sensors

MDPI

Article

Fine-Grained Permeable Surface Mapping through Parallel U-Net

Nathaniel Ogilvie [1], Xiaohan Zhang [1], Cale Kochenour [2] and Safwan Wshah [1,*]

[1] Vermont Artificial Intelligence Laboratory (VaiL), Department of Computer Science, University of Vermont, Burlington, VT 05404, USA; nathaniel.ogilvie@uvm.edu (N.O.); xiaohan.zhang@uvm.edu (X.Z.)
[2] Spatial Analysis Laboratory (SAL), University of Vermont, Burlington, VT 05404, USA; cale.kochenour@uvm.edu
* Correspondence: safwan.wshah@uvm.edu; Tel.: +1-802-656-8086

Abstract: Permeable surface mapping, which mainly is the identification of surface materials that will percolate, is essential for various environmental and civil engineering applications, such as urban planning, stormwater management, and groundwater modeling. Traditionally, this task involves labor-intensive manual classification, but deep learning offers an efficient alternative. Although several studies have tackled aerial image segmentation, the challenges in permeable surface mapping arid environments remain largely unexplored because of the difficulties in distinguishing pixel values of the input data and due to the unbalanced distribution of its classes. To address these issues, this research introduces a novel approach using a parallel U-Net model for the fine-grained semantic segmentation of permeable surfaces. The process involves binary classification to distinguish between entirely and partially permeable surfaces, followed by fine-grained classification into four distinct permeability levels. Results show that this novel method enhances accuracy, particularly when working with small, unbalanced datasets dominated by a single category. Furthermore, the proposed model is capable of generalizing across different geographical domains. Domain adaptation is explored to transfer knowledge from one location to another, addressing the challenges posed by varying environmental characteristics. Experiments demonstrate that the parallel U-Net model outperforms the baseline methods when applied across domains. To support this research and inspire future research, a novel permeable surface dataset is introduced, with pixel-wise fine-grained labeling for five distinct permeable surface classes. In summary, in this work, we offer a novel solution to permeable surface mapping, extend the boundaries of arid environment mapping, introduce a large-scale permeable surface dataset, and explore cross-area applications of the proposed model. The three contributions are enhancing the efficiency and accuracy of permeable surface mapping while progressing in this field.

Keywords: permeable surface mapping; impervious surface mapping; arid environment; aerial imagery; cross-domain adaptation; image segmentation; U-Net

check for updates

Citation: Ogilvie, N.; Zhang, X.; Kochenour, C.; Wshah, S. Fine-Grained Permeable Surface Mapping through Parallel U-Net. *Sensors* **2024**, *24*, 2134. https://doi.org/10.3390/s24072134

Academic Editor: Yangquan Chen

Received: 1 February 2024
Revised: 12 March 2024
Accepted: 18 March 2024
Published: 27 March 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Permeable surface mapping aims to identify surface materials that percolate water, which is a critical component of various environmental and civil engineering applications, including urban planning [1], stormwater management [2], and groundwater modeling [3]. Traditionally, this task involves manually classifying different levels of permeability pixel-wise, which can be labor-intensive and time-consuming. With the development of deep learning computer vision models, semantic segmentation can be one of the potential solutions to increase the efficiency of permeable surface mapping. There are many different ways to implement semantic segmentation, as can be seen in this review of Deep Learning Techniques applied to semantic segmentation [4].

While many existing papers [5–8] studied aerial image segmentation, permeable surface mapping is rarely explored due to difficulties such as (1) lack of a comprehensive

dataset that is used for training and evaluation, (2) heavily unbalanced classes, and (3) hard-to-distinguish classes, for example, permeable brown dirt and the semi-permeable brown road. Some existing methods addressed the mapping of permeable surfaces by binary semantic segmentation. For example, ref. [9] used the inception architecture to classify permeable and impervious surfaces. An alternative approach of binary classification of permeable and impermeable surfaces involves reclassifying pre-existing maps [10]. However, fine-grained permeable surface mapping is rarely studied.

To tackle the above-mentioned difficulties, in this research, our goal is to predict fine-grained permeable surfaces by applying semantic segmentation models, particularly a parallel U-Net [11] model. Our approach involves a two-step process for analyzing aerial imagery. The first stage subjects the input imagery to permeable and non-completely permeable surfaces by a binary segmentation model. Subsequently, the permeable area of the same imagery is further segmented into four remaining levels of permeability. This division allows the second stage to fine-tune its predictions on the fine-grained permeable areas, which encompasses the various permeability levels. The outputs from both models are then merged to produce the final output image. As evidenced in our extensive experiments, this novel parallel U-Net enhances the accuracy of the mapping results, particularly in scenarios where the available dataset is small and heavily weighted toward a single category.

Moreover, this paper presents a novel permeable surface segmentation dataset. This dataset captures three counties in the US. Thus, it is able to evaluate the cross-area performance of the model. The first portion of the dataset comprises aerial images that cover San Mateo County. The second portion of the dataset is from Alameda and Contra Costa Counties. Pixel-wise fine-grained permeable labeling of five distinct classes was also provided for all of the aerial images collected. To our knowledge, this is the first large-scale comprehensive permeable surface dataset in this field. By introducing this dataset, we expect it to advance research on permeable surfaces and inspire future research in this area.

Another goal of our research is to create a model that is more generalized across different domains. Domain adaptation in the context of deep neural networks is a powerful technique that allows the transfer of knowledge gained from one geographical location, or domain, to another [12,13]. Often, data collected at one location may exhibit differences in terms of terrain, climate, vegetation, and even sensor characteristics compared to another location. As a result, a model trained on data from one area may not perform effectively when applied to an unseen area. To this end, we designed a cross-area experiment to demonstrate the generalization of our proposed model. From the experiments, our parallel U-Net model performs better across domains compared to the baseline U-Net.

Our research contributions can be summarized as three-fold:

1. We tackle the problem of permeable surface mapping by proposing a two-stage U-Net semantic segmentation deep learning model. Our novel model is demonstrated to accurately annotate multi-class permeable surfaces.
2. A novel permeable surface dataset with fine-grained permeable-level labeling is collected in three distinct areas of the United States of America. To the best of our knowledge, it is the first dataset with multi-level permeable surface annotations.
3. To advance the application of our model in a real-world environment, a cross-area evaluation protocol is designed for the novel dataset. We also thoroughly studied and attributed the effect of different training techniques on cross-area performance.

2. Related Work

2.1. U-Net-Based Image Segmentation

U-Net [11] was first proposed for precise biomedical image segmentation. It features an encoder–decoder architecture with skip connections to preserve spatial information. Expanding on the U-Net architectures, Liu et al. proposed a cascading U-Net that was applied for Brain Tumor Segmentation, where the architecture involved passing the output from the first U-Net to the subsequent one [14]. Cao et al. proposed Swin-UNet, a U-Net [11] like a swin-transformer-based model, for medical imaging segmentation [15].

Although these models share a similar architecture, they are being applied to different segmentation tasks. Zhang et al. studied lane mark enhancement by adopting a dual U-Net architecture to predict the binary mask of the degraded lane mark regions [16]. In our research, our approach is based on a parallel U-Net model to specifically identify permeable surfaces from aerial imagery. Most U-Net architectures focus on improving a single U-Net [17–22]. Their research often involves adjustments to the convolution mechanism [18], variations in datasets, or the implementation of distinct sampling and attention methods. For example, Sun et al. improved U-Net by incorporating a residual encoder into the U-Net architecture [17]. Nölke et al. used a U-Net for the binary classification of palm tree locations [19]. Hordiiuk et al. combined Xception [23] and U-Net to evaluate ship locations on the water [20]. Wang et al. proposed to segment forest fire smoke from aerial images with a customized U-Net [22]. Nabiee et al. adopted a hybrid U-Net with multiscale feature fusion and a multi-scale skip connection to map war destruction seen in aerial imagery [21].

In summary, in this work, our problem requires fine-grain evaluation on an unbalanced dataset, and we propose a novel parallel U-Net that can focus solely on the fine-grain permeable surface mapping. To achieve this goal, we collect and label a novel permeable surface dataset. These are addressed by training on a new dataset not available to other models that have two different areas of permeable surface maps and by enhancing aerial segmentation by extracting features from an infrared fourth band. Finally, we implement a parallel U-Net architecture that enables fine-grain segmentation and improved cross-area support.

2.2. Satellite Imagery Analysis

A satellite imagery analysis plays a crucial role in various fields ranging from urban planning to disaster management. Over the past decades, researchers have developed numerous techniques and methodologies for extracting valuable information from satellite imagery. German et al. tackles the algal bloom detection from satellite images by calculating the chlorophyll-a concentration from the temporal satellite image sequence [24]. Rashkovetsky et al. combined multiple U-Nets [11] to detect and segment wildfires from multi-band satellite images [6]. Zhang et al. proposed to identify the location of ground images by comparing the learned features from geo-tagged satellite images [25,26]. Li et al. studied poverty mapping from satellite images by proposing a point-to-region dynamic learning framework that can help people find clean water and sanitation services in low-income countries [27]. Li et al. employed YOLOv3 [28] to monitor the agricultural greenhouse effect from satellite images to improve the effectiveness of the modern agricultural industry [29].

In this paper, we focus on using satellite imagery for the fine-grained permeable surface mapping problem. Different from some of the methods mentioned above, for example, refs. [25,27], which only use color channel information, considering the nature of permeable surface mapping, we also take advantage of the infrared band. Our extensive experiments and ablation studies in Section 4 demonstrate the effectiveness of the infrared band in our proposed model.

2.3. Permeable Surface Segmentation Dataset

Several existing datasets have been proposed for permeable surface mapping. For example, Oliazadeh et al. used precipitation measurements to aid in the identification and management of urban runoff [30]. Pelletier et al. assessed the thickness of permeable layers above the Earth's bedrock, although the maps in this study measure the thickness of soil, intact regolith, and sedimentary deposits [31]. Furthermore, Padmanaban et al. utilized the Land Use and Landcover (LULC) classification system, which consists of eight classes that range from barren land to urban areas to wetlands [32]. Lastly, Tootchi et al. aimed to generate a multi-source groundwater and surface water map by amalgamating multiple datasets [33] such as lakes and surface water bodies [34–36], land usage [37], and Hydrological Patterns [38,39].

Our application cannot take advantage of the datasets mentioned above as we have a specific classification system used within our research, especially for permeable surfaces in urban environments. Therefore, we collected and labeled our own dataset, which offers a more in-depth analysis of surface types in an urban setting compared to the previously examined datasets.

3. Methods

3.1. Background

Consider an input image $I_a \in \mathbb{R}^{C\times H\times W}$, where H and W are the image height and width, respectively, and C is the number of channels in the input image, which is 4 in our case. The permeable surface mapping task aims to find a function f such that it assigns one type of permeable surface to each pixel. This can be formulated as follows:

$$f : I_a \in \mathbb{R}^{C\times H\times W} \rightarrow S \in \mathbb{R}^{P\times H\times W}, \tag{1}$$

where S is denoted as the permeable surface segmentation and P is the number of categories of permeable surfaces. In this paper, P is fixed to 5. In existing methods, to find the mapping function f, a deep neural network f_θ is trained to learn from paired data of I_a and P to proximate f, where θ is the parameters of the deep learning model f_θ. In practice, as we discussed in Section 2.1, many existing methods choose U-Net [11] architecture for f_θ and training with multi-class cross-entropy loss as follows:

$$L_{CE}(S, \hat{S}) = -\frac{1}{N}\sum_{n=1}^{N}\sum_{p=1}^{P} log\frac{e^{S_{n,p}}}{\sum_{i=1}^{P} e^{S_{n,i}}}\hat{S}_{n,p}, \tag{2}$$

where $N = H \times W$ is the spatial dimension of the image, S is the prediction from f_θ, and $\hat{S}$ is the ground truth permeable surface segmentation. However, in our case, there is a heavily unbalanced data distribution, which is the nature of the permeable surface mapping task (as mentioned in Section 4.1). The dominant category can easily overwhelm the training procedure, which causes U-Net to suffer from performance degradation.

3.2. Parallel U-Net

To tackle the above-mentioned challenges in permeable surface mapping, we proposed parallel U-Net, which is an enhanced version of U-Net for heavily unbalanced datasets. The key idea of our parallel U-Net is borrowed from the divide-and-conquer concept, which divides the permeable surface mapping into two sub-problems. The first problem involves the intention to segment the fully permeable surface from the other area. The second problem is to perform fine-grained segmentation on semi-permeable surfaces. The final output is simply merging these two predictions. The overview of the proposed parallel U-Net is shown in Figure 1. Formally, this concept can be written as follows:

$$f_\delta : I_a \in \mathbb{R}^{C\times H\times W} \rightarrow S_\delta \in \mathbb{R}^{P_\delta\times H\times W}, \tag{3}$$

$$f_\gamma : I_a \in \mathbb{R}^{C\times H\times W} \rightarrow S_\gamma \in \mathbb{R}^{P_\gamma\times H\times W}, \tag{4}$$

where f_δ parameterized by δ aims to focus on permeable surface and non-permeable surface segmentation, which takes I_a as input and predicts a binary segmentation map while f_γ with learnable parameter γ predicts the fine-grained permeable surface segmentation for other areas. At the inference stage, the final output S can be easily obtained by merging S_δ and S_γ by

$$S_{i,j} = \begin{cases} S_{\delta_{i,j}} & \text{if} \quad S_{\delta_{i,j}} = 1 \\ S_{\gamma_{i,j}} & \text{Otherwise} \end{cases} \tag{5}$$

where i and j are the pixel x-axis and y-axis coordinates. In this simple manner, we can easily merge the output from two models into one single output without any extra calculation.

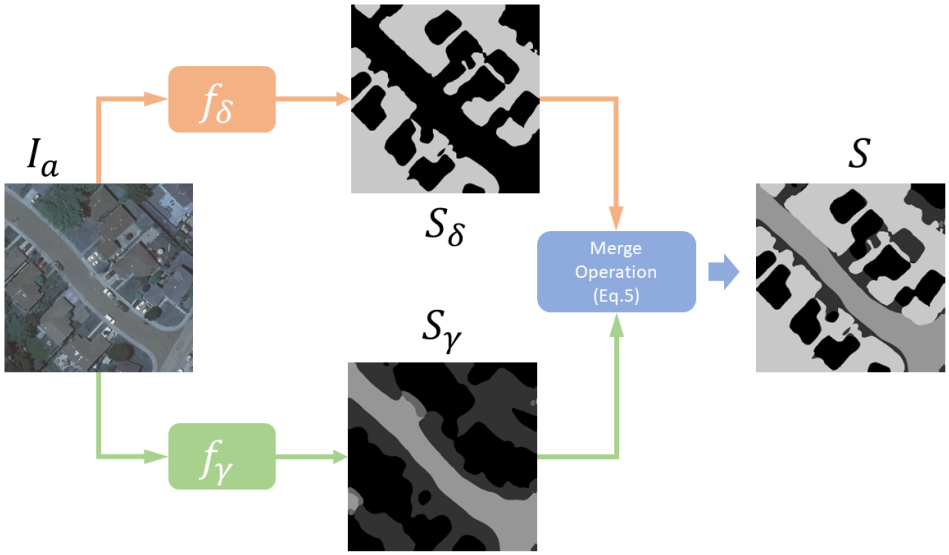


Figure 1. The proposed parallel U-Net architecture. The orange components denote the binary permeable vs. non-permeable segmentation module f_δ and the green components denote the fine-grained permeable surface segmentation module f_γ. The final prediction S is merged from these two outputs.

3.3. Training Objectives

During the training phase, f_δ and f_γ are trained separately for simplicity. Specifically, due to f_δ being only responsible for binary class segmentation, we only adopt binary cross-entropy loss,

$$L_{BCE} = \hat{S}_\delta log(S_\delta) + (1 - \hat{S}_\delta) log(1 - S_\delta). \quad (6)$$

For training f_γ, we adopt the multi-class cross-entropy loss in Equation (2) but ignore the fully permeable surface class as it is handled by f_δ.

4. Experiment

4.1. Dataset

Our dataset contains two main portions. The first portion is collected from San Mateo County, and it is used for training and testing on same-area experiments. The aerial image data contain four channels. The second portion is collected from Alameda and Contra Costa Counties. The aerial imagery contains four channels; the first three channels represent the normal RGB data and the last channel corresponds to the infrared spectrum. For all the aerial images in our dataset, we annotate the permeable surfaces into five different levels, namely, buildings, Other Paved Surfaces, Other Dirt/Gravel Surfaces, paved roads, and all other fully permeable surfaces. The annotation process is in a hybrid style. First, object-based image analysis techniques (OBIA) were used to extract information from land cover using color, tone, texture, pattern, location, size, and shape followed by a classification process to estimate the category of permeable surfaces. Then, a detailed manual review of the dataset was carried out and all observable errors were corrected. Due to the limited capacity of the model and training resources, we tiled the raw large aerial images and the label data into smaller sizes of 512×512. In this manner, we ended up with 2178 images in San Mateo County; specifically, 1465 of them are for training and 713 of them are for testing purposes. Similarly, we ended up with 747 images in Alameda and Contra Costa, which are all for cross-area evaluation. Some randomly sampled processed aerial images as well as their corresponding labels are presented in Figure 2. A summary of the pixel distribution for each class is shown in Table 1 for both San Mateo and Alameda and Contra Costa Counties. This table clearly demonstrates that the data contain a large amount of fully

permeable surfaces but only a minimal amount of dirt/gravel surfaces, which is heavily unbalanced in both San Mateo and Alameda and Contra Costa regions.

Table 1. Statistics of each category of the collected dataset.

Region	Buildings	Other Paved Surfaces	Other Dirt/Gravel Surfaces	Paved Roads	Other Fully Permeable Surfaces
San Mateo	11.36%	8.79%	0.29%	9.24%	70.31%
Alameda and Contra Costa	4.06%	5.48%	1.125%	2.47%	86.86%

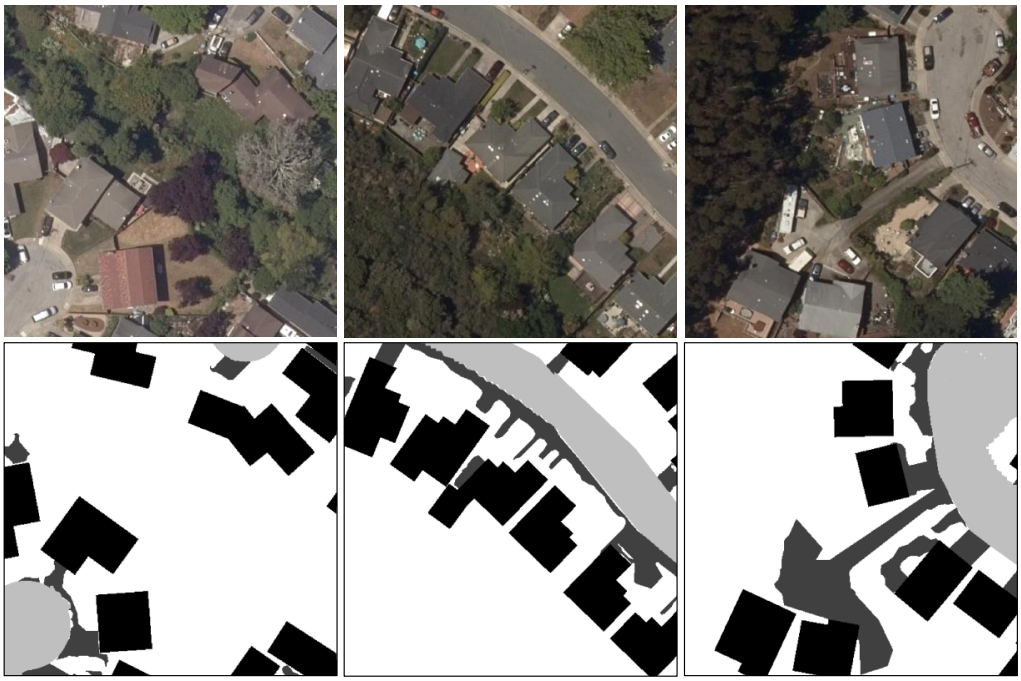


Figure 2. Three randomly sampled data pairs from our permeable surface mapping dataset in the San Mateo region. The images in the first row are aerial imagery in the RGB band. The second row shows the corresponding permeable surface mapping labeling. White means the most permeable surfaces. Black stands for least permeable surfaces.

4.2. Implementation Details and Experiment Setup

We conducted our experiments on the dataset we collected as discussed in Section 4.1. Recall that our dataset has two portions, the San Mateo County and Alameda and Contra Costa Counties. The training set and same-area testing set are from San Mateo County and cross-area testing is from Alameda and Contra Costa Counties. For evaluation, we choose two popular metrics, the F1 score and Intersection-Over-Union (IOU) values. Our model is implemented in Python with the PyTorch 1.9.0 (https://pytorch.org/ (accessed on 5 June 2023)) library and TorchSat 0.0.1 (https://github.com/sshuair/torchsat (accessed on 16 June 2023)) library. We adopted ResNet-152 [40] as the default backbone for our proposed parallel U-Net. Common data augmentations such as random flips and random color jittering are applied during training. The proposed parallel U-Net was trained with the ADAM optimizer [41] with a learning rate of 0.001 and a batch sizer of 32. We trained the model for 100 epochs with a cosine learning rate scheduler. The model was trained on a single Nvidia V100 GPU. The following experiments in this section are performed with the above-mentioned hyper-parameters unless we specify otherwise.

4.3. Main Results

To demonstrate the effectiveness of our proposed parallel U-Net, we benchmark it with two baseline methods, the original U-Net [11], and segmenter [42], which is a transformer-based segmentation module. To enable it, the segmenter was used on our collected dataset with infrared data. We added an extra branch to extract the infrared information before the scale product operation in the segmenter. The result is presented in Table 2. As demonstrated in this table, the segmenter is collapsing due to the heavily unbalanced data distribution. Thus, we did not show the cross-area results of the segmenter model. Another baseline U-Net achieves a reasonable result on both same-area and cross-area experiments. However, we can see that our proposed parallel U-Net exceeds U-Net on both the same-area and cross-area experiments. Specifically, our parallel U-Net significantly outperforms the original U-Net on cross-area benchmarks.

Table 2. Main Results Comparing our Parallel U-Net to Baseline U-Net and Segmenter. The best results are shown in **bold** font.

	Same-Area		Cross-Area	
Method	**Mean F1 Score**	**IOU**	**Mean F1 Score**	**IOU**
Segmenter	0.1597	0.1329	-	-
U-Net	0.6573	0.5670	0.3723	0.2491
Ours	**0.6800**	**0.5799**	**0.4266**	**0.2769**

4.4. Ablation Study

To study the effectiveness of our parallel U-Net model, we conducted several ablation studies to demonstrate its capacity, such as ablation on data augmentation, infrared input, and backbone feature extractors.

4.4.1. Ablation Study on Data Augmentation

Data augmentation plays an important role in semantic augmentation. Similarly, as discussed in Section 4.2, we adopted several data augmentations in training our parallel U-Net. In this section, we conducted experiments to analyze the efficacy of data augmentation methods in permeable surface mapping. First, we compare the results of training with or without data augmentation on both same-area and cross-area benchmarks as shown in Table 3. As indicated in this table, data augmentation improves both same-area and cross-area performance on both the baseline U-Net model and our parallel U-Net model. Noticeably, in the cross-area experiment of our proposed parallel U-Net model, there is a significant improvement from without augmentation to with augmentation. This might be caused by the parallel design that our fine-grained segmentation module f_γ can learn more rich features for augmented semi-permeable surfaces.

Table 3. Ablation results of data augmentation on U-Net and the proposed parallel U-Net. The best results are shown in **bold** font. "AUG" stands for data augmentation.

	Same-Area		Cross-Area	
Method	**Mean F1 Score**	**IOU**	**Mean F1 Score**	**IOU**
U-Net w/o AUG	0.6611	0.5746	0.4080	0.2698
Ours w/o AUG	0.6622	0.5719	0.4107	0.2763
U-Net	0.6573	0.5670	0.3723	0.2491
Ours	**0.6800**	**0.5799**	**0.4266**	**0.2769**

To further study the attributes of each type of data augmentation in the training phase, we conducted another experiment to analyze the contribution of two main data augmentation methods, flipping and color jittering (i.e., randomly adjusting the contrast, brightness, and hue), to the proposed parallel U-Net model. The results are presented in Table 4. As we can see from this table, without any augmentation, the model shows

a poor performance. After adding horizontal flipping and verticle flipping to the model, the same-area performance decreases a little bit but the cross-area increases by a substantial amount. Moreover, we added color jittering to the data augmentation. We found that color jittering helps both same-area and cross-area benchmarks to improve the F1 score and the IOU score. In summary, data augmentation is important for the permeable surface mapping segmentation task. From the experiment, it is recommended to include both flipping and color jittering to achieve the best results.

Table 4. Ablation results of flipping and color jittering data augmentation in training.

	Same-Area		Cross-Area	
Augmentation	**Mean F1 Score**	**IOU**	**Mean F1 Score**	**IOU**
None	0.6658	0.5775	0.3827	0.2324
Flip	0.6622	0.5719	0.4107	0.2763
Flip and Color Jitter	0.6800	0.5799	0.4266	0.2769

4.4.2. Ablation Study on Infrared Band

Different from other satellite image tasks, infrared information is essential in permeable surface mapping. Thus, in this section, we study the importance of infrared information in the proposed parallel U-Net model as shown in Table 5. We experimented on the baseline U-Net and our proposed parallel U-Net. As indicated in the results, infrared band information is critical for both same-area and cross-area benchmarks. Especially in cross-area benchmarks, we can see it improve both baseline U-Net and our proposed parallel U-Net by a large margin. This illustrates that the infrared band contains more invariant information in unseen scenarios.

Table 5. Ablation study on the infrared band in baseline U-Net and our proposed parallel U-Net.

	Same-Area		Cross-Area	
Method	**Mean F1 Score**	**IOU**	**Mean F1 Score**	**IOU**
U-Net w/o Infrared	0.6457	0.5526	0.3566	0.2097
U-Net w/ Infrared	0.6573	0.5670	0.3723	0.2491
Ours w/o Infrared	0.6562	0.5670	0.3844	0.2497
Ours w/ Infrared	0.6800	0.5799	0.4266	0.2769

4.4.3. Ablation Study on Backbone Networks

Finally, we studied the effect of the backbone network on the performance of permeable surface mapping tasks. We varied the backbone feature extractor for the proposed parallel U-Net. We choose four different backbone architectures, namely, ResNet-34, ResNet-50, ResNet-101, and ResNet-152 (default) [40]. The results are presented in Table 6. As we expected, with more convolution layers and the deeper the network, the more accuracy on both same-area and cross-area benchmarks. However, we can also observe that there is a marginal improvement from ResNet-101 to ResNet-152, which suggests that the model almost converges while using ResNet-101 as the backbone. Thus, it is worth considering using ResNet-101 or even using ResNet-50 as the backbone network to balance the accuracy and size of the model in real-world deployment.

Table 6. Ablation study of backbone network on our Parallel U-Net.

	Same-Area		Cross-Area	
Backbone	**Mean F1 Score**	**IOU**	**Mean F1 Score**	**IOU**
ResNet-34	0.6567	0.5679	0.3842	0.2581
ResNet-50	0.6459	0.5542	0.3874	0.2292
ResNet-101	0.6581	0.5635	0.4158	0.2746
ResNet-152	0.6800	0.5799	0.4266	0.2769

4.5. Qualitative Analysis

In conducting a qualitative analysis of the data generated through our implementation of the parallel U-Net, there is evidence that has been observed that parallel U-Net can accurately predict the fine-grained permeable surface mapping under different conditions as shown in Figure 3. The parallel architecture has demonstrated enhanced capabilities in isolating and categorizing permeable features amid semi- or non-permeable surroundings.

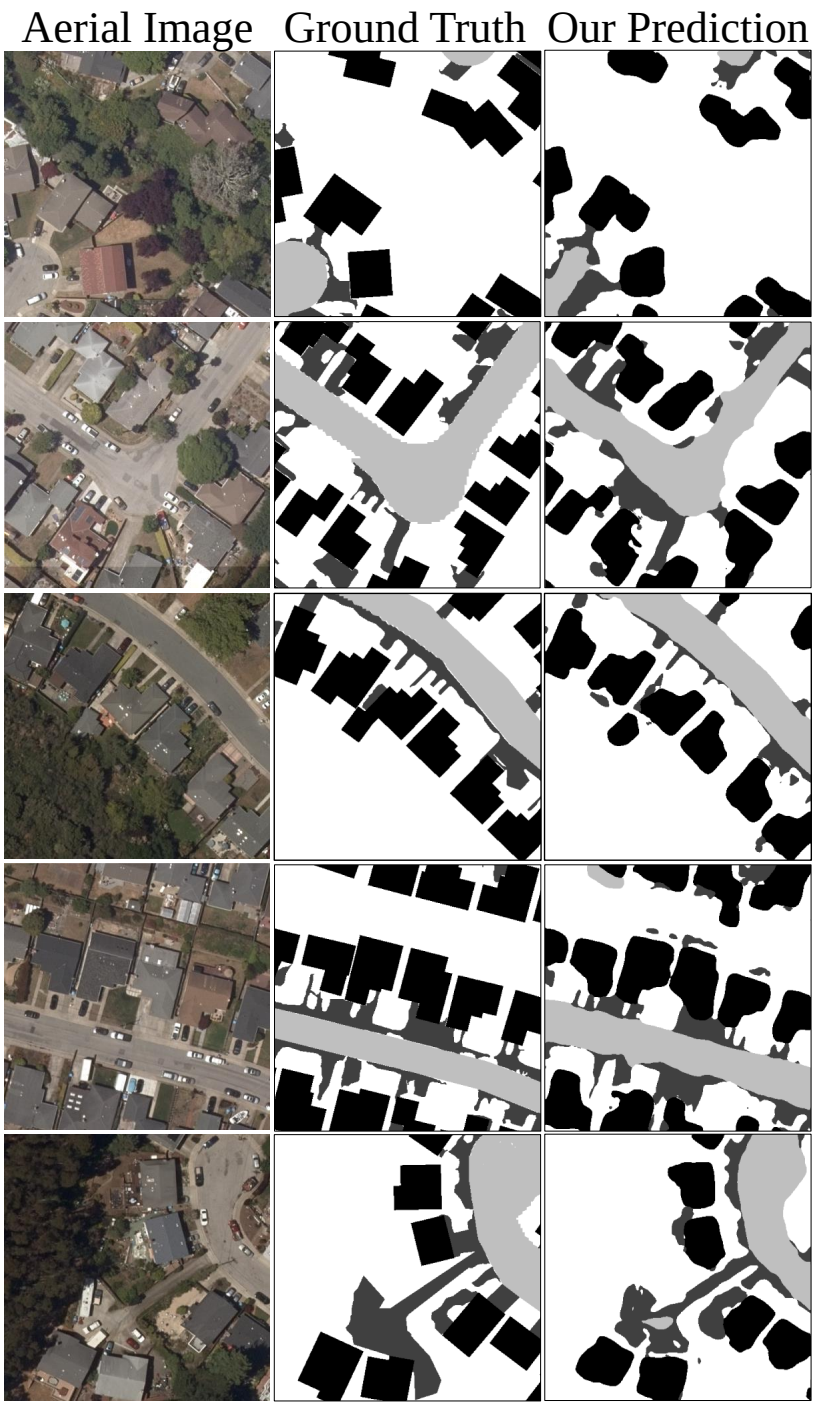


Figure 3. Five random samples' prediction from our proposed parallel U-Net. From left to right is the input aerial image in the RGB band, the ground truth permeable surface mapping, and our predicted permeable surface mapping.

One issue arises in areas where trees overlap roads, presenting challenges in accurately delineating and distinguishing these features. This is highlighted in Figure 4 in the red

box. The intricate nature of such scenarios demands human intervention to ensure precise mapping, as the U-Net output struggles to discern the overlapping elements in the input image. While the parallel U-Net has improved the accuracy of the output when compared to the baseline U-Net as in Section 4.3, the need for manual intervention for the tree–road overlapping issue highlights the importance for future research in this direction and pushes the real-world deployment of fine-grained permeable surface mapping models.

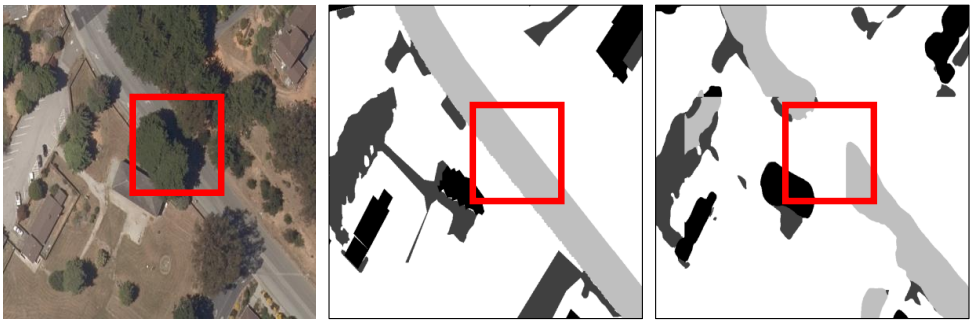


Figure 4. Challenging scenarios when trees and paved road overlap. From left to right is the aerial image, ground truth permeable surface, and our prediction. As shown in the red bounding boxes, this overlapping makes our model fail in predicting a continuous paved road; instead, it predicts fully permeable surfaces since the trees are on top of the road in the aerial image.

4.6. Discussion

The findings presented in Section 4 demonstrate an enhancement in the performance of our parallel U-Net model compared to the baseline U-Net and transformer-based models, particularly in the cross-area evaluation in Alameda and Contra Costa Counties. This improvement demonstrates an outstanding generalization of our proposed parallel U-Net across diverse geographical regions. Specifically, the mean F1 scores and IoU metrics consistently show increased performance for the parallel U-Net in both augmented and non-augmented scenarios, indicating its effectiveness in accurately classifying permeable features in areas with varying characteristics.

In the specific context of Alameda and Contra Costa Counties, the parallel U-Net excels in capturing fine-grained details within the aerial images. This is shown in the higher mean F1 scores and IoU values compared to the baseline U-Net. The improved performance suggests that the parallel architecture can more effectively discern and categorize fine-grained features, providing a detailed and nuanced segmentation of permeable surfaces. This capability is beneficial for applications that demand a higher level of precision, such as urban planning or environmental monitoring in new regions. It is important to acknowledge that while the parallel U-Net creates improvements compared to the baseline U-Net, it may still require manual fixup in certain scenarios such as the tree–road overlapping example shown in Section 4.5. The overall performance boost, especially in cross-area evaluations and fine-grain pixel determination, means that the parallel U-Net is a promising solution for permeable surface applications.

One limitation of the proposed parallel U-Net is the computational efficiency, since the two sub-modules, f_δ and f_γ, are identical. However, f_δ is for a relatively simple task (identifying permeable vs. non-permeable surfaces). Thus, in real-world deployment, it is feasible to prune the weight in f_δ to speed up the inference speed.

5. Conclusions and Future Works

In conclusion, permeable surface mapping stands as an essential component in environmental and civil engineering applications, with implications for urban planning, stormwater management, and groundwater modeling. Permeable surface mapping has traditionally been a labor-intensive manual classification process. The use of deep learning and semantic segmentation offers a new solution to enhance the efficiency of this mapping task. While previous studies have explored various semantic segmentation techniques,

the field of permeable surface mapping in arid environments has remained unexplored. The first complication was the lack of data for training and evaluation; then, within the dataset, there are large class imbalances, and the difficulty of distinguishing between similar surface categories, particularly in arid regions where the color palettes of permeable and non-permeable surfaces often overlap.

In this paper, we introduce a parallel U-Net for fine-grained permeable surface mapping. Our model leverages the successful U-Net model but innovates a novel parallel structure to tackle the heavily unbalanced data distribution in permeable surfaces. This is achieved through the first U-Net deciding between fully permeable surfaces and non-permeable surfaces. Then, the second U-Net categorizes the non-permeable surfaces into various levels of fine-grained permeability. Merging the outputs of these two steps results in higher accuracy and more generalizable maps, particularly in cases where the dataset is limited and dominated by a single category. For future works, considering the different types of terrain and the change in appearance of the land surface at different times of the year, it is feasible to include such data in our proposed dataset. Thus, it can further facilitate future research on permeable surface mapping and attract more researchers to dedicate to this field.

Author Contributions: Conceptualization, N.O.; methodology, N.O. and X.Z.; software, N.O.; validation, N.O.; formal analysis, N.O. and X.Z.; investigation, N.O. and X.Z.; resources, S.W.; data curation, C.K.; writing—original draft preparation, N.O.; writing—review and editing, X.Z. and S.W.; visualization, N.O.; supervision, S.W.; project administration, S.W.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is based upon work supported by the Broad Agency Announcement Program and Cold Regions Research and Engineering Laboratory (CRREL) under Contract No. W913E522C0003.

Data Availability Statement: Data are contained within the article.

Acknowledgments: We extend our sincere gratitude to the Spatial Analysis Lab and its manager, Jarlath O'Neil-Dunne, for generously sharing their data and offering invaluable guidance throughout this project. With heavy hearts, we dedicate this work to the memory of Jarlath, whose spirit guided us even as he passed away during the final stages of this work. His contributions will forever be remembered.

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

References

1. Pafka, E.; Dovey, K. Permeability and interface catchment: measuring and mapping walkable access. *J. Urban. Int. Res. Placemaking Urban Sustain.* **2017**, *10*, 150–162. [CrossRef]
2. Imran, H.; Akib, S.; Karim, M.R. Permeable pavement and stormwater management systems: a review. *Environ. Technol.* **2013**, *34*, 2649–2656. [CrossRef]
3. You, X.; Liu, S.; Dai, C.; Guo, Y.; Zhong, G.; Duan, Y. Contaminant occurrence and migration between high-and low-permeability zones in groundwater systems: A review. *Sci. Total Environ.* **2020**, *743*, 140703. [CrossRef] [PubMed]
4. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Garcia-Rodriguez, J. A review on deep learning techniques applied to semantic segmentation. *arXiv* **2017**, arXiv:1704.06857.
5. Garnot, V.S.F.; Landrieu, L. Panoptic segmentation of satellite image time series with convolutional temporal attention networks. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual, 11–17 October 2021; pp. 4872–4881.
6. Rashkovetsky, D.; Mauracher, F.; Langer, M.; Schmitt, M. Wildfire detection from multisensor satellite imagery using deep semantic segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7001–7016. [CrossRef]
7. Zhang, Z.; Lu, M.; Ji, S.; Yu, H.; Nie, C. Rich CNN features for water-body segmentation from very high resolution aerial and satellite imagery. *Remote Sens.* **2021**, *13*, 1912. [CrossRef]
8. Li, H.; He, Y.; Xu, Q.; Deng, J.; Li, W.; Wei, Y. Detection and segmentation of loess landslides via satellite images: A two-phase framework. *Landslides* **2022**, *19*, 673–686. [CrossRef]
9. Fu, Y.; Liu, K.; Shen, Z.; Deng, J.; Gan, M.; Liu, X.; Lu, D.; Wang, K. Mapping impervious surfaces in town–rural transition belts using China's GF-2 imagery and object-based deep CNNs. *Remote Sens.* **2019**, *11*, 280. [CrossRef]
10. Lee, C.; Kim, K.; Lee, H. GIS based optimal impervious surface map generation using various spatial data for urban nonpoint source management. *J. Environ. Manag.* **2018**, *206*, 587–601. [CrossRef] [PubMed]

11. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
12. Ben-David, S.; Blitzer, J.; Crammer, K.; Pereira, F. Analysis of representations for domain adaptation. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 4–7 December 2006; Volume 19.
13. Ganin, Y.; Lempitsky, V. Unsupervised domain adaptation by backpropagation. In Proceedings of the International Conference on Machine Learning, Lille, France, 7–9 July 2015; pp. 1180–1189.
14. Liu, H.; Shen, X.; Shang, F.; Ge, F.; Wang, F. CU-Net: Cascaded U-Net with loss weighted sampling for brain tumor segmentation. In Proceedings of the Multimodal Brain Image Analysis and Mathematical Foundations of Computational Anatomy: 4th International Workshop, Multimodal Brain Image Analysis 2019, and 7th International Workshop, Mathematical Foundations of Computational Anatomy 2019, Held in Conjunction with International Conference on Medical Image Computing and Computer Assisted Intervention 2019, Shenzhen, China, 17 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 102–111.
15. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-unet: Unet-like pure transformer for medical image segmentation. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 205–218.
16. Zhang, X.; Wshah, S. LanePainter: Lane Marks Enhancement via Generative Adversarial Network. In Proceedings of the 2022 26th International Conference on Pattern Recognition (ICPR), Montreal, QC, Canada, 21–25 August 2022; pp. 3668–3675.
17. Sun, Y.; Bi, F.; Gao, Y.; Chen, L.; Feng, S. A multi-attention UNet for semantic segmentation in remote sensing images. *Symmetry* **2022**, *14*, 906. [CrossRef]
18. Jeppesen, J.H.; Jacobsen, R.H.; Inceoglu, F.; Toftegaard, T.S. A cloud detection algorithm for satellite imagery based on deep learning. *Remote Sens. Environ.* **2019**, *229*, 247–259. [CrossRef]
19. Freudenberg, M.; Nölke, N.; Agostini, A.; Urban, K.; Wörgötter, F.; Kleinn, C. Large scale palm tree detection in high resolution satellite images using U-Net. *Remote Sens.* **2019**, *11*, 312. [CrossRef]
20. Hordiiuk, D.; Oliinyk, I.; Hnatushenko, V.; Maksymov, K. Semantic segmentation for ships detection from satellite imagery. In Proceedings of the 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine, 16–18 April 2019; pp. 454–457.
21. Nabiee, S.; Harding, M.; Hersh, J.; Bagherzadeh, N. Hybrid U-Net: Semantic segmentation of high-resolution satellite images to detect war destruction. *Mach. Learn. Appl.* **2022**, *9*, 100381. [CrossRef]
22. Wang, Z.; Yang, P.; Liang, H.; Zheng, C.; Yin, J.; Tian, Y.; Cui, W. Semantic segmentation and analysis on sensitive parameters of forest fire smoke using smoke-unet and landsat-8 imagery. *Remote Sens.* **2021**, *14*, 45. [CrossRef]
23. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
24. German, A.; Andreo, V.; Tauro, C.; Scavuzzo, C.M.; Ferral, A. A novel method based on time series satellite data analysis to detect algal blooms. *Ecol. Inform.* **2020**, *59*, 101131. [CrossRef]
25. Zhang, X.; Li, X.; Sultani, W.; Zhou, Y.; Wshah, S. Cross-view geo-localization via learning disentangled geometric layout correspondence. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 3480–3488.
26. Zhang, X.; Sultani, W.; Wshah, S. Cross-view image sequence geo-localization. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–7 January 2023; pp. 2914–2923.
27. Li, Z.; Xie, Y.; Jia, X.; Stuart, K.; Delaire, C.; Skakun, S. Point-to-region co-learning for poverty mapping at high resolution using satellite imagery. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 7–14 February 2023; Volume 37, pp. 14321–14328.
28. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
29. Li, M.; Zhang, Z.; Lei, L.; Wang, X.; Guo, X. Agricultural greenhouses detection in high-resolution satellite images based on convolutional neural networks: Comparison of faster R-CNN, YOLO v3 and SSD. *Sensors* **2020**, *20*, 4938. [CrossRef] [PubMed]
30. Oliazadeh, A.; Bozorg-Haddad, O.; Mani, M.; Chu, X. Developing an urban runoff management model by using satellite precipitation datasets to allocate low impact development systems under climate change conditions. *Theor. Appl. Climatol.* **2021**, *146*, 675–687. [CrossRef]
31. Pelletier, J.D.; Broxton, P.D.; Hazenberg, P.; Zeng, X.; Troch, P.A.; Niu, G.Y.; Williams, Z.; Brunke, M.A.; Gochis, D. A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling. *J. Adv. Model. Earth Syst.* **2016**, *8*, 41–65. [CrossRef]
32. Padmanaban, R.; Bhowmik, A.K.; Cabral, P. Satellite image fusion to detect changing surface permeability and emerging urban heat islands in a fast-growing city. *PLoS ONE* **2019**, *14*, e0208949. [CrossRef]
33. Tootchi, A.; Jost, A.; Ducharne, A. Multi-source global wetland maps combining surface water imagery and groundwater constraints. *Earth Syst. Sci. Data* **2019**, *11*, 189–220. [CrossRef]
34. Verpoorter, C.; Kutser, T.; Seekell, D.A.; Tranvik, L.J. A global inventory of lakes based on high-resolution satellite imagery. *Geophys. Res. Lett.* **2014**, *41*, 6396–6402. [CrossRef]
35. Fluet-Chouinard, E.; Lehner, B.; Rebelo, L.M.; Papa, F.; Hamilton, S.K. Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data. *Remote Sens. Environ.* **2015**, *158*, 348–361. [CrossRef]

36. Tulbure, M.G.; Broich, M. Spatiotemporal dynamic of surface water bodies using Landsat time-series data from 1999 to 2011. *ISPRS J. Photogramm. Remote Sens.* **2013**, *79*, 44–52. [CrossRef]
37. Bartholome, E.; Belward, A.S. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977. [CrossRef]
38. Lehner, B.; Döll, P. Development and validation of a global database of lakes, reservoirs and wetlands. *J. Hydrol.* **2004**, *296*, 1–22. [CrossRef]
39. Matthews, E.; Fung, I. Methane emission from natural wetlands: Global distribution, area, and environmental characteristics of sources. *Glob. Biogeochem. Cycles* **1987**, *1*, 61–86. [CrossRef]
40. Targ, S.; Almeida, D.; Lyman, K. Resnet in resnet: Generalizing residual architectures. *arXiv* **2016**, arXiv:1603.08029.
41. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
42. Strudel, R.; Garcia, R.; Laptev, I.; Schmid, C. Segmenter: Transformer for Semantic Segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Virtual, 11–17 October 2021; pp. 7262–7272.

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.