)] [[PubMed](#)]
28. Ciepluch, B.; Jacob, R.; Mooney, P.; Winstanley, A.C. Comparison of the accuracy of OpenStreetMap for Ireland with Google Maps and Bing Maps. In Proceedings of the Ninth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester, UK, 20–23 July 2010; p. 337.
29. Wang, W.; Yang, N.; Zhang, Y.; Wang, F.; Cao, T.; Eklund, P. A review of road extraction from remote sensing images. *J. Traffic Transp. Eng.* **2016**, *3*, 271–282. [[CrossRef](#)]
30. Liu, C.; Szirányi, T.A. Gesture Recognition for UAV-based Rescue Operation based on Deep Learning. In Proceedings of the International Conference on Image Processing and Vision Engineering (IMPROVE 2021), Anchorage, AL, USA, 19–22 September 2021; pp. 180–187.
31. Jetson AGX Xavier Developer Kit. NVIDIA Developer. 2018. Available online: <https://developer.nvidia.com/embedded/jetson-agx-xavier-developer-kit> (accessed on 4 July 2022).
32. Parrot Sequoia+. SenseFly. 2018. Available online: <https://www.sensefly.com/camera/parrot-sequoia/> (accessed on 4 July 2022).
33. Hossain, S.; Lee, D.-J. Deep learning-based real-time multiple-object detection and tracking from aerial imagery via a flying robot with GPU-based embedded devices. *Sensors* **2019**, *19*, 3371. [[CrossRef](#)]

34. Esa.int. ESA—Home. 2019. Available online: <https://www.esa.int/> (accessed on 4 July 2022).
35. USGS. Science for a Changing World. Available online: <https://www.usgs.gov/> (accessed on 4 July 2022).
36. Mansouri, S.S.; Karvelis, P.; Georgoulas, G.; Nikolakopoulos, G. Remaining useful battery life prediction for UAVs based on machine learning. *IFAC Pap.* **2017**, *50*, 4727–4732. [CrossRef]
37. Saif, A.; Dimyati, K.; Noordin, K.A.; Shah, N.S.M.; Alsamhi, S.H.; Abdullah, Q. August. Energy-efficient tethered UAV deployment in B5G for smart environments and disaster recovery. In Proceedings of the 2021 1st International Conference on Emerging Smart Technologies and Applications (eSmarTA), Sana'a, Yemen, 10–12 August 2021; IEEE: New York, NY, USA, 2021; pp. 1–5.
38. Wikipedia Contributors. Biatorbágы. Wikipedia, Wikimedia Foundation. 2021. Available online: <https://en.wikipedia.org/wiki/Biatorb%C3%A1gy> (accessed on 5 July 2022).
39. Google. Google Maps. 2022. Available online: www.google.com/maps (accessed on 5 July 2022).
40. USGS EROS Archive—Commercial Satellites—OrbView 3. U.S. Geological Survey. Available online: www.usgs.gov/centers/eros/science/usgs-eros-archive-commercial-satellites-orbview-3 (accessed on 5 July 2022).
41. Birdwood. Wikipedia. 2021. Available online: <https://en.wikipedia.org/wiki/Birdwood> (accessed on 5 July 2022).
42. Mnih, V. *Machine Learning for Aerial Image Labeling*; University of Toronto: Toronto, ON, Canada, 2013.
43. Demir, I.; Koperski, K.; Lindenbaum, D.; Pang, G.; Huang, J.; Basu, S.; Hughes, F.; Tuia, D.; Raskar, R. Deepglobe 2018: A challenge to parse the earth through satellite images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Lake City, UT, USA, 18–22 June 2018; pp. 172–181.
44. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Singapore, 18–22 September 2015; Springer: Cham, Switzerland, 2015; pp. 234–241.
45. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]
46. Chaurasia, A.; Culurciello, E. Linknet: Exploiting encoder representations for efficient semantic segmentation. In Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA, 10–13 December 2017; IEEE: New York, NY, USA, 2017; pp. 1–4.
47. Zhou, L.; Zhang, C.; Wu, M. D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Lake City, UT, USA, 18–22 June 2018; pp. 182–186.
48. Liu, C.; Szirányi, T. UAV Path Planning based on Road Extraction. In Proceedings of the International Conference on Image Processing and Vision Engineering (IMPROVE 2021), Brussels, Belgium, 16–17 June 2022; pp. 202–210.
49. Goto, T.; Kosaka, T.; Noborio, H. On the heuristics of A* or A algorithm in ITS and robot path-planning. In Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No. 03CH37453), Las Vegas, NV, USA, 27–31 October 2003; IEEE: New York, NY, USA, 2003; Volume 2, pp. 1159–1166.
50. Abhishek, K.; Singh, M.; Ghosh, S.; Anand, A. Weather forecasting model using artificial neural network. *Procedia Technol.* **2012**, *4*, 311–318. [CrossRef]
51. Running, S.W. Estimating terrestrial primary productivity by combining remote sensing and ecosystem simulation. In *Remote Sensing of Biosphere Functioning*; Springer: New York, NY, USA, 1990; pp. 65–86.
52. Myneni, R.B.; Hall, F.G.; Sellers, J.; Marshak, A.L. The interpretation of spectral vegetation indexes. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 481–486. [CrossRef]
53. Papageorgiou, M.; Diakaki, C.; Dinopoulou, V.; Kotsialos, A.; Wang, Y. Review of road traffic control strategies. *Proc. IEEE* **2003**, *91*, 2043–2067. [CrossRef]
54. Definition, Interpretation, and Calculation of Traffic Analysis Tools Measures of Effectiveness—6.0 Recommended MOEs. Available online: <https://ops.fhwa.dot.gov/publications/fhwahop08054/sect6.htm> (accessed on 1 June 2022).
55. Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.M.; Tucker, C.J.; Stenseth, N.C. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* **2005**, *20*, 503–510. [CrossRef] [PubMed]
56. Gupta, V.D.; Areendran, G.; Raj, K.; Ghosh, S.; Dutta, S.; Sahana, M. Assessing habitat suitability of leopards (*Panthera pardus*) in unprotected scrublands of Bera, Rajasthan, India. In *Forest Resources Resilience and Conflicts*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 329–342.
57. Kraetzig, N.M. 5 Things to Know about NDVI (Normalized Difference Vegetation Index). UP42 Official Website. Available online: <https://up42.com/blog/tech/5-things-to-know-about-ndvi#:~{:text=The%20value%20of%20the%20NDVI}> (accessed on 6 July 2022).
58. UW-Madison Satellite Meteorology. Available online: <https://profhorn.meteor.wisc.edu/wxwise/satmet/lesson3/ndvi.html> (accessed on 6 July 2022).
59. Gast, K.; Kram, R.; Riemer, R. Preferred walking speed on rough terrain: Is it all about energetics? *J. Exp. Biol.* **2019**, *222*, jeb185447. [CrossRef] [PubMed]
60. Mohamed, O.; Appling, H. Clinical assessment of gait. *Orthot. Prosthet. Rehabil.* **2020**, *4*, 102–144.
61. Liu, C.; Szirányi, T. Real-time human detection and gesture recognition for on-board UAV rescue. *Sensors* **2021**, *21*, 2180. [CrossRef]
62. Signal for Help. Wikipedia. 2020. Available online: https://en.wikipedia.org/wiki/Signal_for_Hel (accessed on 1 July 2021).
63. Mediapipe. Hands. Available online: <https://google.github.io/mediapipe/solutions/hands.html> (accessed on 1 June 2021).

64. Zhang, F.; Bazarevsky, V.; Vakunov, A.; Tkachenka, A.; Sung, G.; Chang, C.L.; Grundmann, M. Mediapipe hands: On-device real-time hand tracking. *arXiv* **2020**, arXiv:2006.10214.
65. Wright, R.E. *Logistic Regression*; APA: Washington, DC, USA, 1995.
66. Singh, A.; Prakash, B.S.; Chandrasekaran, K. A comparison of linear discriminant analysis and ridge classifier on Twitter data. In Proceedings of the 2016 International Conference on Computing, Communication and Automation (ICCCA), Greater Noida, India, 29–30 April 2016; IEEE: New York, NY, USA, 2016; pp. 133–138.
67. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [[CrossRef](#)]
68. Natekin, A.; Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurorobot.* **2013**, *7*, 21. [[CrossRef](#)]
69. Canziani, A.; Paszke, A.; Culurciello, E. An analysis of deep neural network models for practical applications. *arXiv* **2016**, arXiv:1605.07678.
70. Carreira, J.; Zisserman, A. Quo vadis, action recognition? A new model and the kinetics dataset. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Juan, PR, USA, 17–19 June 1997.
71. Chugh, G.; Bansal, D.; Sofat, S. Road condition detection using smartphone sensors: A survey. *Int. J. Electron. Electr. Eng.* **2014**, *7*, 595–602.
72. Castillo Aguilar, J.J.; Cabrera Carrillo, J.A.; Guerra Fernández, A.J.; Carabias Acosta, E. Robust road condition detection system using in-vehicle standard sensors. *Sensors* **2015**, *15*, 32056–32078. [[CrossRef](#)] [[PubMed](#)]
73. Jokela, M.; Kutila, M.; Le, L. Road condition monitoring system based on a stereo camera. In Proceedings of the 2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing, Cluj-Napoca, Romania, 27–29 August 2009; IEEE: New York, NY, USA, 2009; pp. 423–428.
74. Ranyal, E.; Sadhu, A.; Jain, K. Road condition monitoring using smart sensing and artificial intelligence: A review. *Sensors* **2022**, *22*, 3044. [[CrossRef](#)] [[PubMed](#)]
75. Xie, Q.; Hu, X.; Ren, L.; Qi, L.; Sun, Z. A Binocular Vision Application in IoT: Realtime Trustworthy Road Condition Detection System in Passable Area. In *IEEE Transactions on Industrial Informatics*; IEEE: New York, NY, USA, 2022.
76. Gupta, A.; Anpalagan, A.; Guan, L.; Khwaja, A.S. Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. *Array* **2021**, *10*, 100057. [[CrossRef](#)]
77. Chun, C.; Ryu, S.K. Road surface damage detection using fully convolutional neural networks and semi-supervised learning. *Sensors* **2019**, *19*, 5501. [[CrossRef](#)] [[PubMed](#)]
78. Wang, D.; Liu, Z.; Gu, X.; Wu, W.; Chen, Y.; Wang, L. Automatic Detection of Pothole Distress in Asphalt Pavement Using Improved Convolutional Neural Networks. *Remote Sens.* **2022**, *14*, 3892. [[CrossRef](#)]
79. Rateke, T.; Justen, K.A.; Von Wangenheim, A. Road surface classification with images captured from low-cost camera-road traversing knowledge (rtk) dataset. *Rev. De Inf. Teórica E Apl.* **2019**, *26*, 50–64. [[CrossRef](#)]
80. Yamazaki, Y.; Tamaki, M.; Premachandra, C.; Perera, C.J.; Sumathipala, S.; Sudantha, B.H. Victim detection using UAV with on-board voice recognition system. In Proceedings of the 2019 Third IEEE International Conference on Robotic Computing (IRC), Naples, Italy, 25–27 February 2019; IEEE: New York, NY, USA, 2019; pp. 555–559.
81. Castellano, G.; Castiello, C.; Mencar, C.; Vessio, G. Preliminary evaluation of TinyYOLO on a new dataset for search-and-rescue with drones. In Proceedings of the 2020 7th International Conference on Soft Computing & Machine Intelligence (ISCFMI), Stockholm, Sweden, 14–15 November 2020; IEEE: New York, NY, USA, 2020; pp. 163–166.
82. Cafarelli, D.; Ciampi, L.; Vadicamo, L.; Gennaro, C.; Berton, A.; Paterni, M.; Benvenuti, C.; Passera, M.; Falchi, F. MOBDrone: A Drone Video Dataset for Man OverBoard Rescue. In Proceedings of the International Conference on Image Analysis and Processing, Bangkok, Thailand, 16–17 November 2022; Springer: Cham, Switzerland, 2022; pp. 633–644.
83. Rizk, H.; Nishimur, Y.; Yamaguchi, H.; Higashino, T. Drone-based water level detection in flood disasters. *Int. J. Environ. Res. Public Health* **2021**, *19*, 237. [[CrossRef](#)]
84. Tanim, A.H.; McRae, C.; Tavakol-Davani, H.; Goharian, E. Flood Detection in Urban Areas Using Satellite Imagery and Machine Learning. *Water* **2022**, *14*, 1140. [[CrossRef](#)]
85. Zhang, Z.; Lu, M.; Ji, S.; Yu, H.; Nie, C. Rich CNN Features for water-body segmentation from very high-resolution aerial and satellite imagery. *Remote. Sens.* **2021**, *13*, 1912. [[CrossRef](#)]