The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-2/W8-2024
8th International ISPRS Workshop LowCost 3D - Sensors, Algorithms, Applications, 12–13 December 2024, Brescia, Italy

Synergistic image and point cloud processing of UAV data for urban flood modeling: point cloud smart thinning and curb mapping

Pedro Alberto Pereira Zamboni[1], Hanne Hendrickx[2], Dennis Sprute[3], Holger Flatt [4], Muhtasimul Islam Rushdi[5], Florian Brodrecht[6], Anette Eltner[7]

[1] Institute of Photogrammetry and Remote Sensing, Dresden University of Technology, Germany - pedro.zamboni@tu-dresden.de,
[2]Institute of Photogrammetry and Remote Sensing, Dresden University of Technology, Germany - hanne.hendrickx@tu-dresden.de
[3] Fraunhofer IOSB, Industrial Automation Branch (IOSB-INA), Germany - dennis.sprute@iosb-ina.fraunhofer.de,
[4] Fraunhofer IOSB, Industrial Automation Branch (IOSB-INA), Germany - holger.flatt@iosb-ina.fraunhofer.de)
[5]SPEKTER GmbH, Germany - muhtasim.rushdi@spekter.de
[6]SPEKTER GmbH, Germany - florian.brodrecht@spekter.de
[7]Institute of Photogrammetry and Remote Sensing, Dresden University of Technology, Germany - anette.elttner@tu-dresden.de

Keywords: Flood modeling, image segmentation, point cloud processing, curb mapping.

Abstract

We propose an integrated approach for automatic point cloud thinning and curb mapping in Uncrewed Aerial Vehicle - Structure from Motion (UAV-SfM) point clouds to enhance hydrological modeling in flood-prone urban areas. UAV flights were conducted to generate an initial orthoimage, which was used to train a convolutional neural network (CNN) segmentation model. The trained model was then applied to the UAV images to produce two binary mask sets: one for vegetation and one for streets and sidewalks. These masks were incorporated during photogrammetric 3D reconstruction to estimate camera geometry and generate a dense point cloud. Our results show that vegetation masks did not improve camera geometry estimation. However, by applying UAV masks, we achieved a 15% reduction in total processing time and decreased the number of points by a factor of 2.7. This targeted approach enabled curb detection by focusing on expected curb locations. Curb candidate points were proposed using geometric characteristics of the point cloud, including normal values, linearity, and verticality. Our rule-based method effectively mapped even subtle curb features, providing a rapid, cost-effective solution for large-area curb mapping. Further, we explored the potential of random forest for curb mapping, with promising results. Our approach can support urban flood modeling efforts and strengthen urban resilience for flood-prone communities.

1. Introduction

Hydraulic models are essential tools for identifying flood-prone areas, thus enabling stakeholders to make well-informed decisions. These models rely on high-resolution Digital Elevation Models (DEMs) to accurately simulate water flow dynamics and assess flood risk. However, DEMs often lack the resolution to accurately represent relevant objects, such as curbs, which are important hydraulic structures responsible for directing the runoff down to inlets or other drainage infrastructure. The omission of curbs in hydraulic modeling can result in an underestimation of runoff volumes (Halama et al., 2023). To address this issue, one potential approach is to use high-resolution point clouds for accurate curb mapping.

Uncrewed aerial vehicles (UAVs) equipped with RGB sensors represent a cost-effective platform for the generation of high-resolution point clouds through the utilization of Structure from Motion (SfM) and Multi-View Stereo (MVS) (Eltner et al., 2022). Nevertheless, the generation of high-density point clouds with high levels of detail and precision for large areas can prove to be a very time-consuming and hardware-exhausting task. Segmentation masks produced for the raw UAV images can be applied during point cloud processing to tackle this issue. Previous studies have demonstrated the effectiveness of applying image masks to UAV imagery to filter out specific objects during the SfM process (Park et al., 2022, Pashaei et al., 2023). Furthermore, image segmentation masks can be used to smartly calculate point clouds with different resolutions during the dense calculation by considering image masks for different point densities. This approach has the potential to reduce processing time compared to uniformly handling the entire area at high point density.

Efforts have been made to curb segmentation from point clouds. Researchers have combined point normal analysis with point height difference to map curbs from mobile laser scanning (MLS) data (Zou et al., 2024). Others performed curb segmentation in MLS data based on scanning trajectory, point normals, and point height differences (Na Wang and Zhang, 2022). Moreover, neural networks have been used to segment curb points from LiDAR point clouds (Zhao et al., 2024). While much of the existing research focuses on curb mapping, it predominantly utilizes LiDAR data rather than UAV-SfM point clouds, which offer a less expensive and more accessible alternative. Moreover, there is a lack of research dedicated to the mapping of curbs over large areas, with a specific focus on the hydraulic characteristics of curbs, such as their shape and height. Furthermore, there is a need for strategies to reduce point cloud generation time, especially regarding large areas.

The principal objective of this study is to develop an automated workflow for smart thinning and curb extraction from UAV-SfM point clouds, through the integration of image and point cloud processing techniques. The specific objectives of this study are threefold: 1) to evaluate the impact of using segmentation masks on vegetated areas on the geometric accuracy of the SfM process; 2) to automatically thin point clouds by applying masks to retain high point density only in the most

relevant areas during the MVS dense matching; and 3) to propose an algorithm that automatically maps curbs for hydrological modeling based on point cloud data. A deep learning image segmentation model is employed to generate masks for the UAV images. Vegetation masks are applied during the camera geometry estimation to minimize incorrect tie points during image matching. Meanwhile, street area masks are used to create point clouds with varying densities. From the high-density point clouds, curb points are filtered using rule-based and random forest approaches.

2. Material and methods

We first conducted UAV flights to collect images, creating an initial orthoimage to train and evaluate a deep-learning image segmentation model. Using the trained model, we generated class-specific masks on the UAV images to improve the SfM and MVS accuracy and efficiency during image matching and dense point cloud generation. Finally, we applied two strategies (rule-based and random forest) to map curb points using a high-resolution street and sidewalk point cloud. Figure 1 presents the general workflow.

2.1 Data acquisition

UAV flights were conducted in a small urban area in southern Germany using a Matrice 300RTK equipped with a P1 full-frame camera featuring a 35 mm focal length and a 4.4 μm pixel pitch. The flights occurred under suboptimal lighting conditions in the late afternoon, resulting in long shadows and necessitating the use of a high ISO setting (800–1600). The flight height was 120 m, resulting in a ground sampling distance (GSD) of 1.47 cm. The side and forward overlap was set to 70%. In total, 341 images were collected. Nine ground control points (GCP) were distributed and twelve permanent structures, such as manholes, were used as checkpoints (CP). The root mean square error (RMSE) on the CP was 2 cm.

2.2 Deep learning for image segmentation

An initial orthoimage of the study region was generated, subsequently, the orthoimage was manually labeled and employed to train a deep learning segmentation model. The CFNet network (Zhang et al., 2019) was used with the EfficientNetB4 (Tan and Le, 2019) as the backbone. CFNet is CNN that employs a Co-occurrent feature model to investigate the nuances of context in image segmentation tasks. In regard of CNNs, the term backbone refers to a network used to extract relevant features to encode the input into a feature representation. EfficientNets constitute a group of models that learn their structures through the use of compound scaling methods, thereby potentially increasing efficiency as the network can be optimized in size and processing time.

The orthoimage was labeled into 23 classes merged into five land uses. The semantic classes were crafted by hydrologists to incorporate features and characteristics critical for accurate flood modeling. The street land use comprised street field, street paved, street semi-paved, street unpaved, sidewalk, inlet, and manhole. The river and trenches land use included pipes, bridges, and water. The residential land use consisted of roof flat, roof tiled, roof green, roof tin, building site, garden green, garden stone, sealed surface, and photovoltaic systems. The rural land use considered farmland, greenland, and woodland.

The labeled orthoimage was split into small tiles of 256 x 256 pixels for model training. The model was implemented in TensorFlow and trained for a maximum of 50 epochs (Figure 2). The loss function was set to Sparse Categorical Cross-Entropy, and Adam was selected as the optimizer, using class weight. Class weights were used to reduce class imbalance on the dataset, giving a higher weight for underrepresented classes, and a lower weight for overrepresented classes. Model performance was evaluated in terms of pixel accuracy (equation 1) and Intersection over Union or IoU (equation 2).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$IoU = \frac{TP}{TP + FP + FN} \tag{2}$$

We used the trained segmentation model to generate masks on the UAV images. Two sets of binary masks were created: vegetation masks and street/sidewalk masks. Image segmentation masks were used during point cloud generation in different steps. A series of morphological operations (e.g., erosion and dilation) was used for post-processing to ensure more defined objects.

2.3 3D reconstrucion

We created three point clouds of the area of interest. The first point cloud was generated using a standard processing pipeline without binary masks. In this process, image alignment during SfM was done using all pixels, and the entire area was processed with a high-density point cloud. The second point cloud was created using vegetation masks to exclude these regions during the image matching, and processing the entire area at a high point resolution. The hypothesis is that the image orientation is improved if alignment is based solely on features detected and matched in non-vegetation image regions. The third point cloud was generated by applying both vegetation and street/sidewalk masks. Vegetation masks were used as in the second point cloud, while street and sidewalk masks were applied during the MVS dense matching step. This approach enabled the creation of a high-density point cloud specifically in areas of interest for curb detection, while the medium resolution was maintained for DEM calculations in the remaining reconstructed area.

All processing was conducted using a workstation with 128 GB of RAM, i9-13900k CPU, and a NVIDIA 3070 RTX graphic card.

2.4 Curb extraction using rule based approach

We extracted curb points from the high-density street point cloud using a set of filters. Initially, we applied the cloth simulation filter (Zhang et al., 2016) to isolate non-ground points (i.e., outliers) in the street point cloud. Next, we filtered points based on the confidence values extracted from Agisoft Metashape. Higher confidence values indicate that a point has been confirmed by multiple images, making its measurement more reliable. Afterward, we identified curb points based on point normal values. Point normals represent the orientation of the surface at each point. Abrupt changes in normal values can indicate corners and edges.

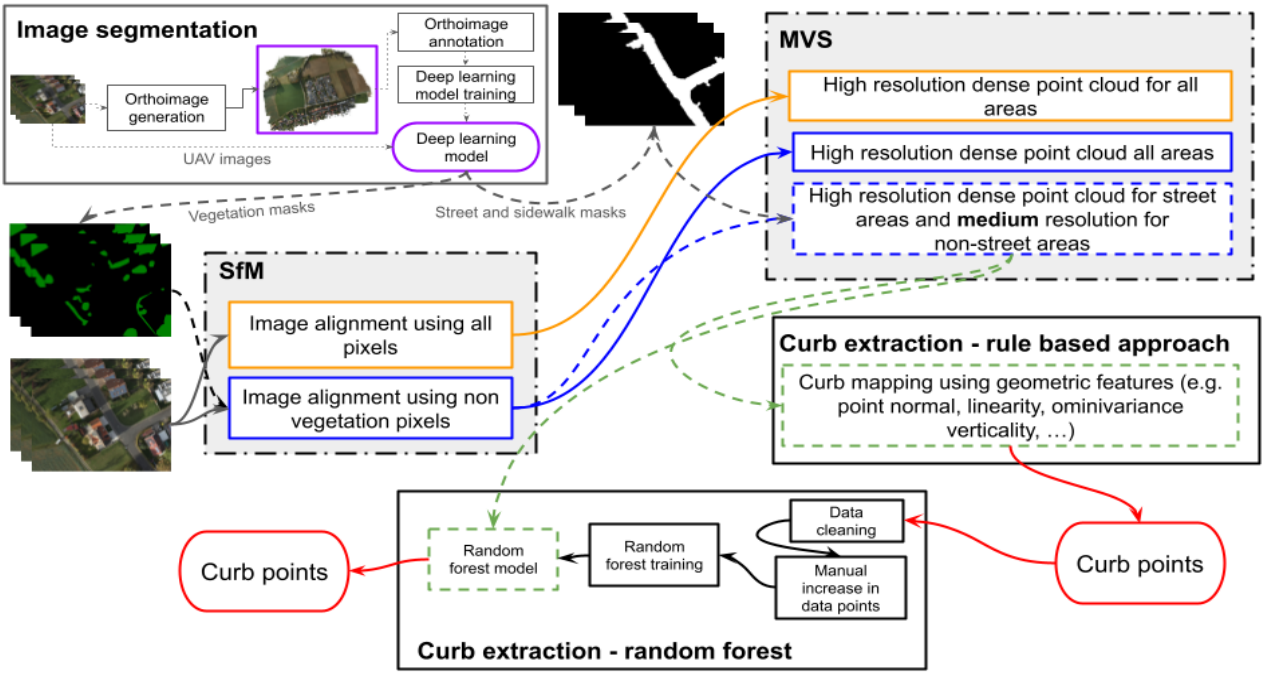

Figure 1. Study Workflow. An initial orthoimage is generated, which is then labeled and used to train a deep learning segmentation model. The trained model is used to generate vegetation and street and sidewalk masks for the original UAV images. During the image alignment stage, vegetation masks are utilized to prevent matching points from being identified in these areas. Subsequently, street masks are employed to generate dense point clouds with varying point densities, keeping high point density for the street and sidewalk area, and medium or the rest of the area. Finally, curb points are segmented on the street and sidewalk point cloud using rule-based and random forest approaches.

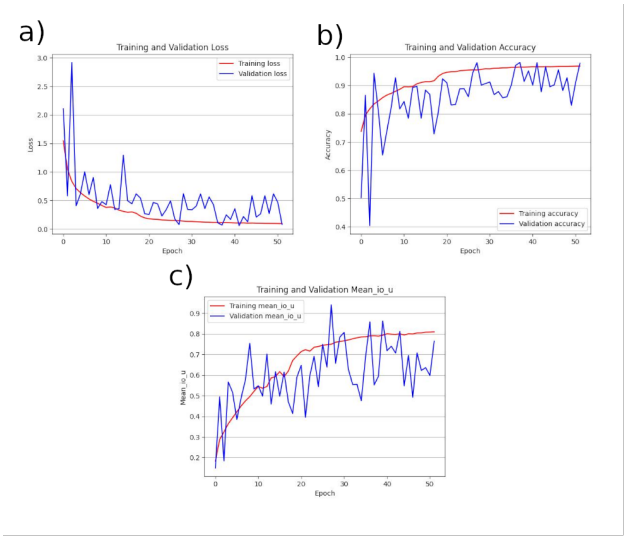

Figure 2. Training and validation metrics. a) loss; b) accuracy; and c) mean IoU. These metrics were used to evaluate the behavior of the network during training and validate training procedures.

We calculated geometric features using a search radius of 0.2 meters, emphasizing linearity, volume density, principle component analysis (PCA), verticality, omnivariance, and first-order momentum. These features are used to reduce remaining faulty curb points by encoding important geometric characteristics of the curbs. Furthermore, curb points were clustered using the HDBSCAN method to reduce remaining outliers in the data. Finally, we computed the curvature for each point, retaining only those with curvature values above the 40th percentile, followed by a down-sampling procedure and a non-maximum suppression.

2.5 Curb mapping using random forest

Random forest (RF) is a bagging algorithm that combines a set of decision trees trained in parallel of a subset of a given dataset. RF takes predictions from different trees and, based on the majority votes of predictions, takes the final predictions.

We applied a RF classifier to automatically map the curbs using the knowledge from section 2.4. We manually remove wrongly assigned curb points from our rule-based approach (section 2.4), and add more 103 points, to create the train and test dataset, with a total of 326 points. Further, we added 1500 negative samples to the dataset. Dataset was split considering 60% for training and validation, and 40% for testing.

We calculate the same set of features from section 2.4 using a multi-scale approach, with a search radius of 0.1, and 0.2. We applied a five-fold cross-validation to define best model parameters (number of trees in the forest and maximum depth of the tree). We used 50 trees and a maximum depth of 50. Further, the classification threshold was decided based on precision (equation 1), recall (equation 3), and F1 score (equation 4).

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1score = \frac{2TP}{2TP + FP + FN} \tag{4}$$

To produce the final curb classification, we applied the trained model over the entire study area. For that, we filter the point cloud using the confidence and the z-point normal values, to narrow down the search area. We visually inspected the final classification to decide the best classification threshold (0.4) to be used when applying the model over the entire area. Afterward, we filter the predicted curb points by computing the curvature for each point, followed by a down-sampling and a non-maximum suppression.

3. Results and discussion

First, we present the results of the image segmentation model. Then, we compare the accuracy and processing time of point cloud generation with and without image masks. Finally, we display the curb mapping using only the street point cloud.

3.1 Segmentation model performance

Training and validation loss, accuracy, and IoU can be seen in Figure 2. The CFNet model achieved an overall accuracy of 0.911, and the mean IoU was 0.738 on the test set. The segmentation model performed well for most of the classes. Only certain classes, such as water, inlet, and garden stone, displayed notably lower IoU scores (Table 1).

Class	Land use	IoU
Background	-	0.883
Building site	Residential	0.755
Farmland	Rural	0.854
Garden stone	Residential	0.574
Greenland	Rural	0.797
Inlet	Street	0.536
Manhole	Street	0.676
PV system	Residential	0.921
Roof flat	Residential	0.712
Roof tiled	Residential	0.885
Roof tin	Residential	0.572
Sealed surface	Street	0.724
Sidewalk	Street	0.625
Street field	Street	0.686
Street paved	Street	0.923
Street semipaved	Street	0.757
Street unpaved	Street	0.869
Water	River and trenches	0.450
Woodland	Rural	0.829

Table 1. Segmentation results (IoU) for relevant classes on the test set.

Figure 3 shows examples of the segmentation results. The results demonstrate that our model was able to segment large structures (e.g., buildings) as well as smaller objects, such as manholes and inlets. The use of class weights is especially beneficial for small objects or underrepresented classes, as these classes and objects receive more ”importance” during model training.

Our results indicate limitations in the segmentation of similar object classes with different land use. For instance, we observed that the model was unable to discriminate between high vegetation in green gardens and woodlands. Furthermore, object occlusion poses a challenge for the model.

We used the trained model to create segmentation masks for the original UAV images. Some artifacts are visible in the segmented UAV masks (Figure 4). These artifacts appear during the merging of patch masks. The UAV images have a higher resolution compared to the patches used for model training. Therefore, UAV images need to be split and merged again during UAV mask generation. One solution would be to consider overlapping patches.

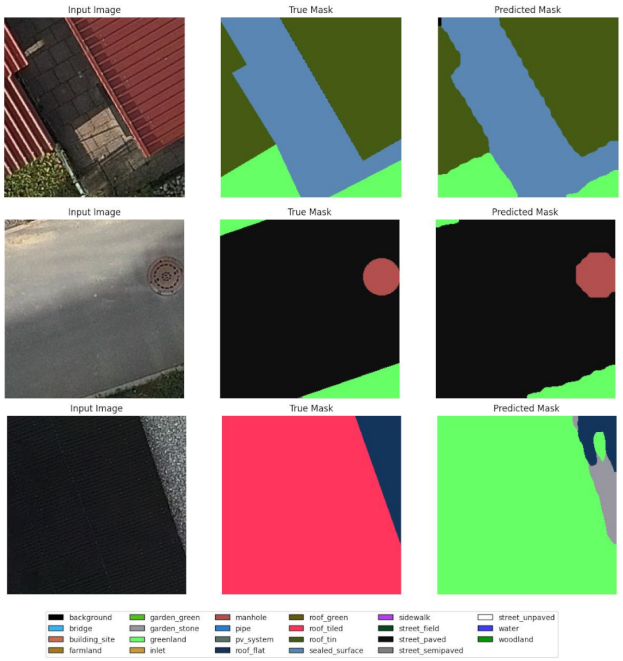

Figure 3. Example of predicted segmentation masks on the test patches. On the top role and middle rows, we observed that the model was able to segment large and small objects accurately. On the last row, is an example of a patch where the model did not produce an accurate prediction.

For point cloud generation, we created two sets of binary masks: one for vegetation and another for streets and sidewalks (Figure 4).

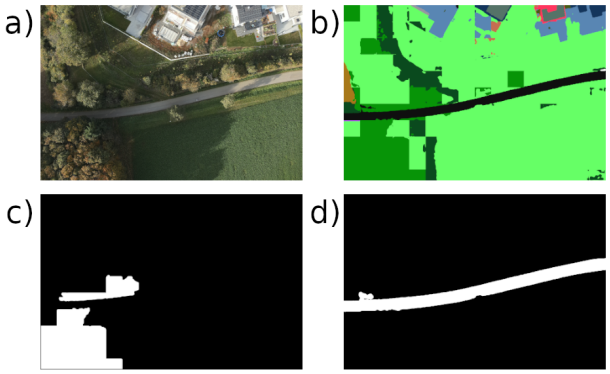

Figure 4. Examplary Segmentation of an original UAV image using our segmentation model. The formation of some square objects can be observed due to the image’s high resolution. Further, it can be noticed only part of the high vegetation was segmented. In a) original RGB UAV image; b) segmented UAV image; c) vegetation maks; and d) street and sidewalk mask.

3.2 SfM and MVS processing time and accuracy

The calculation of three different point clouds was done and compared in terms of accuracy and processing time. Although

	X error (cm)		Y error (cm)		Z error (cm)		Total error (cm)	
	ME	SDE	ME	SDE	ME	SDE	ME	SDE
Without mask	-0.005	0.007	0.01	0.008	0.012	0.015	0.024	0.007
With masks	-0.005	0.006	0.008	0.008	0.025	0.021	0.033	0.013

Table 2. The accuracy (Mean Error, ME) and precision (Standard Deviation of Error, SDE) on 12 Check Points, with and without the use of masks.

Point cloud	Matching time [min]	Alignment time [min]	Depth maps time [min]	Dense matching time [min]	Total processing time [min]	Number of Points [million]
1	75	2	55	31	161	≃408
2	75	2	41	32	150	≃ 397
3	75	2	37 (10*)	5 (7*)	119 (17*)	≃ 58 (92*)

Table 3. Point cloud processing time with and without the use of masks. Point cloud 1: Without mask; 2: with vegetation masks during image alignment, 3: with vegetation masks during image alignment, and streets and sidewalks during the matching (non-street areas processed in medium setting). For point cloud 3, we display results for the street and sidewalk area, and for non-street areas, denoted by *.

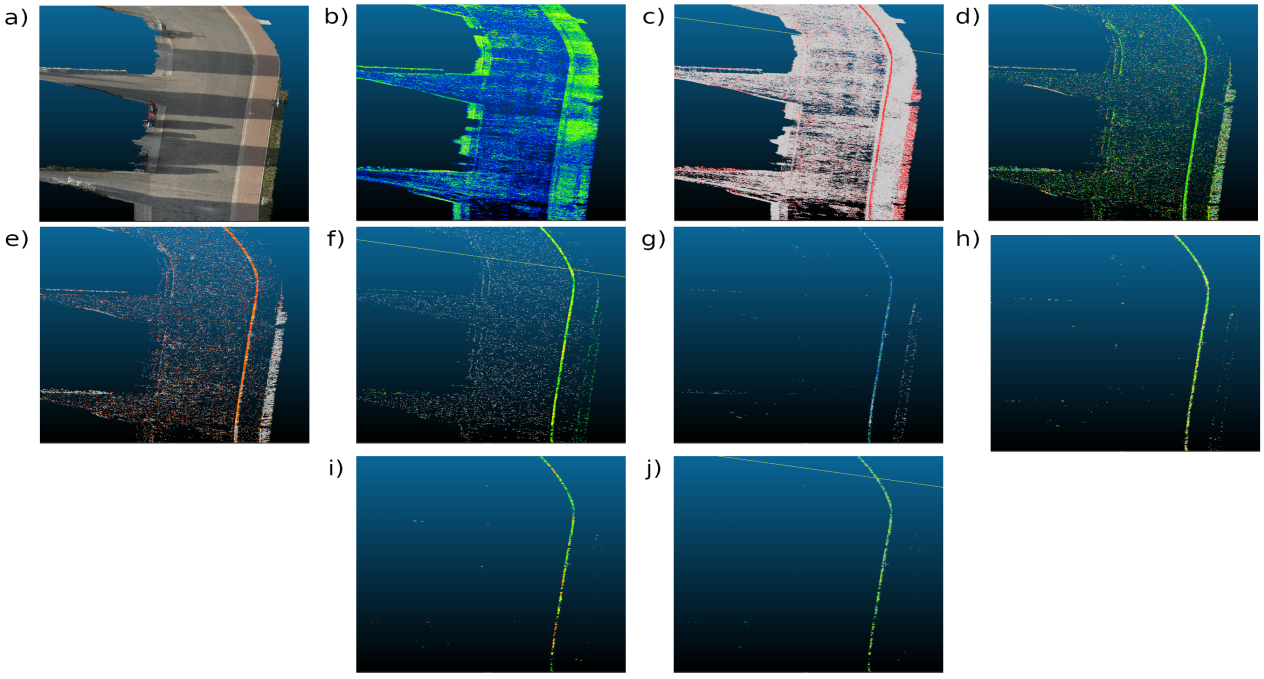

Figure 5. Result of the different processing steps for curb detection. a) Street and sidewalk point cloud; filter applied to the features b) confidence; c) z-normal; d) y-normal; e) linearity; f) surface density; g) omnivariance; h) PCA 2; i) verticality; and j) first order momentum.

the consideration of no masks (i.e., scenario 1) achieved slightly better accuracy in the Z direction during the photogrammetric 3D reconstruction (Table 2), masking of streets and sidewalks significantly reduced processing time during the dense matching, i.e., particularly during the depth map and dense point cloud generation stages.

We observed an unexpected increase in Z error with the use of the vegetation mask (Table 2). The classes used for image labeling and model training were designed by hydrologists, focusing on objects relevant to hydrological processes and for future application in flood modeling. Objects were labeled in more than one class, reflecting different land uses. For instance, high vegetation can be found in both woodland and greenland classes, where the woodland focuses on high vegetation and greenland consists of a mix of grass and high vegetation. Thus, we were unable to fully mask high-vegetation areas during the image matching process. Furthermore, the vegetation land use represented a small fraction of the overall data, which could explain the lack of improvement in CP accuracy. However, considering the GSD of 1.4 cm, all generated point clouds demonstrated satisfactory accuracy and precision performance for the flood models.

Applying vegetation masks did not significantly affect the matching and alignment time during the SfM processing, as vegetation covered only a small portion of the study area, resulting in minimal exclusion during these steps (Table 3). However, the depth map estimation time was reduced by 14 minutes when using vegetation masks alone and by 8 minutes when both vegetation and street/sidewalk masks were applied. By applying high resolution only to street and sidewalk areas and medium resolution to non-street areas, we observed a reduction of 25 minutes in total processing time (Table 3). In this setup, the number of points in the point cloud decreased by a factor of 2.7.

3.3 Curb extraction using rule-based approach

Our rule-based approach selected potential curb points from the street and sidewalk point cloud. We reduced the number of outliers with each filter. We empirically defined threshold values for each filter, adjusting them to refine the candidate curb points.

The initial point cloud consisted of approximately 58 million points. In the first step, we retained points with confidence values higher than 7.5 (Figure 5), reducing the number of points to around 5 million. Subsequently, we used a Z-point normal filter, keeping points with values between 0.82 and 0.96. The Z-normal filter can be effectively used to recognize vertically oriented objects. Flat surfaces have low Z-normal values, while vertical surfaces, such as curbs, have larger values (Figure 5). Afterward, a Y-point normal filter was applied. After both normal value analyses, about half a million potential curb points remained.

A set of geometric features was calculated to further filter the remaining points, which significantly improved the selection of curb points. Linearity measures the alignment of points to a straight line or edge. Curbs typically form long lines along the boundary between streets and sidewalks (Figure 5), where linearity can help discern curb points from more amorphous objects (e.g., cars). We retained points with linearity values greater than 0.74, of which 0.2 million points were preserved.

The surface density feature was applied to remove points in areas with a lower point density. We observed that curb areas exhibited a higher point density (i.e., surface density values higher than 174) compared to other regions (e.g., streets). Omnivariance was used to filter out flat surfaces that were incorrectly assigned as curb points (Figure 5). After these two steps, 50 thousand curb points remained.

The second largest eigenvalue from the PCA of points within a specified search radius reflects the planar spread in a neighborhood and was thus also considered. Curb points showed intermediate values (from 0.063 to 0.148) since they are not entirely flat, considering a search radius of 0.2 meters. Next, verticality was used to delete the remaining points from flat surfaces. Due to their geometric characteristics, curbs have intermediate to high values of verticality. The first-order momentum measures the distribution of points relative to a centroid, indicating alignment along a dominant direction. Curbs generally exhibit a relatively strong alignment in a consistent direction. In this sense, first-order momentum helps to refine curb points, resulting in 18 thousand curb points. Nevertheless, after all the rule-based decisions, it can be seen that still some outliers remain.

Furthermore, due to the original density of the point cloud, even after applying the filters, we observed clusters of points within small distances, necessitating a further reduction in the number of candidate points. We conducted a curvature analysis, followed by down-sampling. Finally, we applied non-maximum suppression to reduce overall noise in the final curb points.

The final curb points can be seen in Figure 7. Our approach successfully extracted curb points, even for very smooth curbs, with a five-centimeter difference between the lowest and highest points. However, we observed outliers in regions outside, primarily, the street area. In these areas, we observed a strong change in curvature close to the street due to vertical structures (e.g., walls, cars, or vegetation). These points can be removed either by reducing the street and sidewalk mask or by classifying the point cloud. Additionally, high vegetation near the street can pose a challenge to our approach, as we cannot identify curb points beneath the canopy.

3.4 Curb mapping using random forest

Results for the random forest on the test set are presented in Figure 6. Using a classification threshold of 0.2, we achieved a precision of 0.65, recall of 0.87, and F1-score of 0.75. Our findings suggest that the model can reasonably distinguish between curb and non-curb points. However, due to the small size of the dataset, our model might be over-fitted to our study area or not able to fully generalize our problem.

We applied the trained model over the entire study area. We conducted a visual inspection to find the best classification threshold to be used, being set at 0.4. Visually, random forest prediction is similar to the candidate curb points after the last geometric filter. Since we used a considerably high threshold classification value, we had a lower number of predicted curb points compared to the last geometric filter. We observed predicted curb points also on the highest part of the curb.

Final curb predictions can be seen in Figure 7. Random forest predictions showed a higher number of wrongly mapped curb points, especially for areas with vegetation close to the streets. Moreover, it can be noticed a higher lack of continued along the curbs when compared to the rule-based approach. Nevertheless, the use of the random forest decreases the need for human action during filter tuning.

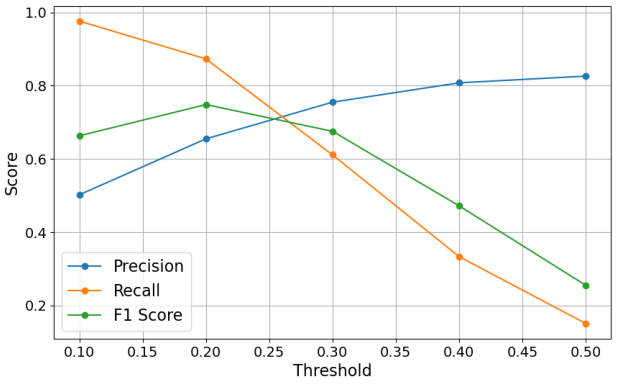

Figure 6. Precision, recall and F1 score curves for the test set. All metrics were calculated using different classification thresholds, ranging from 0.1 to 0.5.

4. Conclusion

We propose an algorithm that integrates image and point cloud processing to automatically map curbs and efficiently thin UAV-SfM point clouds to support hydrological modeling in flood-prone areas. Our findings highlight the advantages of using automatic segmentation to create masks, which significantly accelerates point cloud processing by focusing on areas where curbs are expected, thereby thinning the data and reducing computational load. While we anticipated that masking out vegetation would improve geometric precision, this was not observed in our dataset; however, all point clouds still met the accuracy and precision requirements.

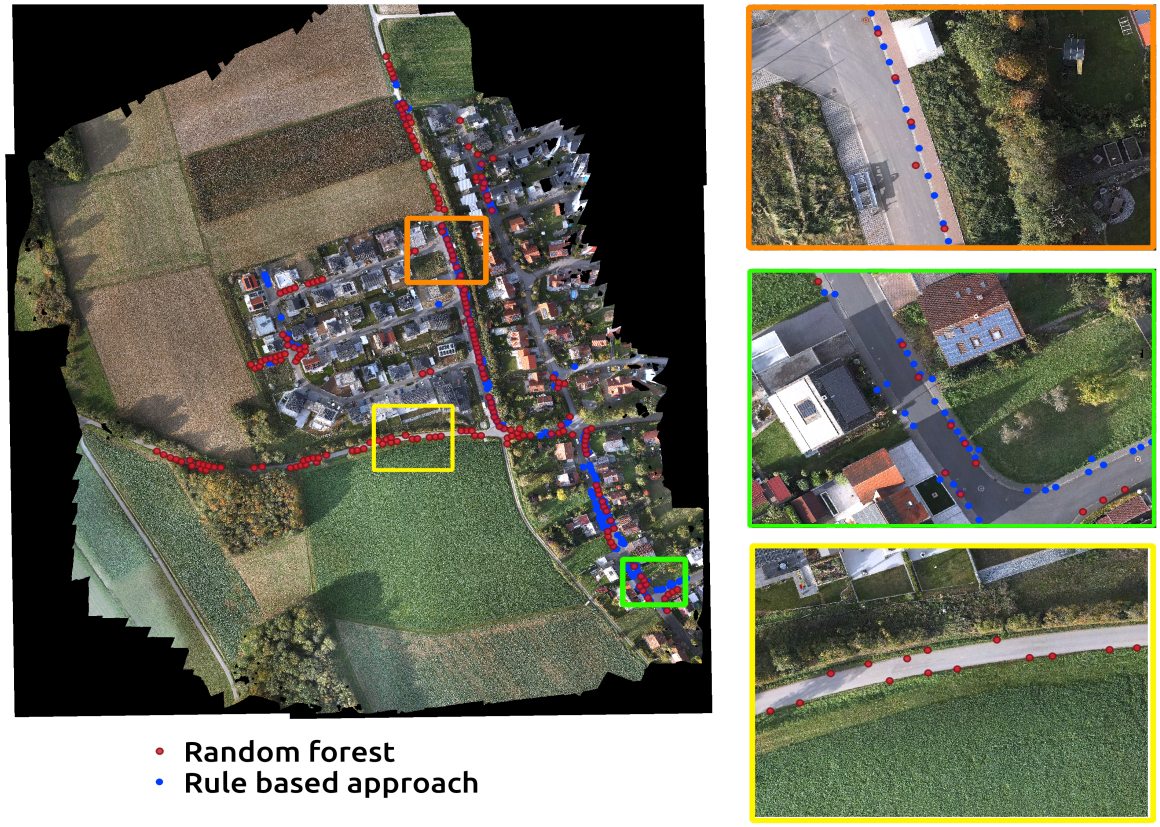

Figure 7. Predicted curb points using the rule-based and random forest approaches. The rule-based approach produced more accurate predictions compared to the random forest approach, further, with higher point density as can be seen in the orange and green boxes. The random forest approach showed a higher number of wrongly assigned curb points (red box).

Our results show that our rule-based approach, which utilizes geometric characteristics of the point cloud, successfully identified potential curb points. The remaining outliers can be further refined by classifying the point cloud or applying skeletonization to the curb points. Notably, our approach was effective in mapping even subtle curbs. Further, random forest showed promising results for our task; however, it presented more wrong predicted curb points than the rule-based approach.

Overall, our approach demonstrates promising potential for efficiently mapping hydraulic structures needed for flood modeling applications. Our work marks an important step toward automating the extraction of key hydraulic features and reducing processing time over large areas, aiding stakeholders in adapting urban landscapes to manage intense rainfall events and enhance urban resilience.

Acknowledgments

This work is part of the project 'Automated fusion and classification of ALS and UAV data for the creation of heavy rain hazard maps' funded by the ZIM programme (Zentrales Innovationsprogramm Mittelstand, grant number KK5549901PR3), Germany's largest innovation programme for small and medium-sized enterprises.

References

Eltner, A., Hoffmeister, D., Karrasch, P., Klingbeil, L., Stöcer, C., Rovere, A., 2022. UAVs in Environmental Sciences. 492. doi.org/10.53186/1028514.

Halama, J., McKane, R., Barnhart, B., Pettus, P., Brookes, A., Djang, K., Phan, V., Chokshi, S., Graham, J., 2023. Improved urban runoff prediction using high-resolution land-use, imperviousness, and stormwater infrastructure data applied to a process-based ecohydrological model. *PLOS Water*, 2, 1-25. doi.org/10.1371/journal.pwat.0000155.

Na Wang, Z. S., Zhang, Z., 2022. Road Boundary, Curb and Surface Extraction from 3D Mobile LiDAR Point Clouds in Urban Environment. *Canadian Journal of Remote Sensing*, 48(4), 504–519. doi.org/10.1080/07038992.2022.2096579.

Park, J., Cho, Y. K., Kim, S., 2022. Deep learning-based UAV image segmentation and inpainting for generating vehicle-free orthomosaic. *International Journal of Applied Earth Observation and Geoinformation*, 115, 103111. doi.org/10.1016/j.jag.2022.103111.

Pashaei, M., Starek, M. J., Berryhill, J., 2023. Application of Semantic Image Segmentation for Efficient UAS-SfM Photogrammetry Mapping. 6983-6986. doi.org/10.1109/IGARSS52108.2023.10282681.

Tan, M., Le, Q., 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. 6105–6114. doi.org/10.48550/arXiv.1905.11946.

Zhang, H., Zhang, H., Wang, C., Xie, J., 2019. Co-Occurrent Features in Semantic Segmentation. 548-557. doi.org/10.1109/CVPR.2019.00064.

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., Yan, G., 2016. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sensing*, 8(6). doi.org/10.3390/rs8060501.

Zhao, G., Ma, F., Qi, W., Liu, Y., Liu, M., 2024. CurbNet: Curb Detection Framework Based on LiDAR Point Cloud Segmentation. doi.org/10.48550/arXiv.2403.16794.

Zou, Z., Lang, H., Lu, J., Ma, Q., 2024. Coarse-to-refined road curb segmentation from MLS point clouds. *Automation in Construction*, 166, 105586. doi.org/10.1016/j.autcon.2024.105586.