

Article

Not peer-reviewed version

Unsupervised Color Based Flood Segmentation in UAV Imagery

[Georgios Simantiris](#) and [Costas Panagiotakis](#) *

Posted Date: 16 April 2024

doi: [10.20944/preprints202404.1049.v1](https://doi.org/10.20944/preprints202404.1049.v1)

Keywords: Flood detection; image segmentation; remote sensing; unmanned aerial vehicle (UAV); unsupervised segmentation



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Unsupervised Color Based Flood Segmentation in UAV imagery

Georgios Simantiris [†] and Costas Panagiotakis ^{*,†}

Department of Management Science and Technology, Hellenic Mediterranean University, P.O. Box 128, 72100 Agios Nikolaos, Crete, Greece; ddk230@edu.hmu.gr

* Correspondence: cpanag@hmu.gr

† These authors contributed equally to this work.

Abstract: We propose a novel unsupervised semantic segmentation method for fast and accurate flood area detection utilizing color images acquired from Unmanned Aerial Vehicles (UAVs). To the best of our knowledge, this is the first fully unsupervised method for flood area segmentation in color images captured by UAVs, without the need of pre-disaster images. The proposed framework addresses the problem of flood segmentation based on parameter-free calculated masks and unsupervised image analysis techniques. First, a fully unsupervised algorithm gradually excludes areas classified as non-flood utilizing calculated masks over each component of the LAB colorspace, as well an RGB vegetation index and the detected edges of the original image. Unsupervised image analysis techniques, such as distance transform, are then applied, producing a probability map for the location of flooded areas. Finally, flood detection is obtained by applying the hysteresis thresholding segmentation. The proposed method is tested and compared with variations, and other supervised methods in two public datasets, consisting of 953 color images in total, yielding high-performance results, with 87.4% and 80.9% overall accuracy and F1-Score, respectively. The results and computational efficiency of the proposed method show that it is suitable for on board data execution and decision-making during UAVs flight.

Keywords: flood detection; image segmentation; remote sensing; unmanned aerial vehicle (UAV); unsupervised segmentation

1. Introduction

Natural disasters had always profound and far-reaching impacts on humanity. In recent years, we observe a climate change leading to escalating weather phenomena, which in turn facilitate natural disasters. Floods occur in the middle of the summer season due to sudden enormous amounts of rain, and dry weather conditions in combination with strong, out-of-season winds open the door for catastrophic, non-controllable wildfires. Earthquakes, volcano eruptions, and hurricanes appear with immense magnitude. All these disasters lead to loss of life and properties, disruption of services such as water supply, electricity, and transportation with further health risks, and have an enormous economic as well as psychological impact on the population.

Efforts to mitigate the impact of natural disasters include early warning systems, improved infrastructure resilience, disaster preparedness education, and international cooperation for humanitarian assistance. Preparedness and response strategies are crucial to minimizing the human toll and facilitating a quicker recovery from such events. Natural disaster detection systems contribute to early warnings, risk reduction, efficient resource allocation, and community preparedness. By leveraging technology and global cooperation, these systems play a vital role in minimizing the impact of disasters on both human populations and the environment.

Technological advancements and collaborative technologies contribute to the sharing of disaster information benefitting from different types of media. Deep Learning (DL) algorithms show promise in extracting knowledge from diverse data modalities, but their application in disaster response tasks remains largely academic. Systematic reviews evaluated the successes, challenges and future opportunities of using DL for disaster response and management, while also examining Machine Learning (ML) approaches, offering guidance for future research to maximize benefits in disaster



response efforts [1,2]. In this work, we are particularly focused on floods. A concise summary of the research conducted is shown in Table 1.

DL methods are increasingly applied to remote sensing imagery to address the limitations of traditional flood mapping techniques. Convolutional layer-based models offer improved accuracy in capturing spatial characteristics of flooding events, while fully connected layer-based models show promise when coupled with statistical approaches. Remote sensing analysis, multi-criteria decision analysis, and numerical methods are replaced with DL models for flood mapping in which flood extent or inundation maps, susceptibility maps, and flood hazard maps determine, categorize, and characterize the disaster, respectively [3]. Furthermore, in a recent review, current DL approaches for flood forecasting and management are critically assessed, pointing out their advantages and disadvantages. Challenges with data availability and potential future research directions are examined. The current state of DL applications in this matter is comprehensively evaluated, showing that they are a powerful tool to improve flood prediction and control [4].

Table 1. A brief overview of the research for this article depicting approach (supervised, unsupervised), modality, and method.

Authors	Year	Approach	Imagery	Method
Chouhan, A. et al. [5]	2023	Supervised	Sentinel-1	Multi-scale ADNet
Drakonakis, G.I. et al. [6]	2022	Supervised	Sentinel-1, 2	CNN change detection
Dong, Z. et al. [7]	2023	Supervised	Sentinel-1	STANets, SNUNet, BiT
Hänsch, R. et al. [8,9]	2022	Supervised	HR satellite RGB	U-Net
He, Y. et al. [10]	2024	Weakly-supervised	HR aerial RGB	End-to-end WSSS framework structure constraints and self-distillation
Hernández, D. et al. [11]	2021	Supervised	UAV RGB	Optimized DNN
Hertel, V. et al. [12]	2023	Supervised	SAR	BCNN
Ibrahim, N. et al. [13]	2021	Semi-supervised	UAV RGB	RGB and HSI color models, k-means clustering, region growing
Inthizami, N.S. et al. [14]	2022	Supervised	UAV video	Improved ENet
Li, Z. et al. [15]	2023	Supervised	Sentinel-1	U-Net
Lo, S.W. et al. [16]	2015	Semi-supervised	RGB (Surveillance camera)	HSV color model, seeded region growing
Munawar, H.S. et al. [17]	2021	Supervised	UAV RGB	Landmark-based feature selection, CNN hybrid
Park, J.C. et al. [18]	2023	Supervised	HR satellite RGB	Swin transformer in a Siamese-UNet
Rahnemoonfar, M. et al. [19]	2021	Supervised	UAV RGB	InceptionNetv3, ResNet50, XceptionNet, PSPNet, ENet, DeepLabv3+
Şener, A. et al. [20]	2024	Supervised	UAV RGB	ED network with EDR block and atrous convolutions (FASegNet)
Shastry, A. et al. [21]	2023	Supervised	WorldView 2, 3 multispectral	CNN with atrous convolutions
Wang, L. et al. [22]	2022	Supervised	True Orthophoto (near infrared), DSM	Swin transformer and DCFAM
Wieland, M. et al. [23]	2023	Supervised	Satelite and aerial	U-Net model with MobileNet-V3 backbone pre-trained on ImageNet
Bauer-Marschallinger, B. et al. [24]	2022	Unsupervised	SAR	Datacube, time series-based detection, Bayes classifier
Filonenko, A. et al. [25]	2015	Unsupervised	RGB (surveillance camera)	Change detection, color probability calculation
Landuyt, L. et al. [26]	2020	Unsupervised	Sentinel-1, 2	K-means clustering, region growing
Li, J. et al. [27]	2022	Unsupervised	Sentinel-2 Landsat	NDWI, U-Net
McCormack, T. et al. [28]	2022	Unsupervised	Sentinel-1	Histogram thresholding, multi-temporal and contextual filters
Trombini, M. et al. [29]	2023	Unsupervised	SAR	Graph-based MRF segmentation

Convolutional Neural Networks (CNNs) proved to be effective in flood detection using satellite imagery. High-quality flood maps are generated with the help of temporal differences from various sensors after CNNs identify changes between permanent and flooded water areas utilizing Synthetic Aperture Radar (SAR) and multispectral images [6,7]. Furthermore, the efficacy of CNNs in semantically segmenting water bodies in highly detailed satellite and aerial images from various sensors, with a focus on flood emergency response applications, is assessed by combining different CNN architectures with encoder backbones to delineate inundated areas under diverse environmental conditions and data availability scenarios. A U-Net model with a MobileNet-V3 backbone pre-trained

on ImageNet consistently performed the best in all scenarios tested, while the integration of additional spectral bands, slope information from digital elevation models, augmentation techniques during training, and the inclusion of noisy data from online sources further improved model performance [23]. In addition, Bayesian convolutional neural networks (BCNN) have been recommended to quantify the uncertainties associated with SAR-based water segmentation, because of their greater flexibility to learn the mean and the spread of the parameter posterior [12]. Finally, a CNN employed to automatically detect inundation extents using the Deep Earth Learning, Tools, and Analysis (DELTa) framework demonstrated high precision and recall for water segmentation despite a diverse training dataset. Supplementary, the effects of surface obstruction due to the inability of optical remote sensing data to observe floods under clouds or flooded vegetation are quantified, suggesting the integration of flood models to improve segmentation accuracy [21].

Since rapid damage analysis and fast coordination of humanitarian response during extreme weather events are crucial, flood detection, building footprint detection, and road network extraction have been integrated into an inaugural remote sensing dataset called SpaceNet 8 and a homonym challenge has been launched [8]. The provided satellite imagery posed real-world challenges such as varying resolutions, misalignment, cloud cover, and lighting conditions and top performing DL approaches focusing on multi-class segmentation showed that swiftly identifying flooded infrastructures such as buildings and roads can significantly shorten response times. The contestants found that simple U-Net architectures yielded the best balance of accuracy, robustness, and efficiency, with strategies such as pre-training and data augmentation proving crucial to improve model performance [9].

U-Nets and their variations were widely used to tackle the problems of water bodies segmentation and flood extent extraction. In [27], a ResNet was used to replace the contracting path of the U-Net, due to its ability to solve vanishing or exploding gradient issues caused by error backpropagation through skip connections of the residual module. The normalized difference water index (NDWI) was employed to create pseudo-labels for training, resulting in an unsupervised DL approach. In [15], another adjusted U-Net was proposed. With carefully selected parameters and training with pre-processed for three-category classification Sentinel-1 images, the proposed method was able to distinguish flood pixels from permanent water and background.

Transformers have also been successfully applied for semantic segmentation in remote sensing images. A novel transformer based scheme employing the Swin Transformer as the backbone to better capture context information and a densely connected feature aggregation module (DCFAM) serving as a novel decoder to restore resolution and generate accurate segmentation maps, proved to be effective in the ISPRS Vaihingen and Potsdam datasets [22]. An improved transformer-based multiclass flood detection model capable of predicting flood events while distinguishing between roads and buildings was introduced, which, with an additional novel loss function and a road noise removal algorithm, achieved superior performance, particularly in road evaluation metrics such as APLS [18]. Finally, the Bitemporal image Transformer (BiT) model scored highest in a change detection approach capturing the changed region better [7].

A multiscale attentive decoder-based network (ADNet) designed for automatic flood identification using Sentinel-1 images outperforms recent DL and threshold-based methods when validated on the Sen1floods11 benchmark dataset. Through detailed experimentation on various dataset settings, ADNet demonstrates effective delineation of permanent water, flood water and all water pixels using both co-polarization (VV) and cross-polarization (VH) inputs from Sentinel-1 images [5].

Dilated or atrous convolutions, which on the one hand increase the network's receptive field while on the other reduce the number of trained parameters needed [30], are utilized in an effort to speed up search and rescue operations after natural disasters such as floods, high tides, and tsunamis. FASegNet, a novel CNN-based semantic segmentation model, was introduced specifically designed for flood and tsunami area detection. FASegNet utilizes encoder and decoder networks with an encoder-decoder-residual (EDR) block to effectively extract local and contextual information. An Encoder-Decoder High-Accuracy Activation Cropping (EHAAC) module minimizes information loss

at the bottleneck, and skip connections transfer information between encoder and decoder networks, outperforming other segmentation models [20].

A novel weak training data generation strategy and an end-to-end weakly supervised semantic segmentation (WSSS) method called TFCSD, challenges urban flood mapping [10]. By decoupling the acquisition of positive and negative samples, the weak label generation strategy significantly reduces the burden of data labeling, enabling prompt flood mapping in emergencies. Additionally, the proposed TFCSD method improves edge delineation accuracy and algorithm stability compared to other methods, especially in emergency scenarios where pre-disaster river data is accessible, or when using the SAM ([31]) assisted interactive labeling method when such data is unavailable.

Satellites such as Sentinel-1 and Sentinel-2 play a key role in flood mapping due to their rapid data acquisition capabilities. Their effectiveness in mapping floods across Europe was evaluated in a study in which the results indicate that observation capabilities vary based on catchment area size and suggest that employing multiple satellite constellations significantly increases flood mapping coverage [32]. The urgent need for real-time flood management systems by developing an automated imaging system using Unmanned Aerial Vehicles (UAVs) to detect inundated areas promptly is addressed, so that emergency relief efforts will not be hindered by current satellite-based imaging systems which suffer from low accuracy and delayed response. By employing the Haar cascade classifier and DL algorithms, a hybrid flood detection model combining landmark-based feature selection with a CNN demonstrated improved performance over traditional classifiers [17].

Specially designed datasets are introduced to address the lack of high-resolution (HR) imagery relevant to disaster scenarios. In [19], FloodNet, a high resolution (HR) UAV imagery dataset, capturing post-flood damage, aims to detect flooded roads and buildings and distinguish between natural and flooded water. Baseline methods for image classification, semantic segmentation, and visual question are evaluated, highlighting its significance for analyzing disaster impacts with various DL algorithms, such as XceptionNet and ENet. In [14], an improved Efficient Neural Network architecture was also the choice to segment the UAV video of flood disaster. The proposed method consists of atrous separable convolution as the encoder and depth-wise separable convolution as the decoder.

To facilitate efficient processing of disaster images captured by UAVs, an AI-based pipeline was proposed enabling semantic segmentation with optimized deep neural networks (DNNs) for real-time flood area detection directly on UAVs, minimizing infrastructure dependency and resource consumption of the network. Experimental results confirm the feasibility of performing sophisticated real-time image processing on UAVs using GPU-based edge computing platforms [11].

It becomes clear that DL methods offer improved segmentation by creating adaptive mapping relationships based on contextual semantic information. However, these methods require extensive manual labeling of large datasets and lack interpretability, suggesting the need to address these limitations for further progress. Traditional ML methods, on the other hand, rely on manually designed mappings. Systematic reviews of water body segmentation over the past 30 years examine the application and optimization of DL methods and outline traditional methods at both the pixel and the image levels [33]. Evaluating the strengths and weaknesses of both approaches prompts a discussion of the importance of maintaining knowledge of classical computer vision techniques. There remains value in understanding and utilizing these older techniques. The knowledge gained from traditional Computer Vision (CV) methods can complement DL, expanding the available solutions. There also exist scenarios in which traditional CV techniques can outperform DL or be integrated into hybrid approaches for improved performance. Furthermore, traditional CV techniques have been shown to have benefits such as reducing training time, processing, and data requirements compared to DL applications [34].

A near decade ago, a method for automatically monitoring flood events in specific areas was proposed using remote cyber-surveillance systems and image-processing techniques. When floods are treated as possible intrusion objects, the intrusion detection mode is utilized to detect and verify flood objects, enabling automatic and unattended flood risk level monitoring and urban inundation

detection. Compared to large-area forecasting methods, this approach offered practical benefits, such as flexibility in location selection, no requirement for real-world scale conversion, and a wider field of view, facilitating more accurate and effective disaster warning actions in small areas [16]. Real-time methods to detect flash floods using stationary surveillance cameras, suitable for both rural and urban environments, became quite popular. Another method used background subtraction to detect changes in the scene, followed by morphological closing to unite pixels belonging to the same objects. Additionally, small separate objects are removed, and the color probability is calculated for the foreground pixels, filtering out components with low probability values. The results are refined using edge density and boundary roughness [25].

Unsupervised object-based clustering was also used for flood mapping in SAR images. The framework segments the region of interest into objects, converts them into a SAR optical feature space, and clusters them using K-means, with the resulting clusters classified based on centroids and refined by region growing. The results showed improved performance compared to pixel and object-based benchmarks, with additional SAR and optical features enhancing accuracy and post-processing refinement reducing sensitivity to parameter choice even in difficult cases, including areas with flooded vegetation [26]. The same techniques were also proposed for flood detection purposes in UAV-captured images. Employing RGB and HSI color models and two segmentation methods: K-means clustering and region growing in a semi-supervised scheme showed potential for accurate flood detection [13].

There is also a datacube-based flood mapping algorithm that uses Sentinel-1 data repetition and predefined probability parameters for flood and non-flood conditions [24]. The algorithm autonomously classifies flood areas and estimates uncertainty values, demonstrating robustness and near-real-time operational suitability. It also contributed to the Global Flood Monitoring component of the Copernicus Emergency Management Service.

Contextual filtering on multi-temporal SAR imagery resulted in an automated method for mapping non-urban flood extents [28]. Using tile-based histogram thresholding and refined with post-processing filters, including multitemporal and contextual filters, the method achieved high accuracy. Additionally, confidence information was provided for each flood polygon, enabling stable and systematic inter-annual flood extent comparisons at gauged and ungauged sites.

Finally, in [29] an unsupervised graph-based image segmentation method has been proposed that aims to achieve user-defined and application-specific segmentation goals. This method utilizes a graph structure over the input image and employs a propagation algorithm to assign costs to pixels based on similarity and connectivity to reference seeds. Subsequently, a statistical model is estimated for each region, and the segmentation problem is formulated within a Bayesian framework using probabilistic Markov random field (MRF) modeling. Final segmentation is achieved through minimizing an energy function using graph cuts and the alpha-beta swap algorithm, resulting in segmentation based on the maximum a posteriori decision rule. In particular, the method does not rely on extensive prior knowledge and demonstrates robustness and versatility in experimental validation with different modalities, indicating its potential applicability across different domains. It was also successfully applied on SAR images for flood mapping.

From our survey, it becomes clear that supervised methodologies are preferred nowadays, as they outnumber unsupervised approaches (see Table 1). Of the unsupervised ones, we found only one that deals with RGB images, but relies on change detection, thus is in need of the pre-disaster image as well.

In this paper, we propose a novel unsupervised method for flood segmentation utilizing color images acquired from UAVs. Without the need of large datasets, extensive labeling, augmentation, and training, the segmentation can be performed directly on the UAV deployed over the disaster area. Therefore, relief efforts can be swiftly directed to damaged sites avoiding time loss, which can be crucial in saving lives and properties. Initially, we employ parameter-free calculated masks over each component of the LAB colorspace utilizing as well an RGB vegetation index and the detected edges of the original image in order to provide an initial segmentation. Next, unsupervised image

analysis techniques, such as distance transform, are adapted to the flood detection problem, producing a probability map for the location of flooded areas. Then, the hysteresis thresholding segmentation method is applied, resulting in the final segmentation. The main contributions of our work can be summarized as follows:

- To our knowledge, this is the first fully unsupervised method for flood area segmentation in color images captured by UAVs. The current work faces for the first time the problem of flood segmentation based on parameter-free calculated masks and unsupervised image analysis techniques.
- The flood areas are given as solutions of a probability optimization problem based on the evolution of an isocontour starting from the high confidence areas and gradually growing according to the hysteresis thresholding method.
- The proposed formulation yields a robust unsupervised algorithm that is simple and effective for the flood segmentation problem.
- The proposed framework is suitable for on-board execution on UAVs, enabling real-time processing of data and decision making during flight, since the processing time per image is about 0.5 sec without the need of substantial computational resources or specialized GPU capabilities.

The proposed system has been tested and compared with other variants and supervised methods on the Flood Area dataset introduced in [35], consisting of 290 color images, yielding high-performance results. Furthermore, experimental results of the proposed method are also reported on the Flood Semantic Segmentation dataset [36], which consists of 663 color images.

The rest of this paper is organized as follows. Sections 2 introduces the datasets used for this article. Section 3 presents our proposed unsupervised methodology. The experimental results and a comprehensive discussion are given in Section 4. Finally, conclusions and future work are provided in Section 5.

2. Materials

We employed two publicly available datasets for this study to demonstrate the robustness and general applicability of our method. First, the dataset used for assessing the efficacy of our approach and facilitating comparative analyses with alternative methodologies is called Flood Area, consisting of color images acquired by UAVs and helicopters [35]. It contains 290 RGB images depicting flood hit areas, as well as their corresponding mask images with the water region segmentations. The ground truth images were annotated by the dataset creators using Label Studio, an open source data labeling software. The images were downloaded selectively from the Internet, thus the dataset exhibits a wide range of image variability, depicting urban, peri-urban, rural areas, greenery, rivers, buildings, roads, mountains, and the sky. Furthermore, there are image acquisitions relatively close to the ground as well as from a very high altitude and from diverse camera rotation angles around the X axis (roll) and Y axis (pitch). If pitch and roll are zeros, this means that the camera is looking down (top-down view). Hereafter, we use the term “camera rotation angle”, to present the angle between the current view plane and the horizontal plane (top-down view). The images have different resolutions and dimensions with height and width ranging from 219 up to 3648 and 330 up to 5472, respectively. Representative images with their corresponding ground truths are shown in Figure 1.

Second, to confirm the universal functionality of our approach, we employed the Flood Semantic Segmentation Dataset [36]. It consists of 600 and 63 color images for training and validation, respectively. Since our method is fully unsupervised, it does not require training, and therefore we used all 663 images for evaluation. Similarly to the initial dataset, this dataset comprises images obtained from UAVs, accompanied by their respective ground truth annotations, portraying diverse flooded scenes captured from various camera perspectives. The image sizes and resolutions also vary, but were all resized and, if necessary zero padded, to 512x512 by the creator, as shown in Figure 2.



Figure 1. Sample images from the Flood Area dataset (top) and their corresponding ground truths (bottom).



Figure 2. Sample images from the Flood Semantic Segmentation dataset (top) and their corresponding ground truths (bottom).

3. Methodology

3.1. System Overview

We propose an approach which gradually removes image areas classified as non-flood, based on binary masks constructed from color and edge information. Our method is fully unsupervised, meaning that there is no training process involved, and thus no need of ground truth labeling. We use the labels provided by the datasets only for evaluation purposes of our method. A repetitive process, consisting of the same algorithmic steps, is applied over each of the components extracted from the color image, in order to identify areas that are not affected by floods. For each component, as described below, a binary map is obtained in which areas identified as non-flood are discarded, leading to a final mask of potential flood areas (PFAs), refined by simple morphological operations. The flood's dominant color is calculated weighting the potential flood area pixels and a hysteresis thresholding yields the final segmentation. An overview of our proposed methodology is graphically depicted in Figure 3. In the following, we analytically present the proposed methodology.

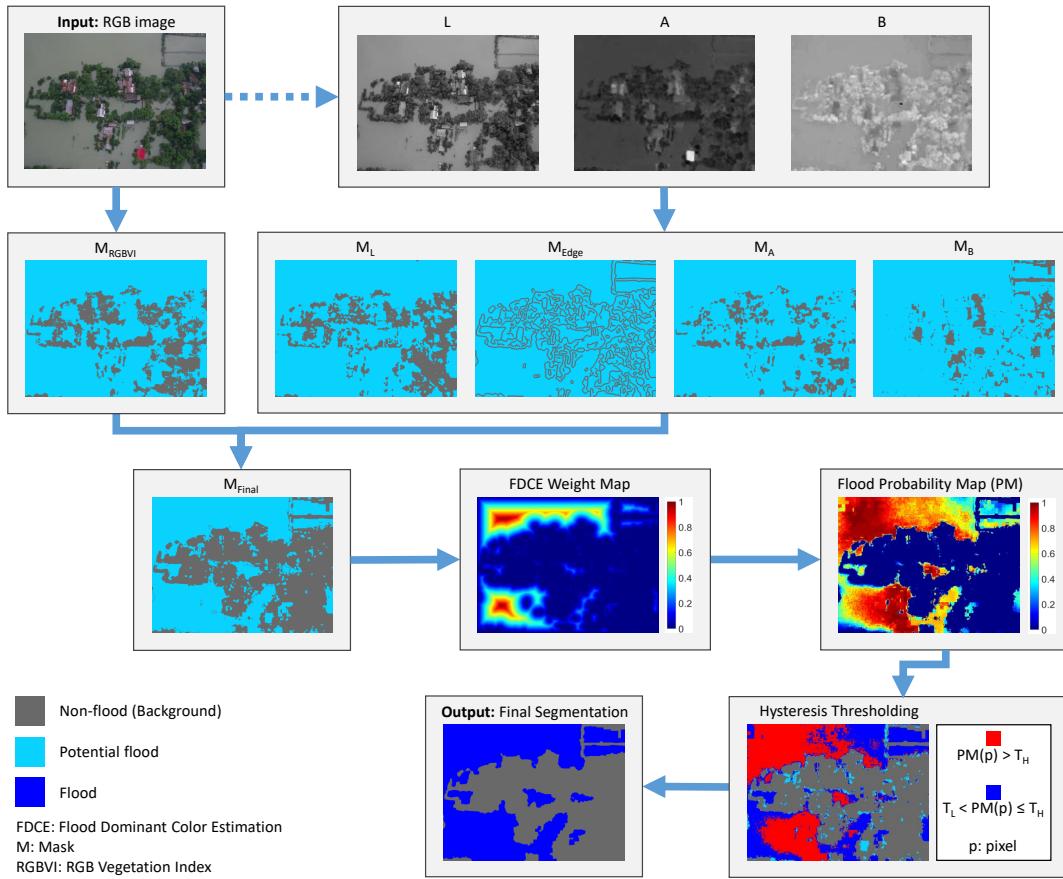


Figure 3. Graphical abstract of our proposed approach.

3.2. RGB Vegetation Index Mask

In both urban and rural areas, the landscape is lush with greenery, largely attributed to the abundant presence of trees and vegetation. Since trees are unlikely to be fully covered by flood events, our first concern is to rule out the greenery, noticing also in our experiments that, using only a flood color approach, vegetation is more likely to be misclassified as flood water. Therefore, we use the RGB Vegetation Index (RGBVI), introduced in [37]. This index was successfully applied in [38] as a first step in detecting and counting trees using region-based circle fitting. RGBVI particularly improves sensitivity to vegetation characteristics while mitigating the impact of interfering factors such as reflection of the background soil and directional effects. However, as shown in Equation 1, it can be influenced by the color quality of the image, e.g. due to bad atmospheric conditions by the time of the image acquisition. It is defined as the normalized difference between the squared green reflectance and the product of blue and red reflectance:

$$RGBVI = \frac{(R_G)^2 - (R_B \times R_R)}{(R_G)^2 + (R_B \times R_R)} \quad (1)$$

where R_G , R_B and R_R denote the red, blue and green reflectance, respectively. In [39], the authors use a threshold of 0.15 on RGBVI for tree detection. In this work, we preferred to set a stricter threshold, e.g. 0.2, so that any value exceeding this threshold is characterized as greenery and therefore is not flooded, resulting in M_{RGBVI} mask. With this mask, we are able to rule out a great amount of image pixels, since visible vegetation cannot be flooded. Particularly in the Flood Area dataset in rural areas up to 92.8% of pixels can be characterized as greenery, and thus non-flood, while on average over the whole dataset 17.94% of pixels are ruled out this way (median value is 14.16%).

In Figure 4, examples of RGBVI masks are shown for (a) urban and (b) rural areas from the Flood Area dataset, where the dim gray color corresponds to the detected greenery. We can clearly notice that, especially in rural areas, a substantial number of image pixels are rightly characterized as trees and vegetation, and therefore these areas cannot be flooded. This technique also works in urban areas where vegetation is present. But there are also cases where the RGB Vegetation Index is not quite efficient, especially when image color quality is poor due to camera rotation angle and/or weather conditions (Figure 4 (c)).

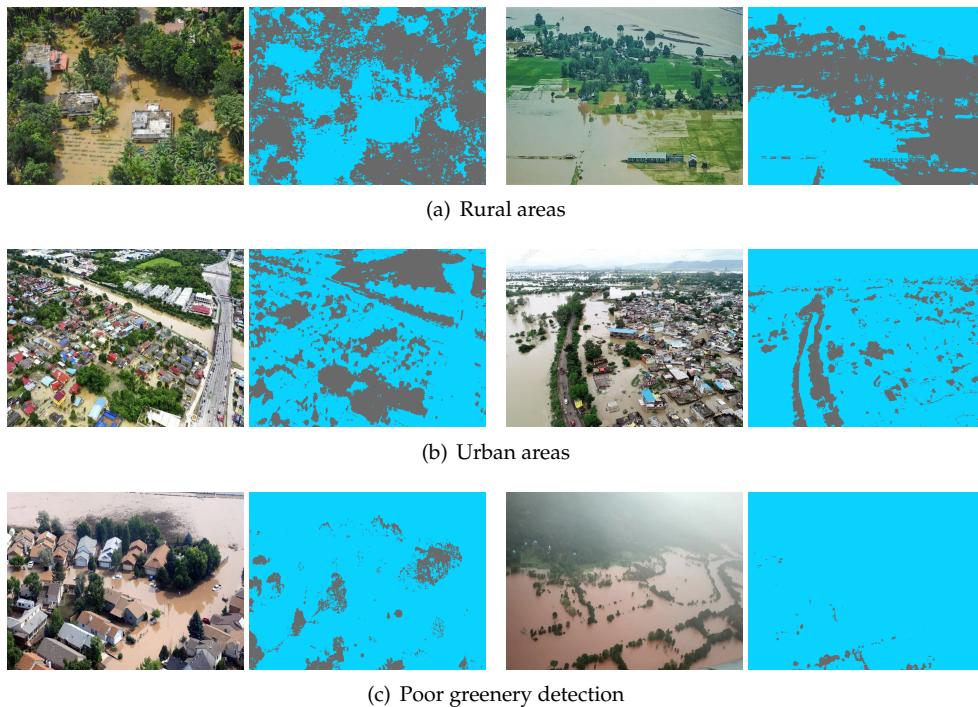


Figure 4. The original images (from the Flood Area dataset) and the corresponding RGBVI masks on their right side. The masks show detected greenery with dim gray color. Examples are presented for (a) urban areas, (b) rural areas, and (c) poor or failed greenery detection. The remaining potential flood areas are shown in cyan.

3.3. LAB Components Masks

The LAB color space offers several advantages over the RGB color space, such as perceptual uniformity, wide color gamut, separation of color and lightness, and robustness to illumination [40]. LAB color space is designed to be perceptually uniform, which means that a small change in the LAB values corresponds to a similar perceptual change in color across the entire color space, making it more suitable for color-based applications where accurate perception of color differences is important. It encompasses a wider range of colors compared to the RGB color space, particularly in terms of the human perceptual color space. This allows for a more accurate representation of colors that fall outside the RGB gamut. The LAB color space is less affected by changes in illumination compared to the RGB color space, since lightness is a separate component. This separation is advantageous when independent control over lightness and color is desired, making it more suitable for applications where lighting conditions vary significantly [41]. Considering that the LAB color space offers greater flexibility and accuracy in color representation and manipulation compared to the RGB color space, and since in our application precise color information is critical, we next convert the image to the LAB color space to process it further.

Using the L, A, and B components, we derive three more masks, where areas can potentially be characterized as non-flood, exploiting the values standard deviation of the data relative to its central tendency for each color component. In flooded areas, the value of each color component in the LAB

color space is usually higher than the corresponding values of the background. So, we select a lower threshold $\Delta_C, C \in \{L, A, B\}$ to create binary masks that detect regions probably belonging to non-flood areas. Δ_C is calculated by subtracting the standard deviation (σ_C) from the mean value (μ_C) over each color component (C) of the LAB color space:

$$\Delta_C = \mu_C - \sigma_C, \quad C \in \{L, A, B\} \quad (2)$$

Figure 5 shows the average value of (a) L, (b) A and (c) B color components computed on flood (blue curve) and background (red curve) pixels for each image of the Flood Area dataset, sorted in ascending order. The yellow curves represent the corresponding threshold Δ_C used to create the three masks. Since flooded areas have elevated values observed in each component of the LAB color space, each mask M_L, M_A and M_B is labeled as non-flood, when the component's pixel value is smaller than Δ_C . In the Flood Area dataset, we observed that the median value, computed over all images in the dataset, of the percentage of pixels that belong to the non-flood class according to the mask M_L, M_A and M_B is only 2.3%, 1.3% and 2.0%, respectively. This means that the M_L, M_A and M_B masks have a very low number of wrong-classified flood pixels, providing a robust initial segmentation for our method.

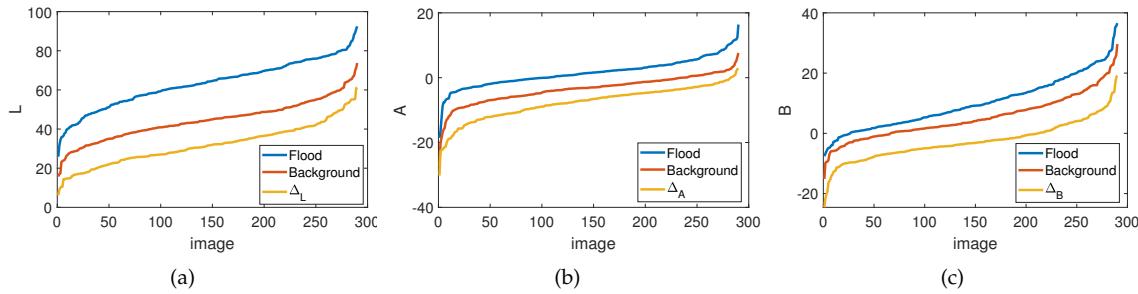


Figure 5. Blue and red curves correspond on the average value of (a) L, (b) A and (c) B color components computed on flood and background pixels respectively, for each image of the Flood Area dataset, sorted in ascending order. The yellow curves show the corresponding Δ_C threshold.

Edges often correspond to significant changes in intensity or color in an image. Detecting these edges allows for the extraction of important features, such as object boundaries, contours, and shapes, which are essential for further analysis and interpretation of image content. Identifying edges helps in highlighting important structures and details, and serves as a fundamental step in image segmentation, which involves partitioning an image into regions with similar characteristics. Edges act as boundaries between different regions, which makes them essential for accurate segmentation and analysis of image content [42]. It is clear that edges can be useful to locate borders between flooded and non-flooded areas and should therefore be excluded from the flood water class. After a simple Canny edge detection on the blurred L component so that dilated edges occur, the fifth mask M_{Edge} is obtained, labeling the edge pixels as non-flood. M_{Edge} mask is also useful, because by detecting the borders of small objects (e.g. buildings, cars, trees, etc.), then the weights of their nearby pixels are reduced in the flood-dominant color estimation, increasing the robustness of the estimation (see Section 3.4).

Examples from the Flood Area dataset of the four LAB component masks M_L, M_A, M_B and M_{Edge} are depicted in Figure 6. We observe that each color component and the detected edges contribute to further classifying image pixels as background (with dim gray color), thus excluding them from potential flood areas (cyan color). Non-flood areas are strengthened, when classified so by multiple components, but also complemented by being detected by one component when others failed to do so. For the whole Flood Area dataset, on average 25.43%, 27.75%, 29.82% and 7.45% of pixels, respectively for the L, A, B component and the edge image are classified as non-flood. Notice the sky in the first image of Figure 6, detected only by the B component as non-flood, and the hill side in the second

image, which as a non-flood area is strengthened by all components. Under any case, the borders of small objects are well detected by the M_{Edge} mask, increasing the robustness of the flood-dominant color estimation, which is described in the next Subsection.

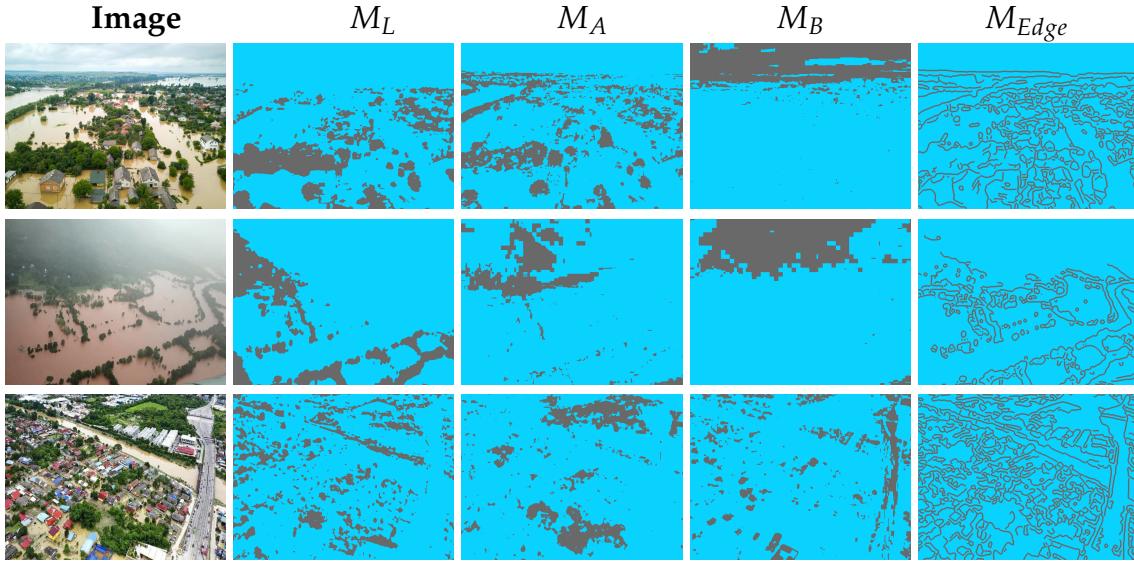


Figure 6. Original images from the Flood area dataset and their corresponding LAB components masks M_L , M_A , M_B , and M_{Edge} from left to right. Note that the edges in M_{Edge} are dilated for illustration purposes. Non-flood areas are depicted with dim gray color, whereas remaining potential flood areas are shown in cyan.

3.4. Flood Dominant Color Estimation

As described in Sections 3.2 and 3.3, five masks with potential non-flooded areas are constructed. There are cases where these masks describe approximately the same areas, but essentially the masks complement each other. When all masks are robust in classifying an area as non-flood then the conclusion for this specific area is strengthened. But if one mask is weak in characterizing the area as non-flood, the other masks function as reinforcement. The final mask (M_{Final}) for the non-flood class is derived by uniting these individual masks, as also depicted in Figure 3:

$$M_{Final} = M_{RGBVI} \cup M_L \cup M_A \cup M_B \cup M_{Edge} \quad (3)$$

The morphological operation of image closing follows, merging adjacent areas that are partially separated and connecting nearby regions. Excluding the non-flood labeled pixels of the final mask leaves us with potential flood areas, which of course have to be refined, as described below.

We opted for a weighted approach to estimate the dominant color of the flood in the image. Taking into account the non-flood area derived from M_{Final} , the Euclidean distance transform [43] assigns to each potential flood pixel a value representing its distance to the nearest boundary pixel (non-flood). This representation of the spatial relationship between the pixels serves as a weight map (W), where the pixels furthest away from non-flood areas receive greater weights such that their color (I_C) will exert a more significant influence on the process of the flood's dominant color estimation.

The weighting variance (σ_C^2) for each color component (C) is calculated on the potential flood area (PFA) as follows:

$$\sigma_C^2 = \frac{N}{N-1} \cdot \frac{\sum_{p \in PFA} W(p) \cdot (I_C(p) - \mu_C)^2}{\sum_{p \in PFA} W(p)} \quad (4)$$

where N and p are the total amount of potential flood area pixels and an image pixel respectively. The mean color value of the potential flood for each component $C \in L, A, B$ is represented by μ_C :

$$\mu_C = \frac{\sum_{p \in PFA} W(p) \cdot I_C(p)}{\sum_{p \in PFA} W(p)} \quad (5)$$

The estimated variance (σ_C^2) can be much higher than the real one, due to the fact that PFA may also contain non-flood pixels. Furthermore, the estimated σ_C^2 in the PFA region should be much lower than the corresponding variance estimated in the entire image ($\bar{\sigma}_C^2$), due to the high color similarity between the flood pixels. So, σ_C^2 is corrected to a predefined percentage (e.g. 20%) of $\bar{\sigma}_C^2$, when it exceeds this value.

In this work, the probability for potential flood area pixels over each color component is defined in Equation 6, via the exponential component of the normal probability distribution function (Gaussian kernel) that is ranged in $[0, 1]$:

$$P_C(p) = e^{-\frac{(I_C(p) - \mu_C)^2}{2\sigma_C^2}} \quad (6)$$

The decisive probability map (PM) is then constructed, as shown in Equation 7

$$PM(p) = \left(P_L(p) \sqrt{P_A(p) \sqrt{P_B(p)}} \right)^{\frac{4}{7}} \quad (7)$$

accounting for the greater significance of the L component with exponent term one. The component A has the second significance, with an exponent term equal to $\frac{1}{2}$. The component B is the least significant, with an exponent term equal to $\frac{1}{4}$. In Equation 7, the exponent term $\frac{4}{7}$ is used for normalization purposes, giving the sum of exponent terms one. The use of different exponent terms on the three color components slightly improves the results of the proposed method, achieving 0.5% and 0.3% higher F1-Score (F_1) (see Section 4.5), compared to the version of the system having equal significance to all color components (see Eq. 8) and equal significance to A and B color components (see Eq. 9), respectively.

$$PM(p) = (P_L(p) P_A(p) P_B(p))^{\frac{1}{3}} \quad (8)$$

$$PM(p) = \left(P_L(p) \sqrt{P_A(p) P_B(p)} \right)^{\frac{1}{2}} \quad (9)$$

Figure 7 illustrates examples of probability maps derived after the corresponding initialization masks and weight maps for examples from the Flood Area dataset. In fact, potential flood areas (cyan color) have a higher probability (red color) compared to the background. Areas classified as non-flood (dim gray color) by Equation 3 exhibit zero weights and probabilities (dark blue color) such that they will not contribute to the flood's dominant color estimation. However, when the flood's color is equivalent to the color of background areas, which were not eliminated by the masks used to initialize our method (see Sections 3.2 and 3.3), there can be the case of high probabilities turning up in non-flood areas, as depicted in the example in the last row, where parts of the road exhibit a tendency towards flood.

At this point, using a single threshold value on the decisive probability map to perform the final segmentation can lead to quite satisfactory results (see Table 3, row UFS-REM). However, we opted for dual thresholds, as described in the following Subsection.

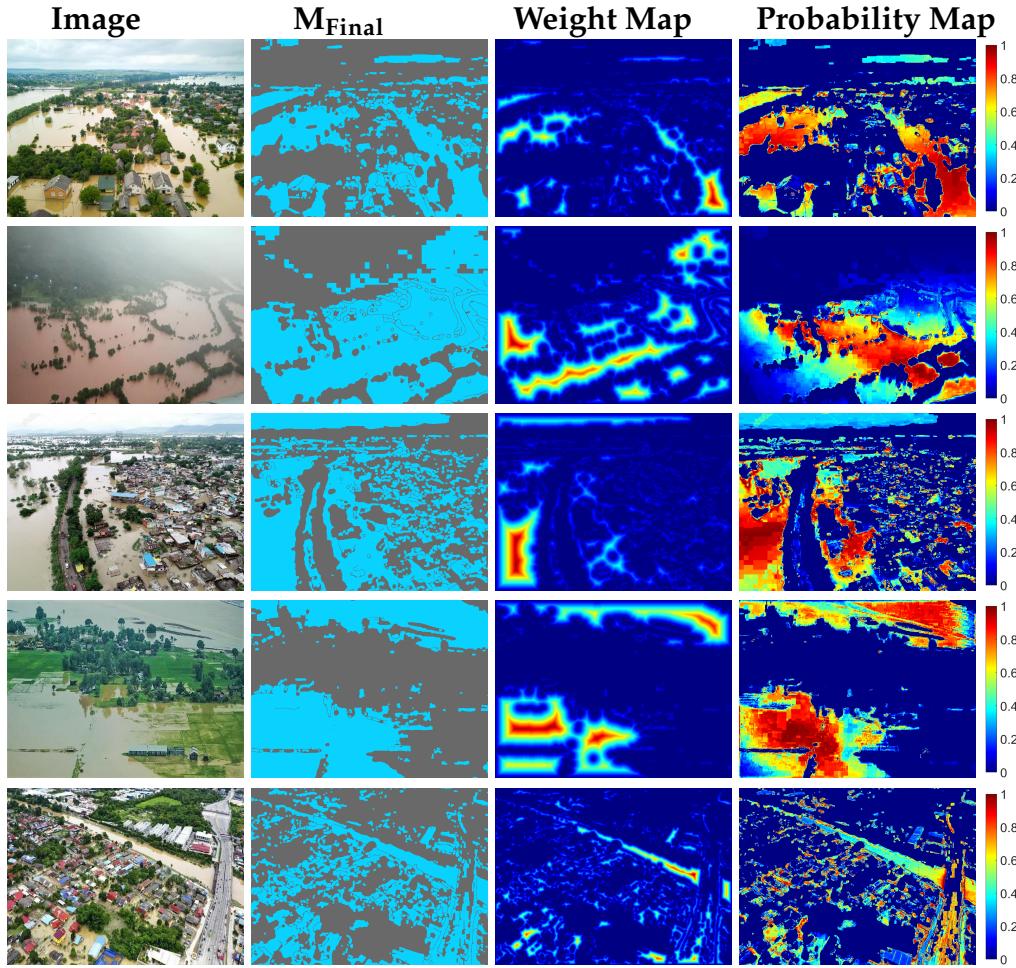


Figure 7. Probability maps (column 4) obtained using potential flood areas of M_{Final} (column 2), weight maps (column 3), as generated by the distance transform and the corresponding images from the Flood Area dataset (column 1). Potential flood area is shown in cyan, and non-flood area in dim gray color. The weights and probabilities range from 0 (dark blue color) to 1 (red color).

3.5. Hysteresis Thresholding

The probability map provides a good indication of the location of flooded areas. This now defines a probability optimization problem. To proceed to the final decision, we preferred two different threshold values to distinguish between actual flood and background. The isocontour evolves starting from the high-confidence areas and gradually grows to segment the flooded area. This technique is known as hysteresis thresholding and was first implemented for edge detection [42]. The main steps of the hysteresis thresholding method are described below:

- We adapt the process for region growing, where the high threshold T_H is applied to the entire probability map to identify pixels with $PM(p) > T_H$ as flood. These regions have high confidence to belong to flood areas, so they can be used as seeds in a region-growing process that is described below.
- Next, a connectivity-based approach is used to track the flood. Starting from the pixels identified in the first step, the algorithm looks at neighboring pixels. If a neighboring pixel has a probability value higher than the low threshold T_L , it is considered part of the flood.
- This process continues recursively until no more connected pixels above the low threshold are found.

The hysteresis effect prevents the algorithm from being too sensitive to small fluctuations in probability. Pixels that fall between the low and high thresholds, but are not directly connected to strong flood

pixels are not considered flood. However, if a weak flood pixel is connected to a strong flood pixel, it is still considered to be part of the flood. In this way, flood continuity is maintained and false detections are reduced, so that flood boundaries are accurately identified.

We have set the low and high thresholds to be at the 1% and 75% marks, respectively. But, as we can observe in Figure 12, any value in the vicinity of $T_L = 0.01$ and $T_H = 0.75$ does not change the outcome substantially, thus making our methodology robust. The middle column of Figure 8 shows the hysteresis thresholding process for four images of the Flood Area dataset in the second column. For the whole set of potential flood areas, red colored are pixels with probability greater than T_H and therefore certain to belong to the flood class, while the blue colored pixels have a probability in the range $PM(p) \in (T_L, T_H]$. For the latter, the flood class inclusion is guaranteed only when they are linked with red pixels. Cyan colored pixels belong to PFAs but fall below the lower threshold, and therefore they will be assigned to the background class. The non-flood areas, according to the M_{Final} mask, are colored with dim gray pixels.

3.6. Final Segmentation

The proposed methodology is completed by obtaining the final segmentation after the hysteresis thresholding is applied to the decisive probability map. For the derived flood areas, an edge correction is performed via the image dilation operation. The connected components of flood and non-flood areas are calculated, and relatively small areas are removed. In particular, if a blob is considered to be flooded but does not exceed about 0.3% of the whole image pixels, then this blob is reclassified as background to reduce noise effects, e.g. small water pits which do not belong to the flood area. Furthermore, background blobs with an area of 0.05% of all the pixels in the image are attributed to the flood class. These blobs can occur amidst of a flood, because of fluctuations in the values of the color components (e.g. shadows), disrupting the continuity and were wrongly classified as background and therefore excluded at the early stages of the method.

Figure 8 shows the final segmentation of our proposed approach (see last column). Blue and dim gray colors represent the segmented flood and background, respectively. Cyan colored pixels from the second column, which did not pass the thresholds, as well as blue pixels, which are not connected to red ones are excluded from the segmented flood area. The same applies to small blue blobs connected to red pixels, because they do not fit the aforementioned area criterion. In the next Section, we present more results and discuss them in detail.

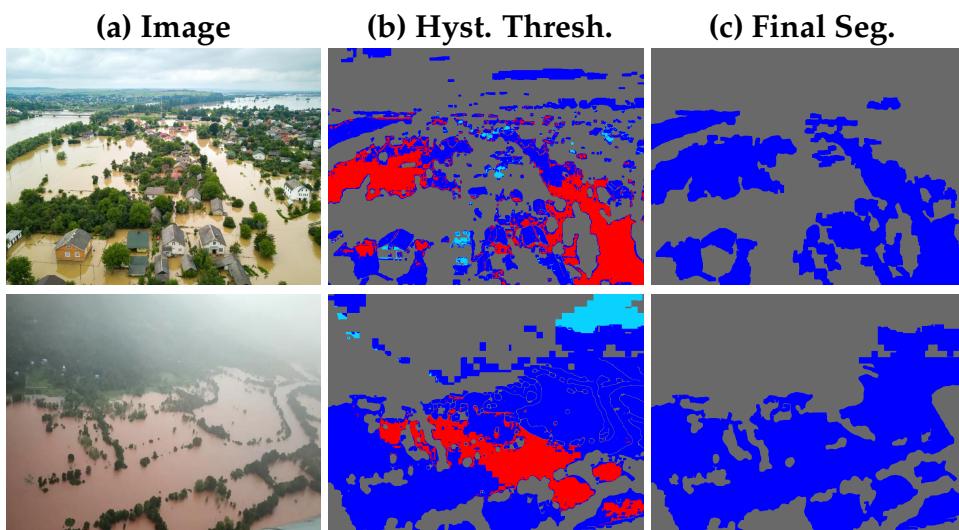


Figure 8. Cont.

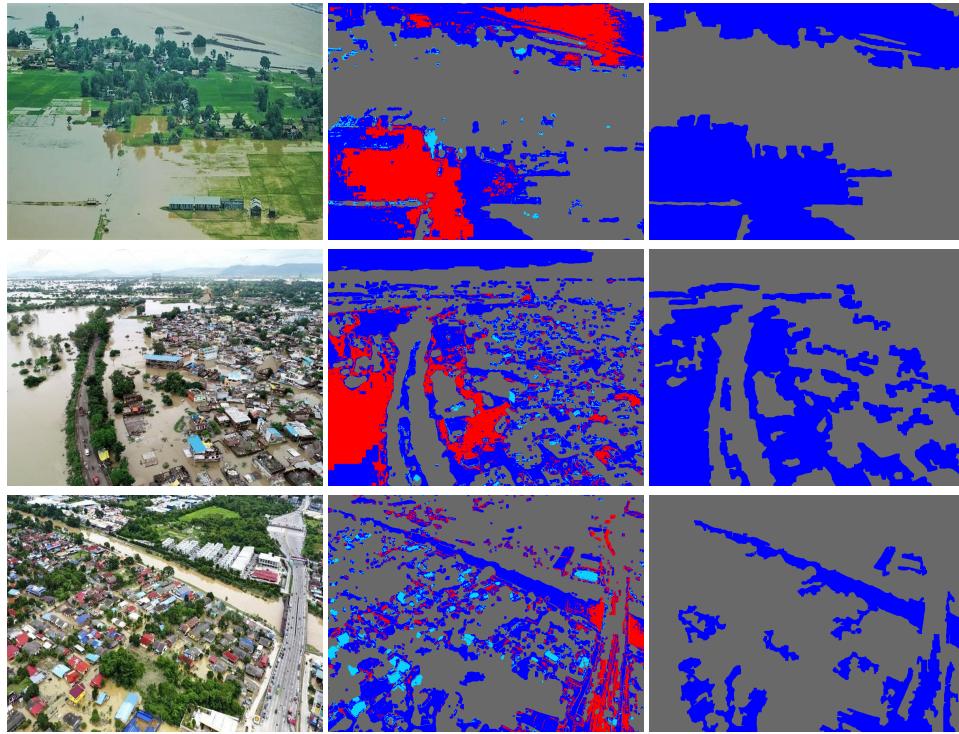


Figure 8. (a) Original image from the Flood Area dataset, (b) the applied hysteresis thresholding on the decisive probability map of the potential flood area and (c) the final segmentation mask. (b) In red and blue are the pixels with $PM(p) > T_H$ and $T_L < PM(p) \leq T_H$ respectively. Cyan colored pixels are with $PM(p) \leq T_L$, they do not surpass the lower threshold, and are subsequently classified as background. The non-flood areas, according to the M_{Final} mask, are colored with dim gray pixels. (c) The last column shows the final segmentation obtained from our proposed method, where the flood is in blue and the background is in dim gray color.

4. Result and Discussion

In this Section, we present our proposed method's results and compare them with selected recent DL approaches, since we did not find any other unsupervised method to challenge this problem. Furthermore, we have conducted an ablation study to measure the contribution of each of our method's modules, as described in Sections 3.2 to 3.6, with respect to performance.

4.1. Evaluation Metrics

The metrics used for evaluation purposes are Accuracy (ACC), Precision (PR), Recall (REC) and F1-Score (F_1), as defined in Equations 10 to 13 below:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$PR = \frac{TP}{TP + FP} \quad (11)$$

$$REC = \frac{TP}{TP + FN} \quad (12)$$

$$F_1 = \frac{2 \cdot PR \cdot REC}{PR + REC} \quad (13)$$

TP , FP , TN , and FN stand for true positive, false positive, true negative, and false negative respectively. Additionally, we have calculated the average value of F1-Score (\bar{F}_1) over the whole dataset, which is given by averaging the corresponding F1-Score of each image of the dataset.

4.2. Implementation

The proposed method has been implemented using MATLAB. All experiments were executed on an Intel I7 CPU processor at 2.3 GHz with 40 GB RAM. The proposed algorithm achieves inference in about half a second per image, without immense calculations that would require the power of GPU cores to perform in the same time range. The code implementing the proposed method, together with the dataset and results, will be publicly available (after the paper acceptance) at the following link ¹.

4.3. Flood Area Dataset: Experimental Results and Discussion

We showcase a series of final segmentations as a result of our proposed approach (UFS-HT-REM) supplemented by evaluation metrics, as described above, and the original and ground-truth images. All images have been adjusted to uniform dimensions (800×600) for illustration purposes only, since our method works for any image size.

Representative outcomes from the Flood Area dataset are shown in Figures 9, 10 and 11. The results show flood segmentation in blue overlaid on the original image with the flood's borders emphasized in dark blue. We also use the same technique to represent the ground truth. In Figure 9, the results of the proposed method yield F_1 and ACC that exceed 90%, showing high performance results with almost perfect flood detection. In Figure 10, F_1 belongs in the range [75%, 90%] providing results, where the flood area detection accuracy is satisfactory. In Figure 11, the results of the proposed method yield F_1 lower than 30%, showing poor segmentation results.

The proposed unsupervised approach works for any tint of flood water and delivers excellent results even when a plethora of small objects protrude the flood, managing to segment around them (see Figure 9 (b) and 10 (d)). As we can observe in the rest of the results, it achieves satisfactory results in urban/peri-urban as well as rural environments, where it accurately segments existing flooded areas, even when they are not labeled in the ground truth, as shown in Figure 10 (e). The method works best, when the acquired image's color quality and the weather, lighting conditions are good, alongside a low camera rotation angle (top-down view). Naturally, should the assumptions underpinning the proposed algorithm not hold, it may result in suboptimal outcomes (see Figure 11). Reasons for bad results are due to the extreme similarity of the flood and the background color, and elevated LAB color components' values of background areas. It is essential that the flood is not green in color, because the vegetation index will confuse it with greenery and therefore mainly exclude it (see Figure 11 (c)).

In the LAB colorspace, the flood has been measured to have elevated values with respect to the background. Therefore, in cases when light is reflected on the surface or the sky is of the same brightness, it is impossible for the method to set these two areas apart (see Figure 11 (a) and (b)). Finally, because the method still relies on color, if there is near-identical coloration of the flood water and other objects (e.g. buildings, rooftops), these objects will be considered part of the flood or will be entirely mistaken as flood omitting the true flood water (see Figure 11 (d)). High ACC scores in the poor segmentation results are due to correctly segmenting large areas of the background.

¹ <https://sites.google.com/site/costaspanagiotakis/research/flood-detection>

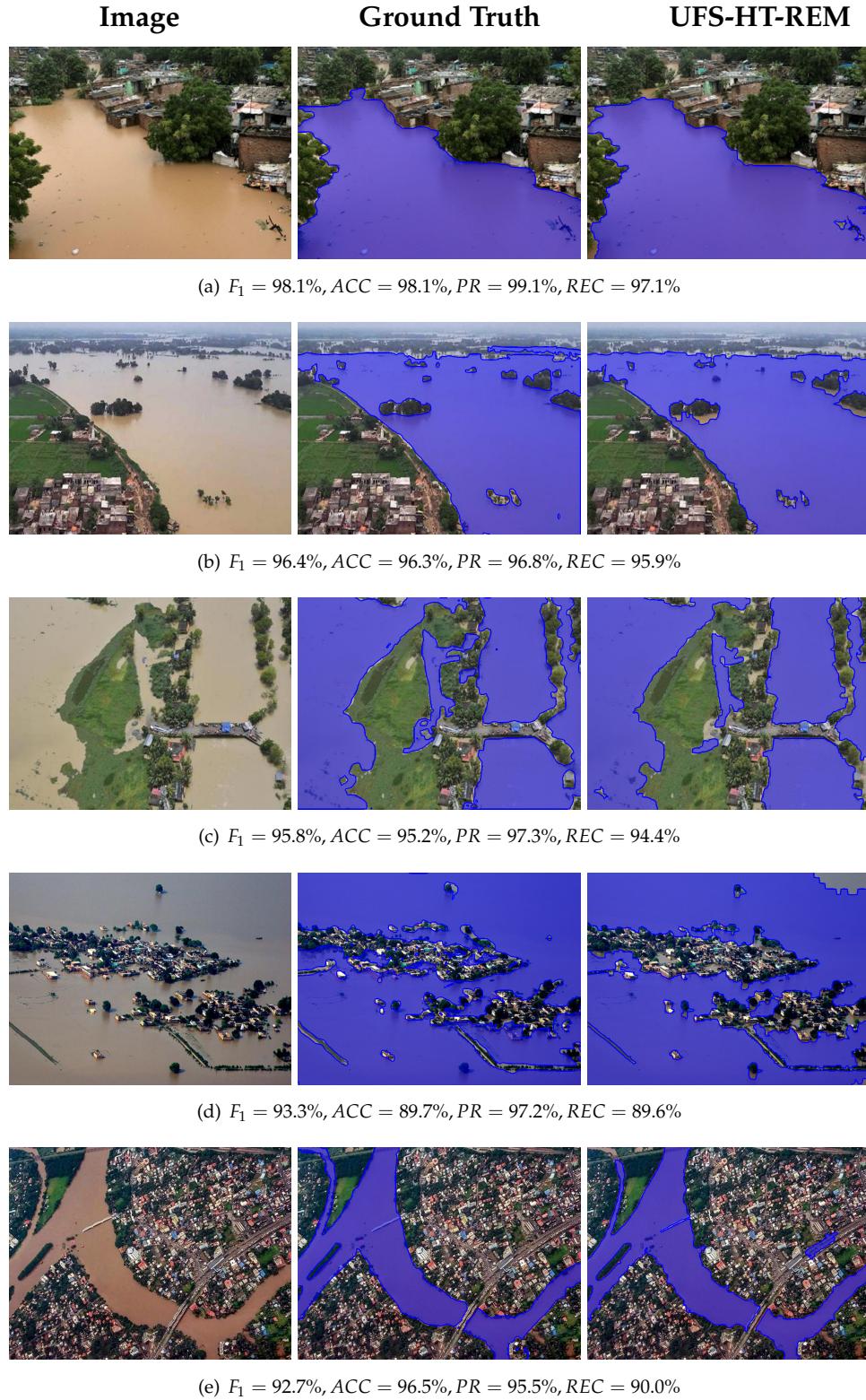


Figure 9. High performance results of the proposed flood segmentation method from the Flood Area dataset. Original images, ground truth, and the final segmentation of our proposed method (UFS-HT-REM).

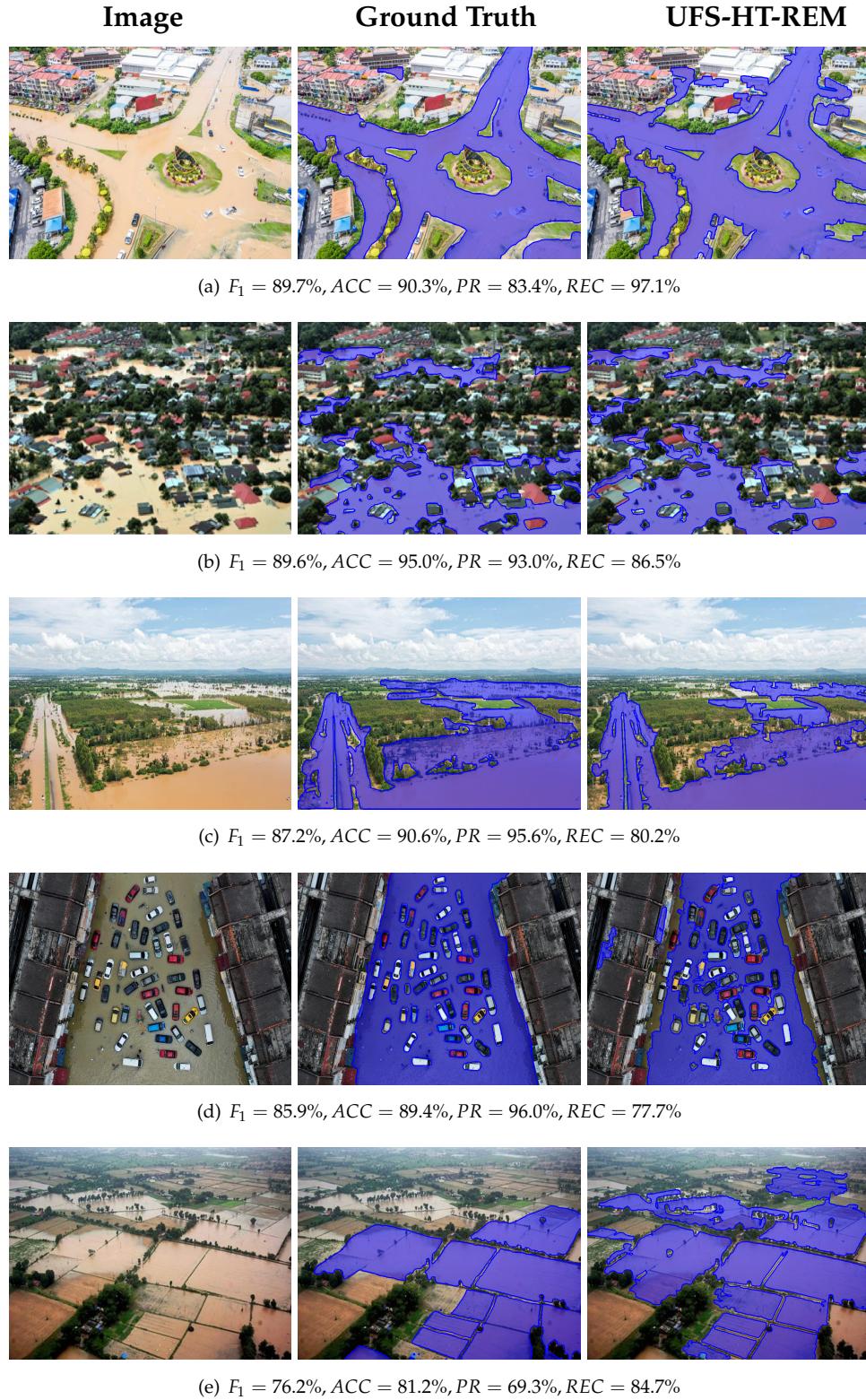


Figure 10. Satisfactory results of the proposed flood segmentation method from the Flood Area dataset. Original images, ground truth, and our proposed method's (UFS-HT-REM) final segmentation.

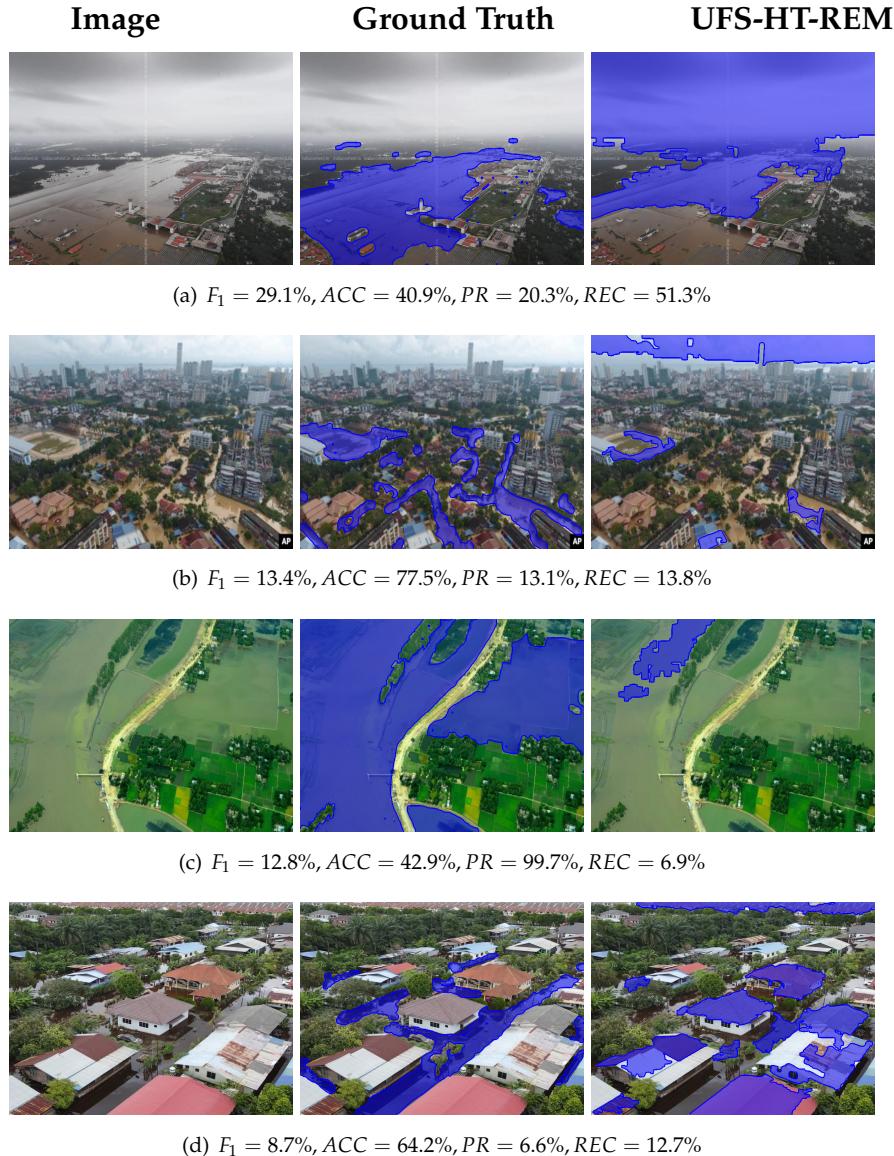


Figure 11. Poor segmentations resulting from the proposed methodology (UFS-HT-REM) from the Flood Area dataset. Original images, ground truth and the final segmentation of our proposed method.

4.4. Exploring the Impact of Environmental Zones and Camera Rotation Angle

Additionally, we study how the environmental zone and the camera rotation angle affect the flooding segmentation by splitting the Flood Area dataset into the following categories:

1. Environmental zone:

- (a) Rural, predominantly featuring fields, hills, rugged mountainsides, scattered housing structures reminiscent of villages or rural settlements, and sparse roads depicted within the images. It consists of 87 out of the 290 images in the dataset.
- (b) Urban and peri-urban, distinctly showcasing urban landscapes characterized by well-defined infrastructure, that conforms to urban planning guidelines, a dense network of roads and high population density reflected in the presence of numerous buildings and structures. It encompasses a collection of 203 images.

2. Camera rotation angle:

- (a) No sky (almost top-down view, low camera rotation angle), distinguished by the absence of any sky elements; specifically, these images entirely lack any portion of the sky or clouding within their composition. It comprises 182 images of the dataset.
- (b) With sky (bird's-eye view, high camera rotation angle), where elements of the sky, such as clouds or open sky expanses, are visibly present within the image composition. It encompasses the remaining 103 images within the dataset.

This categorization has been undertaken to emphasize that the environment, in which the flood is situated, can play a role in the final outcome and to encourage further studies to distinguish in their methodologies urban and rural floods, due to their different characteristics.

As we can observe, in the evaluation metrics for these categories, shown in Table 2, our method performs well in all categories. Slightly better results are achieved when the scenery depicts a rural landscape, with $F_1 = 80.3\%$ being 1.8% higher than in the urban/peri-urban category. The vast greenery in rural areas helps discarding a great amount of pixels from the PFAs utilizing the RGBVI mask (as described in Section 3.2). In addition, flooded areas in such environments are usually large in size, which strengthens the dominant color estimation procedure (see Section 3.4), because plentiful pixels are engaged (see category 1.(a) in Table 2). In urbanized areas, the presence of rooftops, buildings and cars, which can have a similar color to flood, leads to mildly decreased performance with $F_1 = 78.5\%$, given that parts of these objects are misclassified as flood by the algorithm (see category 1.(b) Table 2).

Table 2. Results for the categories of images existing in the Flood Area dataset. The images were divided according to the environmental zone into 1.(a) rural, and 1.(b) urban/peri-urban, and according to the camera rotation angle into 2.(a) no sky (low angle), and 2.(b) with sky (high angle).

Category	ACC	PR	REC	F_1	\bar{F}_1
1.(a) Rural	83.6%	82.3%	78.4%	80.3%	78.2%
1.(b) Urban/peri-urban	85.4%	78.4%	78.7%	78.5%	76.9%
2.(a) No sky	85.2%	81.6%	78.0%	79.8%	77.9%
2.(b) With sky	84.4%	76.1%	79.5%	77.7%	76.1%
All	84.9%	79.5%	78.6%	79.1%	77.3%

Furthermore, as we have stated previously, poor segmentation results have been generated when the image featured the sky or parts of the sky. Since the sky's LAB color component values can fall above Δ_C (as in Eq. 2), and according to our observation that flood exhibits higher values in the LAB colorspace, parts of the sky are not excluded from the PFAs in this case. Their pixels are involved in the dominant color estimation, and they are commonly segmented as flood. In Table 2, rows 2.(a) and 2.(b) present the evaluation metrics for the categories no sky and with sky, with F1-Scores being 79.8% and 77.7% respectively, in which clearly the no sky group prevails by 2.1%, as expected. This categorization culminates to the deduction that when the rotation angle of the UAV's camera is minded, so that the acquired image does not depict any part of the sky, but is facing only terrain (a fact that can be controlled by the human operator or the navigation software), the segmentation result is improved.

4.5. Ablation Study

The ablation study outlines the significance of each of the proposed method's modules and is reported in Table 3. To do so, we have reported experiments of the proposed method variants, conducted with the Flood Area dataset. The proposed method (UFS-HT-REM) includes all modules and yields the best performance in the following metrics $ACC = 84.9\%$, $PR = 79.5\%$, $F_1 = 79.1\%$ and $\bar{F}_1 = 77.3\%$.

Table 3. Ablation study highlighting the contribution of the method's modules. All experiments were conducted with the Flood Area dataset.

Method	ACC	PR	REC	F_1	\bar{F}_1
UFS-HT-REM	84.9%	79.5%	78.6%	79.1%	77.3%
UFS-HT-REM (weights by Eq. 9)	84.7%	79.2%	78.5%	78.8%	77.0%
UFS-HT-REM (weights by Eq. 8)	84.2%	78.6%	78.7%	78.6%	76.7%
UFS-HT	83.4%	76.3%	79.4%	77.8%	75.9%
UFS-REM	82.5%	75.7%	79.1%	77.3%	74.9%
UFS-HT-REM - L	79.9%	68.9%	82.0%	74.8%	72.8%
UFS	78.6%	68.9%	81.1%	74.5%	71.6%
M_{Final} Mask	76.0%	64.9%	82.7%	72.7%	69.8%
UFS-HT-REM - B	74.6%	66.3%	76.1%	70.8%	67.8%
UFS-HT-REM - A	68.6%	56.7%	88.0%	68.9%	66.0%
UFS(Otsu)-HT-REM	72.6%	66.9%	36.8%	47.5%	43.5%

First, we examine the simplification of the probability map (PM) estimation defined in Equation 7, including equally all the LAB color components (UFS-HT-REM weights by Eq. 8) and the AB color components (UFS-HT-REM weights by Eq. 9) for the decisive probability map computation. In both cases, the performance of the method is slightly degraded, since the reduction of F_1 is less than 0.5%. This shows that the proposed Equation 7 is robust and can be replaced by a simpler formula without significant changes. Additionally, it shows that the initialization process is solid, leaving in the PFAs a majority of pixels that are truly flood. From the LAB color components, luminance L is of greater importance, as proven by experiments exploiting only one color component:

- L (UFS-HT-REM-L with $F_1 = 74.8\%$)
- A (UFS-HT-REM-A with $F_1 = 68.9\%$)
- B (UFS-HT-REM-B with $F_1 = 70.8\%$)

The core methodology, as described in Sections 3.2 to 3.4, performs quite adequately (UFS), resulting 4.6% lower F_1 compared to UFS-HT-REM. Additionally, by adding either the small area removal segment (UFS-REM) or the hysteresis thresholding technique (UFS-HT) separately performance is improved about 3% compared to the UFS. It holds that the hysteresis thresholding exert slightly better influence, resulting 0.5% higher F_1 compared to UFS-REM. Overall, each of the proposed components were carefully selected to enhance the final segmentation result.

Finally, the proposed initialization of pixels as potential flood or non-flood using Δ_C thresholding (see Equation 2) is replaced by the Otsu thresholding [44] (UFS(Otsu)-HT-REM), resulting 31.6% lower F_1 compared to the UFS-HT-REM. This proves that exploiting the standard deviation of the color data relative to its central tendency of the LAB components is a crucial step of the proposed method. This is also proved by the evaluation of the M_{Final} Mask (see Eq. 3) without any other step of the proposed method, which yields $F_1 = 72.7\%$ which is only 6.4% lower compared to the corresponding F_1 of UFS-HT-REM.

To show the stability of the proposed system under different values of system parameters, we performed the following sensitivity test on the two parameters of the hysteresis thresholding technique. Figure 12 depicts the average values of ACC , REC , PR and F_1 computed on the Flood Area dataset for different values of (a) T_L with $T_H = 0.75$, and (b) T_H with $T_L = 0.01$. In any scenario, the method's performance measured with the values of F_1 and ACC is almost stable, while, as expected, only the values of REC and PR decrease slightly and increase, respectively.

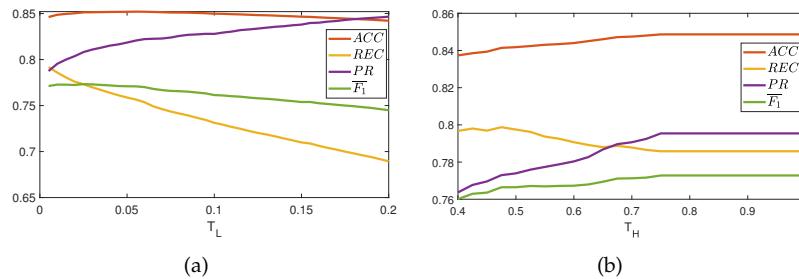


Figure 12. The average values of ACC , REC , PR , F_1 and \bar{F}_1 computed on the Flood Area dataset for different values of (a) T_L (with $T_H = 0.75$) and (b) T_H (with $T_L = 0.01$).

4.6. Comparison with DL Approaches

Compared to selected DL approaches, our results are reported in Table 4. We included the best performing FASeNet [20] ($F_1 = 90.9\%$), the intermediate scoring UNet [45] ($F_1 = 90\%$) and the worst performing so far HRNet [46] ($F_1 = 88.3\%$). Although we do not outperform any DL method, UFS-HT-REM scores in close proximity. However, our methodology does not require training. In practical terms, it does not need a significant amount of data with an accurate labeling of the ground truth, which is a subjective and time-consuming process. Also, our proposed approach is parameter-free, whereas the DL methods demand at least 640 thousand to 31 million trainable parameters to be estimated. We subside only 3.7%(9.2%) of the HRNet that performs the worst and 6.6%(11.8%) of the FASeNet in accuracy (F1-Score), respectively. It is a good compromise taking into account the simplicity of our methodology.

Table 4. Comparison of our proposed approach with selected DL approaches on the Flood Area dataset. The metrics used are Accuracy (ACC), Precision (PR), Recall (REC), and F1-Score (F_1) (calculated as in Eq. 13) expressed in percentage, and Trainable Parameters (Tr. Par.) expressed in millions (M).

Method	ACC	PR	REC	F_1	Tr. Par.
FASeNet	91.5%	91.4%	90.3%	90.9%	0.64 M
UNet	90.7%	90.0%	90.1%	90.0%	31.05 M
HRNet	88.6%	84.8%	92.0%	88.3%	28.60 M
Ours	84.9%	79.5%	78.6%	79.1%	0 M

4.7. Flood Semantic Segmentation Dataset: Experimental Results and Discussion

To demonstrate the generalization of the proposed algorithm, we used the second dataset, as described in Section 2. It comprises more than twice the number of images compared to the initial one. Without any image pre-processing and modification of the code or parameters, the algorithm performed even higher, reaching 88.5% and 81.7% in accuracy and F1-Score respectively. Consequently, we encountered a 3.6% (2.6%) increase in ACC (F_1). This demonstrates that the observations on which our method is based are universally applicable and resilient. Moreover, it proves that when controllable variables like the camera's rotation angle are considered, an observation we found to be consistent within this dataset due to the reduced occurrence of sky portions in images, segmentation outcomes exhibit enhancement. Quantitative metrics in comparison with the first dataset are presented in Table 5. Also, we provide the weighted average in the last row to give an overview of the overall performance of the algorithm in segmenting the flood for 953 images, which depict various scenes and were acquired with different camera settings, e.g. camera rotation angle, focal length, etc.

Representative results for the Flood Semantic Segmentation dataset are shown in Figure 13. Flood segmentation is shown in blue overlaid on the original image with the flood's borders emphasized in dark blue. The same technique is used to present the ground truth as well. The best flood segmentation scored 99.6% in F_1 . As we can observe, the excellent segmentations are able to capture the flood in its full or almost full extend (Figure 13 (a) and (b)). Furthermore, high performing results segment

flood details in challenging environments characterized by interference of numerous natural or man-made obstacles (Figure 13 (c) and (d)). Of course, the same issues leading to poor performance exist, as described in Section 4.3. Figure 13 (e) presents such a low-performance segmentation, where the luminosity of the sky prevents its exclusion from the PFAs. This results in the sky's pixels being assigned significant weights consequently influencing subsequent probabilities, and ultimately causing them to be missclassified as flood during the hysteresis thresholding isocontour evolution procedure.



Figure 13. Representative results of the proposed methodology (UFS-HT-REM) from the Flood Semantic Segmentation dataset. Original images, ground truth, and the final segmentation of the proposed method are shown from left to right.

Table 5. Quantitative findings for the Flood Semantic Segmentation dataset (FSSD) in comparison with the Flood Area dataset (FAD) and for the union of the two datasets (calculated as the weighted average due to the varying number of images). The number of images for each dataset (*Images*), Accuracy (ACC), Precision (PR), Recall (REC), and F1-Score (F_1) (calculated as in Eq. 13) expressed in percentage are reported.

Dataset	Images	ACC	PR	REC	F_1	\bar{F}_1
FSSD	663	88.5%	79.8%	83.7%	81.7%	79.4%
FAD	290	84.9%	79.5%	78.6%	79.1%	77.3%
FSSD \cup FAD	953	87.4%	79.7%	82.2%	80.9%	78.8%

5. Conclusions and Future Work

Overall, we presented a fully unsupervised approach for flood detection in color images acquired by flying vehicles such as UAVs and helicopters. The method progressively eliminates image regions identified as non-flood using binary masks generated from color and edge data. Our method operates in a fully unsupervised manner, with no need for training and ground truth labeling. We iteratively apply the same algorithmic steps to each color component. Subsequently, a binary map is generated for each component, discarding regions identified as non-flood and producing a final mask of potential flood areas (PFAs), refined through basic morphological operations. By weighting the pixels within the PFAs, we calculate an estimation of the dominant color of the flood, and a hysteresis thresholding technique is employed to achieve the final segmentation through probabilistic region growing of an isocontour. To the best of our knowledge, it is the first unsupervised approach to tackle this problem.

In this work, we showed that the following simple features suffice to accurately solve the problem of unsupervised flood detection. First of all, the flood's color is similar wherever it appears within the image, and this color differs from the background. Almost always, the flood's color is not green, assuming tree-like vegetation to be covered with water is extreme. Finally, in the LAB colorspace, the flooded area exhibits a higher value in at least one of the color components than the background. Color quality and camera rotation angle of the captured image contribute to the solidity of our observations, and thus a good amount of control over the flying vehicle while capturing the images will support the aforementioned inferences.

Experimental results confirmed that our proposed approach is robust, performs well in metrics, and is comparable to recent DL approaches, although not outperforming them. Furthermore, we introduced a categorization of the dataset according to the depicted scenery and camera rotation angle, into rural and urban/per-urban, and no sky and with sky, respectively. We showed that our approach performs well in all categories, it is slightly excels in segmenting floods in rural environments and is better suited for acquired images that do not contain sky, which is a controllable factor when maneuvering the UAV. The inference time is about half a second and it does not require GPU core processing capabilities. The method is suitable for on-board execution and the flood segmentation provided can be used to better guide relief efforts preventing loss of lives and mitigating the flood's impact on the infrastructure.

In future research, we plan on to extend this work, detecting flooded buildings and roads. This will lead to a refinement of already existing flood segmentations, and a correction of erroneous segmentations, which the method now produces when the observations it relies on do not apply in the image. Combined with suitable methodologies which identify buildings and roads, such as [47], and cross-correlating the results, we will be able to (a) avoid missclassifications of rooftops, building, and road pixels which have a color extremely similar to the flood, thus anticipating to attain improved outcomes and elevated scores in accurately segmenting the flood event, and (b) identify damaged buildings, when most of their circumference is adjacent to the flood, and flooded roads, when there exist discontinuities in the structure. These will help to even better assess the situation in the flood hit area and more accurately guide disaster assistance, evacuation, and recovery efforts. Furthermore, we plan to exploit the gained knowledge, in order to construct a specialized DL architectures, directing the

network's attention towards the flood and even incorporating our classical computer vision approach into hybrid deep learning frameworks, tackling the problem.

Author Contributions: The authors contributed equally to this work. Conceptualization, G.S. and C.P.; methodology, G.S. and C.P.; software, G.S. and C.P.; validation, G.S. and C.P.; formal analysis, C.P.; investigation, G.S. and C.P.; resources, C.P.; writing—original draft preparation, G.S.; writing—review and editing, G.S. and C.P.; visualization, G.S. and C.P.; supervision, C.P.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The code implementing the proposed method together with our results, and the links to the datasets are publicly available at the following link <https://sites.google.com/site/costaspanagiotakis/research/flood-detection>.

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

Abbreviations

The following abbreviations are used in this manuscript:

ADNet	Attentive Decoder Network
AI	Artificial Intelligence
APLS	Average Path Length Similarity
BCNN	Bayesian Convolutional Neural Network
BiT	Bitemporal image Transformer
CNN	Convolutional Neural Network
CV	Computer Vision
DCFAM	Densely Connected Feature Aggregation Module
DELTa	Deep Earth Learning, Tools, and Analysis
DL	Deep Learning
DNN	Deep Neural Network
DSM	Digital Surface Model
EDN	Encoder Decoder Network
EDR	Encoder Decoder Residual
EHAAC	Encoder-Decoder High-Accuracy Activation Cropping
ENet	Efficient Neural Network
HR	High Resolution
ISPRS	International Society for Photogrammetry and Remote Sensing
LSTM	Long Short-Term Memory
ML	Machine Learning
MRF	Markov Random Field
NDWI	Normalized Difference Water Index
PFA	Potential Flood Area
PSPNet	Pyramid Scene Parsing Network
RGBVI	RGB Vegetation Index
ResNet	Residual Network
SAM	Segment Anything Model
SAR	Synthetic Aperture Radar
UAV	Unmanned Aerial Vehicle
VH	Vertical-Horizontal
VV	Vertical-Vertical
WSSS	Weakly Supervised Semantic Segmentation

References

1. Algiriyyage, N.; Prasanna, R.; Stock, K.; Doyle, E.E.; Johnston, D. Multi-source multimodal data and deep learning for disaster response: a systematic review. *SN Computer Science* **2022**, *3*, 1–29.
2. Linardos, V.; Drakaki, M.; Tzionas, P.; Karnavas, Y.L. Machine learning in disaster management: recent developments in methods and applications. *Machine Learning and Knowledge Extraction* **2022**, *4*.

3. Bentivoglio, R.; Isufi, E.; Jonkman, S.N.; Taormina, R. Deep learning methods for flood mapping: a review of existing applications and future research directions. *Hydrology and Earth System Sciences* **2022**, *26*, 4345–4378.
4. Kumar, V.; Azamathulla, H.M.; Sharma, K.V.; Mehta, D.J.; Maharaj, K.T. The state of the art in deep learning applications, challenges, and future prospects: A comprehensive review of flood forecasting and management. *Sustainability* **2023**, *15*, 10543.
5. Chouhan, A.; Chutia, D.; Aggarwal, S.P. Attentive decoder network for flood analysis using sentinel 1 images. 2023 International Conference on Communication, Circuits, and Systems (IC3S). IEEE, 2023, pp. 1–5.
6. Drakonakis, G.I.; Tsagkatakis, G.; Fotiadou, K.; Tsakalides, P. OmbriaNet—supervised flood mapping via convolutional neural networks using multitemporal sentinel-1 and sentinel-2 data fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2022**, *15*, 2341–2356.
7. Dong, Z.; Liang, Z.; Wang, G.; Amankwah, S.O.Y.; Feng, D.; Wei, X.; Duan, Z. Mapping inundation extents in Poyang Lake area using Sentinel-1 data and transformer-based change detection method. *Journal of Hydrology* **2023**, *620*, 129455.
8. Hänsch, R.; Arndt, J.; Lunga, D.; Gibb, M.; Pedelose, T.; Boedihardjo, A.; Petrie, D.; Bacastow, T.M. Spacenet 8—the detection of flooded roads and buildings. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 1472–1480.
9. Hänsch, R.; Arndt, J.; Lunga, D.; Pedelose, T.; Boedihardjo, A.; Pfefferkorn, J.; Petrie, D.; Bacastow, T.M. SpaceNet 8: Winning Approaches to Multi-Class Feature Segmentation from Satellite Imagery for Flood Disasters. IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2023, pp. 1241–1244.
10. He, Y.; Wang, J.; Zhang, Y.; Liao, C. An efficient urban flood mapping framework towards disaster response driven by weakly supervised semantic segmentation with decoupled training samples. *ISPRS Journal of Photogrammetry and Remote Sensing* **2024**, *207*, 338–358.
11. Hernández, D.; Cecilia, J.M.; Cano, J.C.; Calafate, C.T. Flood detection using real-time image segmentation from unmanned aerial vehicles on edge-computing platform. *Remote Sensing* **2022**, *14*, 223.
12. Hertel, V.; Chow, C.; Wani, O.; Wieland, M.; Martinis, S. Probabilistic SAR-based water segmentation with adapted Bayesian convolutional neural network. *Remote Sensing of Environment* **2023**, *285*, 113388.
13. Ibrahim, N.; Sharun, S.; Osman, M.; Mohamed, S.; Abdullah, S. The application of UAV images in flood detection using image segmentation techniques. *Indones. J. Electr. Eng. Comput. Sci* **2021**, *23*, 1219.
14. Inthizami, N.S.; Ma'sum, M.A.; Alhamidi, M.R.; Gamal, A.; Ardhianto, R.; Jatmiko, W.; others. Flood video segmentation on remotely sensed UAV using improved Efficient Neural Network. *ICT Express* **2022**, *8*, 347–351.
15. Li, Z.; Demir, I. U-net-based semantic classification for flood extent extraction using SAR imagery and GEE platform: A case study for 2019 central US flooding. *Science of The Total Environment* **2023**, *869*, 161757.
16. Lo, S.W.; Wu, J.H.; Lin, F.P.; Hsu, C.H. Cyber surveillance for flood disasters. *Sensors* **2015**, *15*, 2369–2387.
17. Munawar, H.S.; Ullah, F.; Qayyum, S.; Heravi, A. Application of deep learning on uav-based aerial images for flood detection. *Smart Cities* **2021**, *4*, 1220–1242.
18. Park, J.C.; Kim, D.G.; Yang, J.R.; Kang, K.S. Transformer-Based Flood Detection Using Multiclass Segmentation. 2023 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2023, pp. 291–292.
19. Rahnemoonfar, M.; Chowdhury, T.; Sarkar, A.; Varshney, D.; Yari, M.; Murphy, R.R. Floodnet: A high resolution aerial imagery dataset for post flood scene understanding. *IEEE Access* **2021**, *9*, 89644–89654.
20. Şener, A.; Doğan, G.; Ergen, B. A novel convolutional neural network model with hybrid attentional atrous convolution module for detecting the areas affected by the flood. *Earth Science Informatics* **2024**, *17*, 193–209.
21. Shastry, A.; Carter, E.; Coltin, B.; Sleeter, R.; McMichael, S.; Eggleston, J. Mapping floods from remote sensing data and quantifying the effects of surface obstruction by clouds and vegetation. *Remote Sensing of Environment* **2023**, *291*, 113556.
22. Wang, L.; Li, R.; Duan, C.; Zhang, C.; Meng, X.; Fang, S. A novel transformer based semantic segmentation scheme for fine-resolution remote sensing images. *IEEE Geoscience and Remote Sensing Letters* **2022**, *19*, 1–5.
23. Wieland, M.; Martinis, S.; Kiefl, R.; Gstaiger, V. Semantic segmentation of water bodies in very high-resolution satellite and aerial images. *Remote Sensing of Environment* **2023**, *287*, 113452.

24. Bauer-Marschallinger, B.; Cao, S.; Tupas, M.E.; Roth, F.; Navacchi, C.; Melzer, T.; Freeman, V.; Wagner, W. Satellite-Based Flood Mapping through Bayesian Inference from a Sentinel-1 SAR Datacube. *Remote Sensing* **2022**, *14*, 3673.

25. Filonenko, A.; Hernández, D.C.; Seo, D.; Jo, K.H.; others. Real-time flood detection for video surveillance. IECON 2015-41st annual conference of the IEEE industrial electronics society. IEEE, 2015, pp. 004082–004085.

26. Landuyt, L.; Verhoest, N.E.; Van Coillie, F.M. Flood mapping in vegetated areas using an unsupervised clustering approach on sentinel-1 and-2 imagery. *Remote Sensing* **2020**, *12*, 3611.

27. Li, J.; Meng, Y.; Li, Y.; Cui, Q.; Yang, X.; Tao, C.; Wang, Z.; Li, L.; Zhang, W. Accurate water extraction using remote sensing imagery based on normalized difference water index and unsupervised deep learning. *Journal of Hydrology* **2022**, *612*, 128202.

28. McCormack, T.; Campanyà, J.; Naughton, O. A methodology for mapping annual flood extent using multi-temporal Sentinel-1 imagery. *Remote Sensing of Environment* **2022**, *282*, 113273.

29. Trombini, M.; Solarna, D.; Moser, G.; Dellepiane, S. A goal-driven unsupervised image segmentation method combining graph-based processing and Markov random fields. *Pattern Recognition* **2023**, *134*, 109082.

30. Yu, F.; Koltun, V. Multi-Scale Context Aggregation by Dilated Convolutions, 2016, [[arXiv:cs.CV/1511.07122](https://arxiv.org/abs/cs/1511.07122)].

31. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.Y.; Dollár, P.; Girshick, R. Segment Anything, 2023, [[arXiv:cs.CV/2304.02643](https://arxiv.org/abs/cs/2304.02643)].

32. Tarpanelli, A.; Mondini, A.C.; Camici, S. Effectiveness of Sentinel-1 and Sentinel-2 for flood detection assessment in Europe. *Natural Hazards and Earth System Sciences* **2022**, *22*, 2473–2489.

33. Guo, Z.; Wu, L.; Huang, Y.; Guo, Z.; Zhao, J.; Li, N. Water-body segmentation for SAR images: past, current, and future. *Remote Sensing* **2022**, *14*, 1752.

34. O'Mahony, N.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Hernandez, G.V.; Krpalkova, L.; Riordan, D.; Walsh, J. Deep learning vs. traditional computer vision. Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1. Springer, 2020, pp. 128–144.

35. Karim, F.; Sharma, K.; Barman, N.R. Flood Area Segmentation. <https://www.kaggle.com/datasets/faizalkarim/flood-area-segmentation>. Accessed on November 2023.

36. Yang, L. Flood Semantic Segmentation Dataset. <https://www.kaggle.com/datasets/lihuayang111265/flood-semantic-segmentation-dataset>. Accessed on April 2024.

37. Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M.L.; Bareth, G. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation* **2015**, *39*, 79–87.

38. Markaki, S.; Panagiotakis, C. Unsupervised Tree Detection and Counting via Region-Based Circle Fitting. ICPRAM, 2023, pp. 95–106.

39. Ashapure, A.; Jung, J.; Chang, A.; Oh, S.; Maeda, M.; Landivar, J. A comparative study of RGB and multispectral sensor-based cotton canopy cover modelling using multi-temporal UAS data. *Remote Sensing* **2019**, *11*, 2757.

40. Chavolla, E.; Zaldivar, D.; Cuevas, E.; Perez, M.A. Color spaces advantages and disadvantages in image color clustering segmentation. *Advances in soft computing and machine learning in image processing* **2018**, pp. 3–22.

41. Hernandez-Lopez, J.J.; Quintanilla-Olvera, A.L.; López-Ramírez, J.L.; Rangel-Butanda, F.J.; Ibarra-Manzano, M.A.; Almanza-Ojeda, D.L. Detecting objects using color and depth segmentation with Kinect sensor. *Procedia Technology* **2012**, *3*, 196–204.

42. Canny, J. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence* **1986**, pp. 679–698.

43. Fabbri, R.; Costa, L.D.F.; Torelli, J.C.; Bruno, O.M. 2D Euclidean distance transform algorithms: A comparative survey. *ACM Computing Surveys (CSUR)* **2008**, *40*, 1–44.

44. Xu, X.; Xu, S.; Jin, L.; Song, E. Characteristic analysis of Otsu threshold and its applications. *Pattern recognition letters* **2011**, *32*, 956–961.

45. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18. Springer, 2015, pp. 234–241.

46. Sun, K.; Xiao, B.; Liu, D.; Wang, J. Deep high-resolution representation learning for human pose estimation. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 5693–5703.
47. Grinias, I.; Panagiotakis, C.; Tziritas, G. MRF-based segmentation and unsupervised classification for building and road detection in peri-urban areas of high-resolution satellite images. *ISPRS journal of photogrammetry and remote sensing* **2016**, *122*, 145–166.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.