

# A VERSATILE UAV NEAR REAL-TIME MAPPING SOLUTION FOR DISASTER RESPONSE – CONCEPT, IDEAS AND IMPLEMENTATION

Phillipp Fanta-Jende\*, Daniel Steininger, Felix Bruckmüller and Christoph Sulzbachner

Center for Vision, Automation and Control – Assistive & Autonomous Systems, AIT – Austrian Institute of Technology, Vienna, Austria – {firstname.name}@ait.ac.at

## Commission I/II, WG: UAS and Small Multi-sensor Platforms: Concepts and Applications (I/II)

**KEY WORDS:** UAV, disaster response, design concept, orthorectification, semantic segmentation

### ABSTRACT:

In recent years, the proliferation and further development of unmanned aerial vehicles (UAVs) led to a great number of key technologies, advances and opportunities especially in the realm of time-critical applications. UAVs as a platform provide a unique combination of flexibility, affordability and sensor technology which enables the design of cost-effective and intriguing services particularly for disaster response. This contribution presents a concept for UAV-based near real-time mapping system for disaster relief to provide decision-making support for first responders particularly for possible disaster scenarios in Austria. We outline our system concept and its respective architecture, discuss requirements from a stakeholder perspective as well as legal regulations and initiatives at an EU level. In the methodology section of this paper, the preliminary data processing pipeline with respect to the near real-time orthomosaic generation and the semantic segmentation network are presented. Lastly, first experimental results of the pipeline are shown, and further advances are discussed.

## 1. INTRODUCTION

UAVs have proven to be an important instrument for various disaster scenarios (Alamouri et al., 2019; Erdelj et al., 2017). Unlike traditional geodata acquisitions techniques, such as aerial photogrammetry or satellite surveys, UAVs feature a considerably more agile operational readiness allowing for immediate on-site data capture, processing and visualisation which ultimately expedites decision-making support for first responders.

From a technical point of view, these merits pose certain trade-offs or rather compromises. Requirements and limitations on image orientation accuracy, image resolution as well as processing performance let alone the derivation of tertiary products such as semantic labels need to be well balanced for such application-driven developments.

To this end, we design the system with the intention to share the processing load between an air and a ground segment. While the UAV platform itself is equipped with RGB as well as thermal infrared sensors, positioning instruments and onboard processing units, the ground segment features high-performance processing capabilities for more demanding tasks.

Analogously, the algorithmic workflow varies based on the mapping scenario at hand. At its core, the workflow performs image orientation, sparse reconstruction and orthomosaic generation in near real-time using the onboard processing capacities of the UAV.

For additional information on the situation, further processing is conducted successively and in parallel on the ground station. These processes are usually computationally expensive and are thus performed on the ground station.

## 2. RELATED WORK

UAVs have become a substantial asset for various disaster scenarios, such as monitoring floods (Luo et al., 2015; Zhang et al., 2016), assessing the damages caused by earthquakes (Duarte et al., 2017; Kerle et al., 2019; N. Kerle et al., 2019; Suzuki et

al., 2008), mapping landslides (Lindner et al., 2016; Rossi et al., 2018; Tanteri et al., 2017) or searching missing persons (Miyano et al., 2019; Silvagni et al., 2017; Waharte and Trigoni, 2010).

Depending on the disaster scenario, different data and respective processing steps are required to support first responders. In general, an orthomosaic, i.e. a georeferenced and geometrically rectified composition of images, is a helpful tool to provide an overview of the situation and can serve as the basis for further analyses. Depending on the system setup, i.e. processing capabilities of the UAV, data downlink etc., and time criticality of the disaster event, either offline or online processing comes into play.

In the case of offline processing, the UAV only collects data and needs to land for data transmission and processing. The standard photogrammetric workflow or structure from motion techniques can be used to reconstruct the scene for ortho image generation (Ghosh and Kaabouch, 2016; Saeed Yahyanejad, 2013; Schönberger and Frahm, 2016). For this kind of workflow, various commercial solutions are available<sup>1</sup>.

Disaster events, however, usually require information on the situation as fast as possible. To this end, the data processing chain needs to be (near) real-time. Depth reconstruction, however, is a computationally expensive component in the orthomosaic generation procedure.

Bu et al. (2016) circumvent this step by utilising projective transformations, i.e. homographies, to warp acquired images onto a virtual ground plane. Although this procedure is effective in rather flat and consistent terrain, it inevitably leads to geometric distortions in the resulting orthomosaic in mountainous or urban areas. Alternatively, Hein et al. (2019) resort to SRTM data<sup>2</sup> with a ground sampling distance of 90

<sup>1</sup> Drone Deploy, <https://www.dronedeploy.com/>, Pix4D, <https://www.pix4d.com/>, particularly Pix4D React, <https://www.pix4d.com/product/pix4dreact>, Agisoft, <https://www.agisoft.com/>

<sup>2</sup> <https://www2.jpl.nasa.gov/srtm/>

metres to provide height information for the orthorectification process of acquired UAV images. Certainly, this rather low resolution results in inaccuracies and distortions in the orthomosaic similar to the aforementioned procedure if applied in areas of uneven terrain.

Another noteworthy method is to reconstruct the scene using stereo cameras mounted on a UAV (Fan et al., 2019). This method, however, is only applicable for close-range mapping since the base of the cameras is physically limited.

To achieve a stereo base wide enough for typical flying altitudes for large-area UAV mapping (100 m or more), virtual stereo pairs can be used (Hinzmann et al., 2018). In this case, using an efficient stereo block matching algorithm enables rapid dense reconstruction as well as orthomosaic generation.

The method used in the experiments in this paper is based on the works of Kern (2018) and Bobbe et al. (2017). The authors utilise ORB-SLAM2 (Mur-Artal and Tardos, 2016) to compute a sparse point cloud of the scene which is then densified by interpolation or using a plane-sweeping algorithm.

To understand the content of the images, semantic segmentation can be used. It provides an assignment of specific class labels to input images on a per-pixel basis. This task has especially benefited from recent advances in machine learning and GPU parallelisation. Especially the latter facilitated the application of Convolutional Neural Networks (CNNs) to complex learning problems (Chen et al., 2017; Long et al., 2014). Since semantic segmentation requires large amounts of training data, another import prerequisite was the increasing development of densely label datasets in multiple application domains (Cordts et al., 2016).

### 3. CONCEPT AND BACKGROUND

Various crisis and disaster scenarios, such as natural disasters, industrial accidents, searches of persons, leakage of pollutants or mass movements, benefit from the use of airborne systems. In most cases, critical security situations require immediately available and detailed situation information on a large scale.

From a stakeholder perspective, the system needs to be designed fit-for-purpose, i.e. satisfies the requirements of completeness, readiness, usability, coherence and reliability. Since various disaster scenarios need to be accounted for, different algorithms need be used to derive the information required. Consequently, the design has to be modular and ideally extendable.

This is especially true regarding the scene understanding part of our system. For instance, the suppression of forest fires requires the mapping of hot spot locations with implicit change detection mechanics. In this case, thermal imaging is incorporated, and the UAV's flight pattern is adjusted accordingly to map the area in constant intervals. Floods on the other hand pose different challenges, such as reliable information on the trafficability in the area. To this end, semantic segmentation can be helpful to understand which areas are affected by flooding and which areas are traversable for disaster relief teams.

Whereas certain data products (i.e. orthomosaic and surface model) are always obtained and constitute the fundamental basis of this system and further analyses, other processing methods or sensors (e.g. thermal) are optional.

Figure 1 illustrates the design concept of our system. The separation into air and ground segment relates to hard- and software components. The air segment's priority is to provide information as fast as possible and relay that information and raw data to the ground segment where further task-related analysis and the respective visualisation takes place.

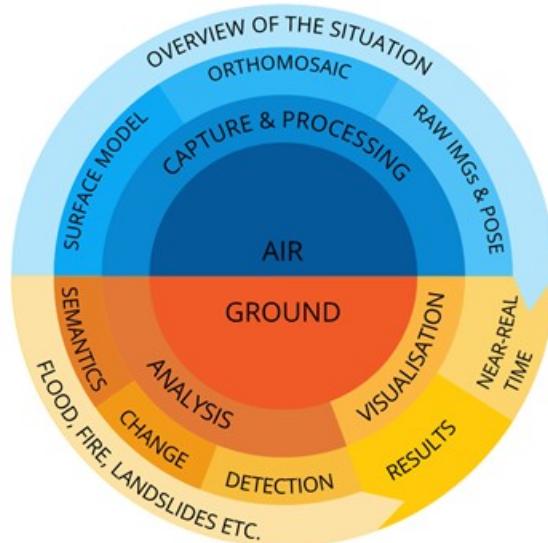


Figure 1. UAV mapping design scheme

#### 3.1 Communication link

To connect both parts efficiently, a communication link tightly couples both segments. For our system, the microhard pDDL2450 Wireless OEM Ethernet & Serial Digital Data Link<sup>3</sup> provides a direct radio link between the air and ground segment. Using a ground-based tracking antenna, it provides an ethernet interface with a bandwidth of 20 Mbit over up to 10 km.

Alternatively, a VPN connection between an onboard processing unit and a ground processing unit utilising a 3G/4G/5G connection on both ends can be used. However, using mobile network connections requires working mobile internet infrastructure in the area of operation.

#### 3.2 Air segment

The air segment consists of an UAV in combination with a sensor package, a processing unit and a communication link. At AIT, we develop two different fixed-wing UAVs for long range aerial mapping.

The smaller development platform is based on the commercial off-the-shelf product Skywalker EVE-2000<sup>4</sup> and serves solely as a testing platform for interface and algorithm design (see Figure 2). With a wingspan of 2.24 m and a flight ready weight of 7 kg, it reaches a flight time of 30 minutes. It features a single RGB camera facing downwards in combination with a single processing unit based on a Nvidia Jetson TX2<sup>5</sup>.



Figure 2. Skywalker EVE-2000

<sup>3</sup> <http://www.microhardcorp.com/pDDL.php>

<sup>4</sup> <http://skywalkermodel.com/en/89.html>

<sup>5</sup> <https://developer.nvidia.com/embedded/jetson-tx2>

The second platform is based on an airframe developed in conjunction with the Institute of Aviation at the FH Joanneum Graz. With a wingspan of 3.8 m and a flight ready weight of up to 30 kg, it offers the possibility to carry multiple different sensors and one or more onboard processing units based on Nvidia Jetson TX2, Nvidia Jetson Xavier or NUC sized mini PCs (see Figure 3). The airframe is optimised for stall speeds of less than 15 m/s which is desirable for on-board data processing since a certain image overlap is required for 3D reconstruction of the terrain. The modular payload bay allows for a flexible integration of various sensors including RGB/near-infrared/thermal-infrared cameras, radar-sensors or lidar-sensors. Depending on the payload, a flight time between 20 and 50 minutes is achievable.

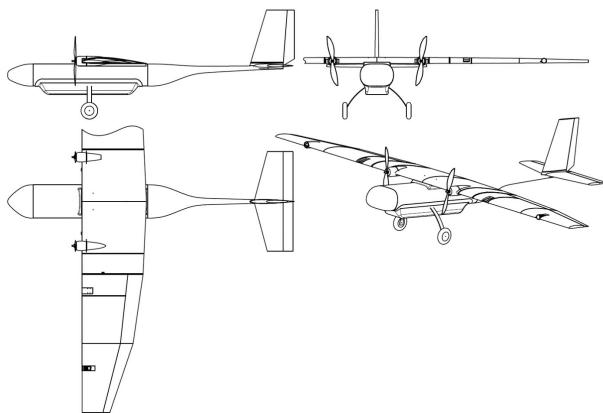


Figure 3. Blueprint of larger UAV

Both platforms are controlled by the onboard flight controller Pixhawk 2<sup>6</sup> running the Arduplane flight control stack<sup>7</sup>. This setup offers either manual control using an RC radio/receiver pair, in this case an Fr-Sky Taranis<sup>8</sup>, manual control over the onboard processing unit or autonomous flight controlled by the flight controller itself or the onboard processing unit. Communication between the flight controller and the onboard processing unit is based on the MAVLINK communication protocol<sup>9</sup> using a serial interface connection. The state estimation and sensor fusion are conducted by an Extended Kalman Filter.

All onboard and ground processing units communicate using ROS<sup>10</sup> utilising a simple network connection. Moreover, ROS assigns a timestamp to all sensor readings which simplifies the synchronisation of all data inputs for further processing.

### 3.3 Ground segment

The ground segment consists of a high-performance processing unit and a visualisation unit. Data from the air segment is processed by a 24-core Intel Xeon server with 384 GB of RAM and four Nvidia GeForce RTX 2080 TI graphic cards. This hardware allows for running advanced machine learning algorithms in order to provide disaster-specific scene analyses. A custom visualisation system offers easy to use inspection of generated maps and data in addition to information generated by the scene analyses for first responders.

### 3.4 Legal regulations

In 2008, a regulation was introduced to ensure that drones are safely integrated into European airspace. The regulation establishes common safety rules for civil aviation and amends the mandate of the European Aviation Safety Agency (EASA), replacing the regulatory framework of 2008. After 2008, EU Member States were responsible for the regulatory approach for drones of up to 150 kg, which led to a fragmented regulatory framework within the EU Member States. In 2019, the European Commission adopted EU-wide rules on technical requirements for drones (EASA, 2015a, 2015b). These regulations are classified into three categories: open, specific and certified with different safety requirements appropriate to the risk (The European Commission, 2019). According to the EASA roadmap, the EU-wide regulations will be implemented in the EU Member States in July 2020 (EASA, 2019).

For risk assessment, a Specific Risk Operation Assessment (SORA) is provided, which is a guideline for the operation of an unmanned aerial vehicle system according to a specific operational concept (JARUS, 2019). In this context, risk is understood to be the combination of the probability of an event and its associated severity. Safety is defined as a condition in which the risk is considered acceptable. The risk on the ground and in the air shall be mitigated to an acceptable level by an appropriate combination of design and operational means of mitigation. These mitigations have to meet a level of robustness corresponding to the established risk classes on the ground and in the air. The level corresponds to an appropriate combination of integrity and safety levels. The integrity level is the safety gain achieved by the mitigation and the reliability level is the method to demonstrate that the integrity level has been achieved.

In previous research projects, a preliminary multi-stage risk assessment for beyond-line-of-sight operation was completed and our custom platform was certified according to the latest harmonised European regulations.

## 4. METHODOLOGY

Although the system is currently under development and many methods and components are yet to be developed and determined, a brief overview of the functionality of the orthomosaic generation as well as exemplary scene analysis approaches is given.

### 4.1 Orthomosaic generation

The orthomosaic generation is an essential part of our system. First, it is designed to work in (near) real-time to provide an overview of the disaster situation as fast as possible. Second, the approach implicitly georeferences all image data and serves as a basis for further analysis. Third, the orthomosaic provides a base layer for all subsequent data visualisations.

The underlying framework is based on ROS, where all sensor readings, i.e. GNSS/INS, images etc., are synchronised and so-called sensor messages can be easily fed into an arbitrary number of other processes during runtime.

As mentioned earlier, first experiments with respect to the orthomosaic generation have been conducted using the works of Kern (2018) and Bobbe et al. (2017). Their method is also based on ROS for the communication, i.e. data transfer, synchronisation etc. between the different steps of the orthomosaic generation. Starting from geotagged images with respective EXIF-headers or ROS messages containing the images and corresponding metadata, a visual SLAM process using ORB-SLAM2 is triggered. ORB-SLAM creates a sparse

<sup>6</sup> [https://docs.px4.io/v1.9.0/en/flight\\_controller/pixhawk-2.html](https://docs.px4.io/v1.9.0/en/flight_controller/pixhawk-2.html)

<sup>7</sup> <https://ardupilot.org/plane/>

<sup>8</sup> <https://www.frsky-rc.com/product/taranis-x9d-plus-2019/>

<sup>9</sup> <https://mavlink.io/en/>

<sup>10</sup> <https://www.ros.org/>

point cloud using ORB detection and description. In parallel, the platform's attitude with respect to its relative orientation is determined.

By introducing the absolute position of each image from either the EXIF-header or the ROS message, the sparse point cloud as well as the platform's attitude determined by ORB-SLAM is transformed into a global reference frame. Currently, the process does not rely on IMU information for attitude determination and a visual-inertial pose determination is labelled future work.

The sparse point cloud is then either densified by simple interpolation or plane sweeping. Although the plane sweeping approach generally returns more details, the result is noisier than the interpolation approach (see Figure 4). Especially in the context of applications in mountainous or vegetated areas, the reconstruction quality is crucial. Moreover, height information is a valuable feature for semantic labelling tasks. In order to achieve good reconstruction quality in these cases while maintaining (near) real-time processing capability, future endeavours will focus on the integration of more powerful yet robust densification methods (e.g. Knöbelreiter et al. (2019); Tonioni et al. (2018)).



Figure 4. Densification approach: input image (left), plane sweeping (centre), interpolation (right). Image courtesy Kern (2018)

The actual orthorectification is conducted by retrieving the images' RGB information by pixelwise backprojection of the height information into the respective image frame.

Finally, the orthomosaic is composed using a region blending approach. The cells constituting the mosaic are selected by a probabilistic method based on elevation variance, elevation hypothesis and the number of observations.

The process is able to output all intermediary results, such as point clouds, densified tiles, orthorectified images as well as a complete orthomosaic or surface model as a final result.

#### 4.2 Semantic segmentation

As a preliminary model architecture for our initial Semantic Segmentation experiments, Deep Layer Aggregation introduced by (Yu et al., 2017) was integrated into our learning and inference pipeline. The main advantage of this architecture is its elegant integration of scale levels and compact model size, which facilitate the run-time performance required for our application scenarios.

While an increasing number of datasets is available tailored to certain applications, such as surveillance and autonomous driving, only a small fraction of them is applicable to our setting, which requires pixel-wise annotations from non-canonical viewpoints in highly unstructured environments. One widely known example is DOTA (Xia et al., 2017), a large-scale dataset for object detection in aerial images. However, the image data is captured from large distances and no annotations are provided for background classes such as trees or roads, which renders the adaption for a fine-grained scene analysis in Search & Rescue operations difficult.

A more suitable data basis is provided by the Semantic Drone Dataset (TU Graz (ICG), 2019), which consists of densely labelled images captured from a UAV perspective. In order to

specialise the training data to flooding and forest fires, we aggregated the 24 source labels of this dataset to our target classes of vegetation, grass, building, person, vehicle, water and traversable, the latter containing regions such as paved-area, dirt or gravel. This approach allows for a continuous and seamless integration of further datasets, as well as a specialisation to the domains of flooding and forest fires, individually.

### 5. PRELIMINARY RESULTS

In this section, we present early results of the orthomosaic generation as well as the semantic labelling method. The data has been acquired by a Skywalker EVE-2000 in a rural area south of Vienna, Austria. The area is predominantly flat and comprises fields, regular as well as farm roads and a recycling yard with a few buildings and vehicles.

In total, 650 images were captured (without a gimble) at an average altitude of 100 metres with five frames per second and a resolution of 1920 x 1080. The UAV flew 1.5 circles over the area following the farm and paved roads leaving a gap in the centre (see Figure 6).

#### 5.1 Orthomosaic generation

Since the orthomosaic generation is not fully optimised yet and the images were not resampled before processing, the procedure had to be locked at two frames per second. Although most of the steps in the orthomosaic generation process potentially allow for buffering and caching the data and do not require absolute real-time computation, feature tracking during the visual SLAM pipeline is prone to interruption using higher framerates.

The orthomosaic (see Figure 6) has been computed with a ground sampling distance of 5 cm. The overall quality is acceptable given the fact that a (near) real-time situation has been simulated. Some areas could not be properly reconstructed (see red circle in Figure 5). This is possibly due to the higher acquisition altitude during the second circle. The visual SLAM procedure could not close the loop and respective gaps propagated to later steps causing a clear break in the resulting orthomosaic.

Moreover, some blending issues are visible in the orthomosaic (green circle in Figure 6). In general, poorly textured areas with repeated patterns are difficult to reconstruct. Integrating inertial measurements will likely allow for a more robust model and hence reconstruction.

Interestingly, a shadow cast by a wind turbine is properly projected into the orthomosaic (orange circle in Figure 6) ascertaining a good reconstruction result in this region.

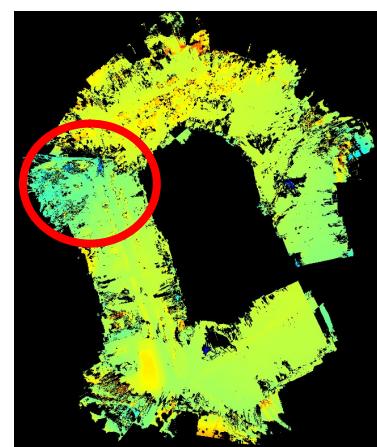


Figure 5. Computed DSM of the test area



Figure 6. Orthomosaic of test area

## 5.2 Semantic segmentation

We conducted our semantic labelling experiments based on the Deep Layer Aggregation model variant dla-34. The 400 publicly available images were split using a ratio of 80:20 for training and validation, respectively.

Standard data augmentation techniques were applied, including random resizing, horizontal/vertical mirroring and cropping to a size of 640 pixels. With a batch size of 24, we achieve a mean-intersection-over-union (mIoU) of 55.4% on the validation set after 60 training epochs on a test bench with three NVIDIA 2080 RTX GPUs. Tests conducted on a single GPU resulted in a frame rate of 11.2 frames per second for inferring the labelled images. Optionally, pixel-level confidences can be generated, which can serve as an essential map-making cue for our SLAM-based approach in combination with the provided distinction between static and dynamic objects.

As visible in Figure 7, the model is able to sufficiently identify traversable regions and distinguish them from classes such as water, tree or building. Although there is a tendency to confuse labels in visually similar image regions, most difficult cases can be resolved correctly, such as water and persons in non-canonical perspectives.

Compared to the validation examples, the test dataset is more challenging regarding flight altitude and the unusual appearance of most classes, as depicted in Figure 8. Besides some minor issues in distinguishing between grass and vegetation, some image regions resemble different classes in the training dataset (e.g. similarities in texture and colour between regions of fields and concrete areas). However, ambiguous regions can be identified using the confidence map, which clearly shows higher confidence values in correctly labelled regions. While the initial results are already promising, we plan to further improve the reliability of our scene analysis module with additional training and test images to overcome the domain gap and increase data variability.

Evidently, further improving the semantic segmentation requires the acquisition and annotation of image data specialised on flooding and forest fires. While a higher variability and more thorough coverage of underrepresented classes is generally important to mitigate data gaps, especially the latter scenario will furthermore benefit from including IR-images as an additional modality. Based on the preliminary results, an ablation study will be conducted to identify and evaluate other model architectures regarding their applicability for on-board real-time processing.

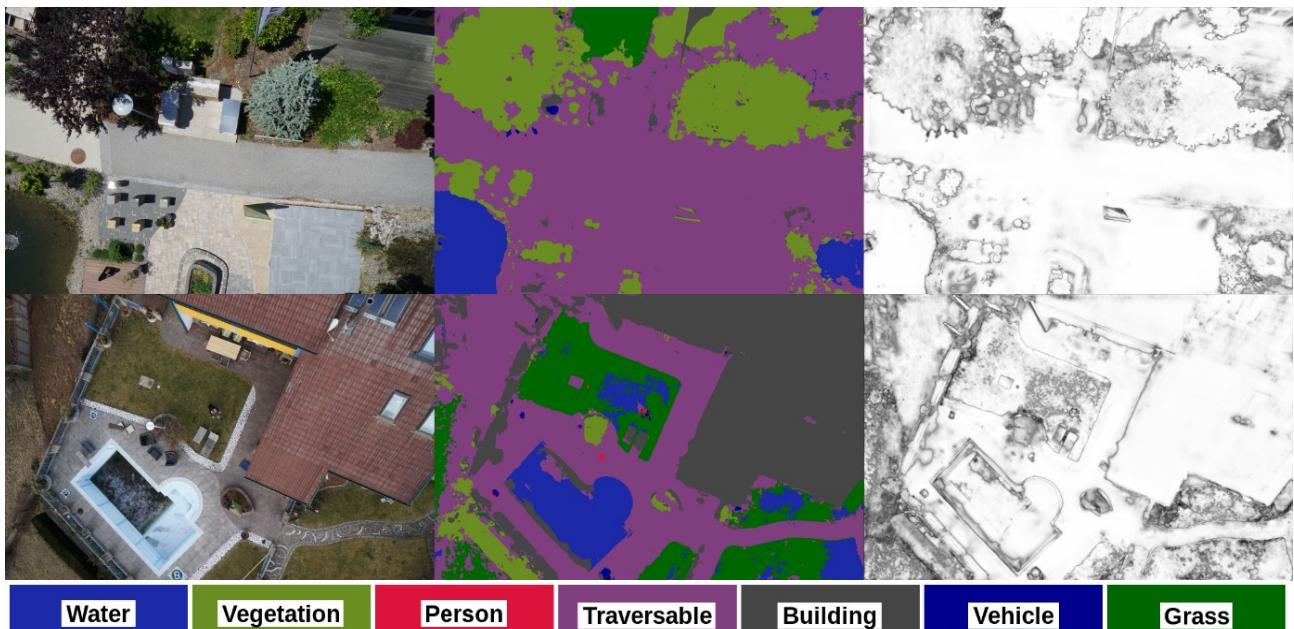


Figure 7. Results on validation images of SemanticDroneDataset (left to right: input image, segmentation results, confidence map)

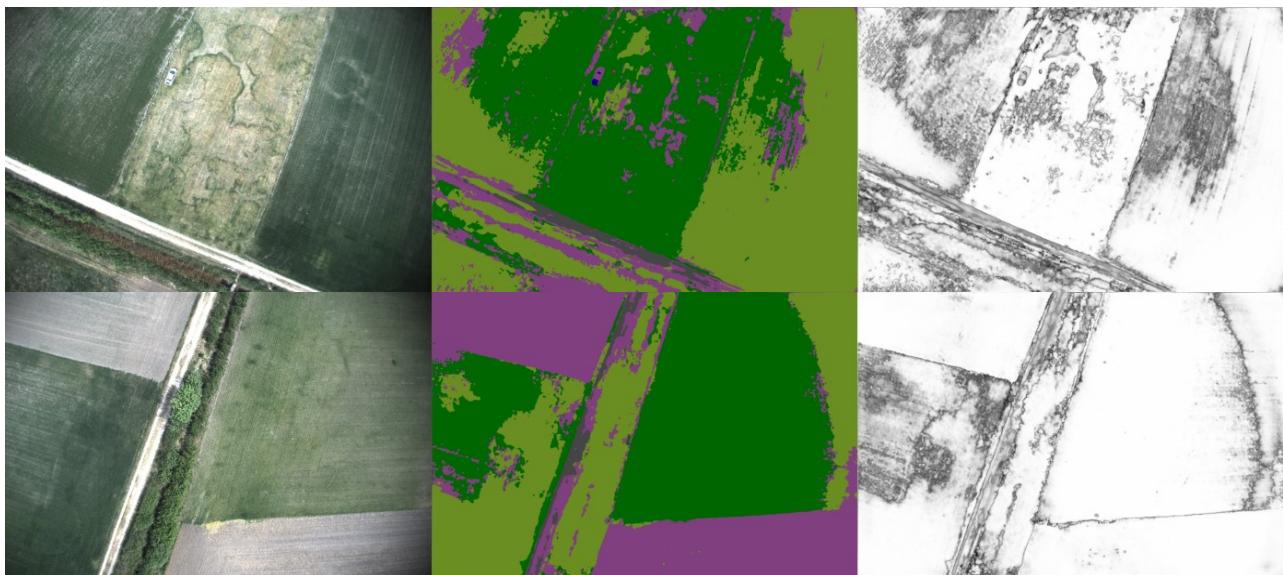


Figure 8. Initial results on test dataset (from left to right: input image, segmentation results, confidence map)

## 6. CONCLUSION AND OUTLOOK

This paper presented the concept, the design and a brief experimental study of a modular UAV-based near real-time mapping system for disaster purposes. We outlined our hardware and software architecture, summarised the legal framework in The European Union and discussed two essential parts of our pipeline, orthomosaic generation and semantic segmentation, in the methodology and experimental section. In the future, we conduct task-related campaigns, i.e. mapping water bodies and mountainous terrain as well as acquire thermal imagery. The orthomosaic generation will be further optimised with respect to speed and reconstruction quality as well as be ported to an embedded system to be executed directly on the UAV hardware. Special emphasis will be placed on finding the right balance between quality and performance with running on an embedded system. The semantic segmentation will highly benefit from training data ideally fit for our special purpose.

## REFERENCES

Alamouri, A., Gerke, M., Batzdorfer, S., Becker, M., Bestmann, U., Bobbe, M., Khedar, Y., Blume, T., Schattenberg, J., Schmiemann, J., 2019. THE JOINT RESEARCH PROJECT ANKOMMEN – EXPLORATION USING AUTOMATED UAV AND UGV. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLII-2/W13, 165–172. <https://doi.org/10.5194/isprs-archives-XLII-2-W13-165-2019>

Bobbe, M., Kern, A., Khedar, Y., Batzdorfer, S., Bestmann, U., 2017. An Automated Rapid Mapping Solution Based on ORB SLAM 2 and Agisoft Photoscan API. Int. Micro Air Veh. Conf. Flight Compet. 2 286–292.

Bu, S., Zhao, Y., Wan, G., Liu, Z., 2016. Map2DFusion: Real-time incremental UAV image mosaicing based on monocular SLAM, in: IEEE International Conference on Intelligent Robots and Systems. Institute of Electrical and Electronics Engineers Inc., pp. 4564–4571. <https://doi.org/10.1109/IROS.2016.7759672>

Chen, L.-C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking Atrous Convolution for Semantic Image Segmentation.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-December, 3213–3223.

Duarte, D., Nex, F., Kerle, N., Vosselman, G., 2017. TOWARDS A MORE EFFICIENT DETECTION OF EARTHQUAKE INDUCED FAÇADE DAMAGES USING OBLIQUE UAV IMAGERY. d-nb.info XLII-2/W6, 93–100. <https://doi.org/10.5194/isprs-archives-XLII-2-W6-93-2017>

EASA, 2019. Drones - regulatory framework timeline | EASA [WWW Document]. URL <https://www.easa.europa.eu/drones-regulatory-framework-timeline> (accessed 4.30.20).

EASA, 2015a. A-NPA 2015-10: Introduction of a regulatory framework for the operation of drones | EASA [WWW Document]. URL <https://www.easa.europa.eu/document-library/notices-of-proposed-amendment/npa-2015-10> (accessed 4.30.20).

EASA, 2015b. Opinion of a technical nature: Introduction of a regulatory framework for the operation of unmanned aircraft | EASA [WWW Document]. URL <https://www.easa.europa.eu/document-library/opinions/opinion-technical-nature> (accessed 4.30.20).

Erdelj, M., Natalizio, E., Chowdhury, K.R., Akyildiz, I.F., 2017. Help from the Sky: Leveraging UAVs for Disaster Management. IEEE Pervasive Comput. <https://doi.org/10.1109/MPRV.2017.11>

Fan, R., Jiao, J., Pan, J., Huang, H., Shen, S., Liu, M., 2019. Real-Time Dense Stereo Embedded in A UAV for Road Inspection.

Ghosh, D., Kaabouch, N., 2016. A survey on image mosaicing techniques. J. Vis. Commun. Image Represent. 34, 1–11. <https://doi.org/10.1016/j.jvcir.2015.10.014>

Hein, D., Kraft, T., Brauchle, J., Berger, R., 2019. Integrated UAV-Based Real-Time Mapping for Security Applications. ISPRS Int. J. Geo-Information 8, 219. <https://doi.org/10.3390/ijgi8050219>

Hinzmann, T., Schönberger, J.L., Pollefeys, M., Siegwart, R., 2018. Mapping on the Fly: Real-Time 3D Dense Reconstruction, Digital Surface Map and Incremental Orthomosaic Generation for Unmanned Aerial Vehicles. Springer, Cham, pp. 383–396.

https://doi.org/10.1007/978-3-319-67361-5\_25

JARUS, 2019. Specific Risk Operation Assessment [WWW Document]. URL <http://jarus-rpas.org/content/jar-doc-06-sora-package> (accessed 4.30.20).

Kerle, N., Nex, F., Duarte, D., Vetrivel, A., 2019. Uav-based structural damage mapping-results from 6 years of research in two european projects, in: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. International Society for Photogrammetry and Remote Sensing, pp. 187–194. <https://doi.org/10.5194/isprs-archives-XLII-3-W8-187-2019>

Kerle, N., Gerke, Duarte, Vetrivel, 2019. UAV-Based Structural Damage Mapping: A Review. *ISPRS Int. J. Geo-Information* 9, 14. <https://doi.org/10.3390/ijgi9010014>

Kern, A., 2018. Real-time Photogrammetry using monocular SLAM for Unmanned Aerial Vehicles.

Knöbelreiter, P., Vogel, C., Pock, T., 2019. Self-Supervised Learning for Stereo Reconstruction on Aerial Images.

Lindner, G., Schraml, K., Mansberger, R., Hübl, J., 2016. UAV monitoring and documentation of a large landslide. *Appl. Geomatics* 8, 1–11. <https://doi.org/10.1007/s12518-015-0165-0>

Long, J., Shelhamer, E., Darrell, T., 2014. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 640–651.

Luo, C., Nightingale, J., Asemota, E., Grecos, C., 2015. A UAV-cloud system for disaster sensing applications, in: IEEE Vehicular Technology Conference. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/VTCSpring.2015.7145656>

Miyano, K., Shinkuma, R., Mandayam, N.B., Sato, T., Oki, E., 2019. Utility Based Scheduling for Multi-UAV Search Systems in Disaster-Hit Areas. *IEEE Access* 7, 26810–26820. <https://doi.org/10.1109/ACCESS.2019.2900865>

Mur-Artal, R., Tardos, J.D., 2016. ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. *IEEE Trans. Robot.* 33, 1255–1262. <https://doi.org/10.1109/TRO.2017.2705103>

Rossi, G., Tanteri, L., Tofani, V., Vannocci, P., Moretti, S., Casagli, N., 2018. Multitemporal UAV surveys for landslide mapping and characterization. *Landslides* 15, 1045–1052. <https://doi.org/10.1007/s10346-018-0978-0>

Saeed Yahyanejad, 2013. Orthorectified Mosaicking of Images from Small-scale Unmanned Aerial Vehicles. undefined.

Schönberger, J.L., Frahm, J.-M., 2016. Structure-from-Motion Revisited.

Silvagni, M., Tonoli, A., Zenerino, E., Chiaberge, M., 2017. Multipurpose UAV for search and rescue operations in mountain avalanche events. *Geomatics, Nat. Hazards Risk* 8, 18–33. <https://doi.org/10.1080/19475705.2016.1238852>

Suzuki, T., Miyoshi, D., Meguro, J.I., Amano, Y., Hashizume, T., Sato, K., Takiguchi, J.I., 2008. Real-time hazard map generation using small unmanned aerial vehicle, in: Proceedings of the SICE Annual Conference. pp. 443–446. <https://doi.org/10.1109/SICE.2008.4654695>

Tanteri, L., Rossi, G., Tofani, V., Vannocci, P., Moretti, S., Casagli, N., 2017. Multitemporal UAV Survey for Mass Movement Detection and Monitoring, in: Advancing Culture of Living with Landslides. Springer International Publishing, pp. 153–161. [https://doi.org/10.1007/978-3-319-53498-5\\_18](https://doi.org/10.1007/978-3-319-53498-5_18)

The European Commission, 2019. Commission Implementing Regulation (EU) 2019/947 of 24 May 2019 on the rules and procedures for the operation of unmanned aircraft [WWW Document]. URL <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1588063082075&uri=CELEX:32019R0947> (accessed 4.30.20).

Tonioni, A., Tosi, F., Poggi, M., Mattoccia, S., Di Stefano, L., 2018. Real-time self-adaptive deep stereo.

TU Graz (ICG), 2019. Semantic Drone Dataset v1.1 [WWW Document]. URL <http://dronedataset.icg.tugraz.at/> (accessed 5.4.20).

Waharte, S., Trigoni, N., 2010. Supporting search and rescue operations with UAVs, in: Proceedings - EST 2010 - 2010 International Conference on Emerging Security Technologies, ROBOSEC 2010 - Robots and Security, LAB-RS 2010 - Learning and Adaptive Behavior in Robotic Systems. pp. 142–147. <https://doi.org/10.1109/EST.2010.31>

Xia, G.-S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J., Datcu, M., Pelillo, M., Zhang, L., 2017. DOTA: A Large-scale Dataset for Object Detection in Aerial Images. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 3974–3983.

Yu, F., Wang, D., Shelhamer, E., Darrell, T., 2017. Deep Layer Aggregation. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 2403–2412.

Zhang, J., Xiong, J., Zhang, G., Gu, F., He, Y., 2016. Flooding disaster oriented USV & UAV system development & demonstration, in: OCEANS 2016 - Shanghai. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/OCEANSAP.2016.7485676>