2022

A Comparison of Deep Learning Algorithms on Image Data for Detecting Floodwater on Roadways

Sarp Salih

Kuzlu Murat
Old Dominion University, mkuzlu@odu.edu

Zhao Yanxiao

Cetin Mecit
Old Dominion University, mcetin@odu.edu

Original Publication Citation

Sarp, S., Kuzlu, M., Zhao, Y., Cetin, M., & Guler, O. (2022). A comparison of deep learning algorithms on image data for detecting floodwater on roadways. *Computer Science and Information Systems, 19*(1), 397-414. https://doi.org/10.2298/CSIS210313058S

Computer Science and Information Systems 19(1):397–414 https://doi.org/10.2298/CSIS210313058S

A Comparison of Deep Learning Algorithms on Image Data for Detecting Floodwater on Roadways*

Salih Sarp[1], Murat Kuzlu[2], Yanxiao Zhao[1], Mecit Cetin[2], and Ozgur Guler[3]

[1] College of Engineering, Virginia Commonwealth University,
Richmond, USA
{sarps, yzhao7}@vcu.edu
[2] Batten College of Engineering & Technology,
Old Dominion University, Norfolk, VA, USA
{mkuzlu, mcetin}@odu.edu
[3] eKare, Inc., Fairfax, VA, USA
oguler@ekare.ai

Abstract. Object detection and segmentation algorithms evolved significantly in the last decade. Simultaneous object detection and segmentation paved the way for real-time applications such as autonomous driving. Detection and segmentation of (partially) flooded roadways are essential inputs for vehicle routing and traffic management systems. This paper proposes an automatic floodwater detection and segmentation method utilizing the Mask Region-Based Convolutional Neural Networks (Mask-R-CNN) and Generative Adversarial Networks (GAN) algorithms. To train the model, manually labeled images with urban, suburban, and natural settings are used. The performances of the algorithms are assessed in accurately detecting the floodwater captured in images. The results show that the proposed Mask-R-CNN-based floodwater detection and segmentation outperform previous studies, whereas the GAN-based model has a straightforward implementation compared to other models.

Keywords: Floodwater detection; Mask-R-CNN; GAN; object detection and segmentation.

1. Introduction

Monitoring and sensing roadway conditions automatically are critical not only for self-driving vehicles but also for informing the traveling public. Due to various environmental and weather-related factors (e.g., heavy snow, rain, storm surge), roadway conditions can dramatically change, and a given road segment may be inundated or may operate under a reduced capacity. The vulnerability of roadways to floodwater is increasingly becoming a major concern for numerous communities, especially for those living in the coastal regions, since recurrent flooding occurs more frequently due to sea-level rise, storm surge, heavy rain, and tidal flooding [1]. As many societal functions depend on a functioning transportation infrastructure, i.e., routing of emergency vehicles and delivery of goods and services to support commerce, there needs to be a scalable and effective system to monitor or predict the inundations

* This is an extended version of a conference paper, INISTA 2020.

on roadways [2]. The main goal of this paper is to contribute to the development of such a system by showing how image-based sensing and detection techniques could be utilized to detect floodwater present on the roadways.

Over the last five years, object detection and segmentation have been evolving rapidly, where new approaches are being invented, and new application areas are emerging. Previous studies for the detection and segmentation of floodwater focused on remote sensing methods, which leverage aerial photographs and radar data [3-4]. A survey of large areas can be done with these methods, whereas detailed information about an area, especially a roadway, needs a more localized approach. Synthetic Aperture Radar (SAR) images are used by Kang et al. [5] with Fully Convolutional Network (FCN) implementation. These approaches lack important local information, such as the type, severity, and extent of floodwater, which are crucial for safe driving. Ultrasonic rangefinder and passive infrared temperature sensors are utilized by Mousa et al. [6] to detect floods in urban cities with Artificial Neural Network (ANN). Their implementation requires special sensor placement, which will increase the cost and effort to detect and segment the floodwater. Authors in [7] propose a flood detection model where Region-Based Convolutional Neural Networks (R-CNN) architecture is used to preprocess the images.

In this paper, the recently proposed Mask-R-CNN and Generative Adversarial Networks (GAN) [8] methods will be utilized to detect and segment the floodwater on roadways. Mask-R-CNN is built on the Faster-R-CNN algorithm and improves the segmentation performance [9]. Mask-R-CNN is a deep learning algorithm belonging to the Region-Based Convolutional Neural Networks (R-CNN) family of object detection and semantic segmentation models. As the latest evolution in the R-CNN family, Mask-R-CNN fuses localization, classification, and segmentation in a compact and fast algorithm. GAN algorithm is another deep learning algorithm that is widely used in a variety of applications. It is used in a conditional setting to segment flooded areas in the images. The implementation of Mask-R-CNN and GAN models on detecting and segmenting the floodwater on roadways has not been investigated previously. The main contributions of this study include: (i) Applying Mask-R-CNN and GAN models to both detect whether a given image contains floodwater and, if so, to segment the image to delineate the floodwater; and (ii) Demonstrating that the proposed approaches yield more accurate results than alternative approaches. The proposed deep learning models do not require manual extractions of any features and are tested on various images collected under different conditions.

The remainder of this paper is organized as follows. Section 2 introduces the background of proposed algorithms Mask-R-CNN and GAN. Section 3 presents the methodology. Section 4 includes data collection steps and the preparation of the dataset for processing. Results and related discussions are examined in Section 5. Conclusions and future work are provided in Section 6.

2. Background

Image segmentation is one of the essential processes in the image processing field [10]. Image segmentation tasks include the division of the image into various regions by utilizing similar and specific properties in the image [11]. Deep learning techniques

contribute significantly to image processing, especially in image segmentation. In this section, R-CNN and GAN methods will be discussed in detail.

2.1. Regions with CNN features (R-CNN)

One of the common model families for object detection and segmentation is R-CNN [12]. Four different methods have evolved over the years, namely R-CNN, Fast-R-CNN, Faster-R-CNN, and lastly, Mask-R-CNN.

R-CNN, Regions with CNN features, was developed with a three-step structure in 2014 [13]. The first part of this structure is the selective search, which proposes bounding boxes of around 2000 regions called 'Region of Interest' (RoI) [13]. These proposed regions are fed to CNN to compute the features of each proposed RoI. This step requires very high computing power and is time-consuming. Class-specific linear Support Vector Machines (SVM) are employed for the classification of each region.

The Fast-R-CNN method, Fast Region-based Convolutional Neural Network, is an improved version of R-CNN, which requires a long time for object detection. The calculation of the ConvNet forward-pass independently for each RoI is the main reason why R-CNN is slow [14]. The Fast-R-CNN takes an image and produces a convolutional (Conv) feature map from the entire image, which is shared for each RoI. On the other hand, R-CNN calculates the feature map for each RoI, which takes a lot of time. Fast-R-CNN then extracts the feature vector for the corresponding RoI, which is then fed to a fully connected (FC) layer. FC layer has two outputs: SoftMax probability output for classification and a real-valued output for the position of the bounding box of classes.

The Faster-R-CNN is developed to speed up Fast-R-CNN. The significant drawback of Fast-R-CNN is the separate generation of region proposals. Girshick [14] indicates that the generation of region proposals can be implemented by Conv feature maps used in Fast-R-CNN. Region Proposal Networks (RPNs) are developed as an attention mechanism by Ren et al. [15] and fused with Fast-R-CNN, which harnesses shared computation, to speed up the object detection. The dual structure of Faster-R-CNN decreased the time needed for object detection.

2.2. Generative Adversarial Networks (GAN)

GAN algorithm is developed by Goodfellow [8] that consists of two hostile networks, namely Generator (G) and Discriminator (D) networks. These two networks are simultaneously trained to pick the underlying statistical data distribution of the training set.

Different GAN models are proposed after its first introduction for various image-related applications. One of the most used GAN models, DCGAN, is proposed by Radford et al. [16]. DCGAN utilizes the feature extraction capabilities of CNNs to generate natural images. The elimination of the fully connected layer, use of batch normalization, replacing pooling layers with strided and fractional-strided convolutions improve the produced image quality.

The use of GAN in a conditional setting is implemented by forcing the generated images to follow certain data distribution. Conditional Generative Adversarial Networks (cGAN) models have been used to successfully tackle a wide variety of problems [17]. Future state prediction [18], image manipulation by user constraints [19], super-resolution [20], inpainting [21], and style transfer [22] are some of the application areas that generate significant results.

3. Methodology

This study employs Mask-R-CNN and cGAN based architectures to segment the inundated areas on roadways. These two deep learning techniques will be implemented separately using the same dataset. Our study will reveal that both models have a high potential to achieve successful segmentation results.

3.1. Mask-R-CNN

Mask-R-CNN is the fourth generation of the R-CNN object detection family by adding a third branch that implements fine pixel-to-pixel alignment for segmentation. The Faster-R-CNN framework is modified to predict an object mask [23]. The RoI pooling layer is changed with the Region of Interest Align (RoIAlign) layer to align the extracted features precisely with the input [9]. A fully convolutional neural network is applied to each RoI to produce a segmentation mask. Mask and class predictions are also decoupled to improve instance segmentation [24]. As shown in Fig. 1, localization, classification, and segmentation are the three tasks that Mask-R-CNN implements.

The Mask-R-CNN has a two-stage procedure that consists of a region proposal network and three networks fed by RoI Align, as seen in Fig. 1. The proposed Mask-R-CNN model in this paper is built utilizing a Feature Pyramid Network (FPN) [25] for feature extraction and adopted ResNet101 [26] as the backbone in the first stage. This CNN has four layers of convolution and one deconvolution layer. The RoIAlign layer keeps the size of the feature map the same and avoids misalignment by avoiding quantization. Anchors are used for region proposals. They are preassigned, and our model follows the same steps as in [9]. The model checks if there is any object inside the anchor, then it moves to each pixel in the feature map. After refining the anchor coordinates, it returns the bounding boxes as object proposals. Multi-task loss (L), bounding-box loss (L_{box}), classification loss (L_{cls}), and mask loss (L_{mask}) are identical to the previous studies [9-15]. The multi-task loss is defined as $L = L_{box} + L_{cls} + L_{mask}$ [27]. The Tensorboard visualization tool is used to track these loss metrics. The features are fed into three networks. The mask prediction branch has four convolutions and one deconvolution layer followed by the ReLU activation function to predict flood masks. The returned predicted mask has a size of 28x28 to decrease the computational complexity, which provides faster instance segmentation. On the other hand, image segmentation on downscaled images (28x28) could result in artifacts after being scaled up for inference. There are two common FC layers on the classification and bounding-box branch. This assists the classification network to classify only the inside of the bounding box. The final FC layers consist of 2 layers.

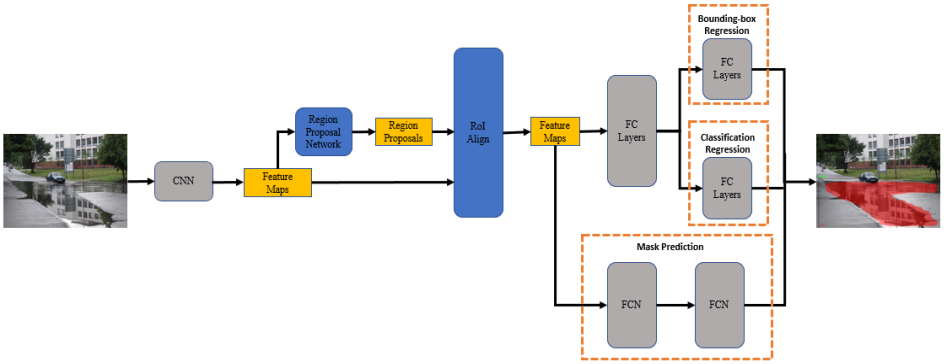


Fig. 1. Framework of the proposed method

3.2. Conditional Generative Adversarial Networks (cGAN)

Conditioning the output to a given input is achieved by cGAN, which evolves from the vanilla GAN model. In the vanilla GAN model, generator and discriminator networks form the GAN architecture. These two separate networks are trained simultaneously to compete with each other. The generator network uses the given random uniform or Gaussian noise (z) to mimic the training data distribution for generating the fake image (y), G: z $\rightarrow$ y [28]. The generated images by the generator network are then classified as real or fake by the discriminator network during the training stage, i.e., D: y $\rightarrow$[0, 1]. This two-player zero-sum game continues until satisfactory image generation is achieved [29]. Training of GAN algorithms is subject to research, and it is a continuing effort to handle efficient training of GAN models [30]. After the training, the generator network is used for the generation of new images.

The cGAN architecture utilizes an additional conditioning input image (x) which provides further benefits in addition to the vanilla GAN architecture. cGAN model outshines other GAN-based architectures because of its simplicity, fast and straightforward implementation. Generator network gets input image (x) and noise (z) to produce an image (y) that shares similar characteristics with the input image and similar data distribution with the training set, G: {x, z}$\rightarrow$ y [17]. Discriminator network also observes the input image (x) so as to determine if the output of the generator image (y) is real or fake, D: {x, y}$\rightarrow$ [0,1] [31]. Generator network design is based on U-Net [32] encoder-decoder with skip connections architecture. Discriminator network utilizes PatchGAN architecture to increase the generated image quality [17].

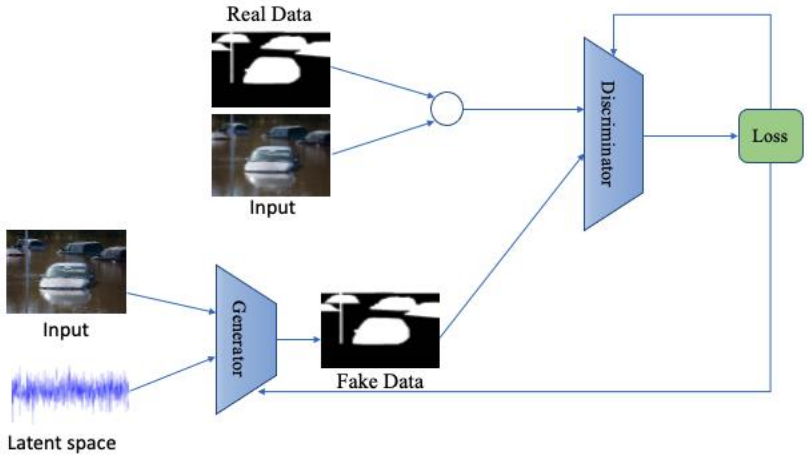


Fig. 2. Framework of the proposed method

The GAN algorithm utilizes four different loss functions during the training to generate decent fake images. D_real and D_fake loss functions for real and fake samples directly update the discriminator network. Generator network weights are updated with two loss functions: adversarial loss (G_GAN) from the discriminator network and L1 loss (G_L1). Training and update of the discriminator network are straightforward, whereas generator network training is done indirectly via the discriminator network. This is one of the reasons why the training of GAN is difficult [33]. Adversarial loss objective is depicted as:

$$\mathrm{G\_GAN(G, D)} = \mathbb{E}_{x,y}[\log D(x,y)] + \mathbb{E}_{x,z}\left[\log\left(1 - D\big(x, G(x,z)\big)\right)\right] \quad (1)$$

The overall objective loss function is expressed with regularizing parameter (λ) as:

$$G^* = argmin_G max_D G\_GAN(G,D) + \lambda G\_L1(G) \quad (2)$$

4. Data Collection

The data collection, preparation, validation, and training environment details will be presented in this section.

4.1. Data Collection and Preparation

The flood image dataset, which includes urban, suburban, and natural areas, is manually labeled for training and testing [34]. The variety of scenes improves the applicability of the algorithm implemented in this paper. The size of the images and labels is shaped as 385x512 pixels for the Mask-R-CNN-based model and the cGAN based model. Flood

images are hand-labeled at the pixel level where flooded areas have the pixel value of one, and the rest of the image has the pixel value of zero. MATLAB labeler tool is used for the labeling. Around 500 images contain floodwater, while another 250 images are dry. Twenty-five (25) crowdsourced images from the internet are used for testing. Some of the images used in this study can be seen below in the result section. Red-colored areas are detected and segmented as floodwater.

4.2. Environment

We implemented the proposed Mask-R-CNN model using the Keras deep learning and Tensorflow libraries with Python 3.6. We used the pre-trained weights for MS COCO and Matterport's code [35] as a starting point. Our model is trained using an Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40 GHz with 128 GB memory and NVIDIA Quadro K4200 with 4 GB dedicated and 64 GB shared memory. We trained our model for 20 epochs which took approximately 24 hours. We use a learning rate of 0.001 where the original paper used 0.02 because it causes weights to blow up in Tensorflow. The rest of the hyperparameters were the same as those in the original Mask-R-CNN article [14].

cGAN based model is implemented using PyTorch machine learning library with Python 3.8. Pix2pix framework [18] is utilized to create the proposed cGAN-based segmentation model. The study with the cGAN based model is realized using Intel(R) Core(TM) i7-10750H CPU @ 2.60 GHz with 32 GB memory and NVIDIA GeForce GTX 1650 Ti GPU with 4 GB dedicated and 16 GB shared memory. We trained our model for 200 epochs which took approximately 2 hours. We choose a batch size of 20 to fully leverage GPU power. That's why the training of the GAN algorithm took less time. Adam optimizer and a learning rate of 0.002 are employed during the training.

4.3. Validation

The performances of the algorithms are evaluated based on four common metrics: accuracy, precision, recall, and F1-score [36]. Since a binary classification problem is being solved, each pixel will fall into one of the two possible categories: "dry" indicating no floodwater and "flood" for the opposite. The performance metrics are defined below :

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1score = 2x\frac{Precision\ x\ Recall}{Precision + Recall} \tag{6}$$

Where True Positive (TP) is the number of correctly predicted flood pixels; False Positive (FP) is the number of pixels that are incorrectly classified as dry; True Negative (TN) is the number of correctly predicted flood pixels; False Negative (FN) is the number of pixels that are incorrectly classified as dry.

5. Result and Discussion

This section will present and discuss the outputs of the Mask-R-CNN and cGAN based models in terms of performance. The dataset consisting of 441 flood and 238 dry images is divided into training (75%) and validation (25%) subsets.

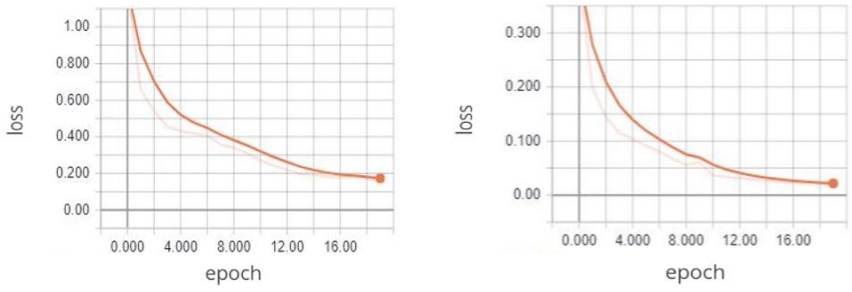
5.1. Results of Mask-R-CNN based model

The proposed Mask-R-CNN-based model encapsulates three different networks. The three networks address predicting the bounding box of the RoI, the classification, and the instance segmentation. All these three networks are trained together during training.

Accuracy is the measurement of correctly classified instances, which is depicted in equation (3). According to the results from the testing samples, the floodwater/dry classification accuracy of the Mask-R-CNN model is 99.2%, and the segmentation accuracy of the flooded area is 93.0%. Compared to cGAN based model and a recent study [2], Mask-R-CNN yields better accuracy in segmentation and classification tasks.

The loss function is used to evaluate how well the machine learning algorithms perform in the training phase. The model weights are regulated to minimize the loss function. Figs. 3 and 4 show the loss graphs of the total, bounding box, classification, and segmentation mask loss. Each figure shows the number of epochs on the X-axis and the output of the loss function on the Y-axis.

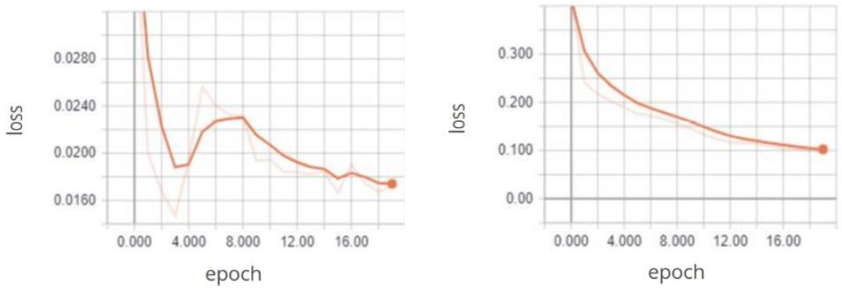
Furthermore, the bold line shows a smoothed version of the actual output shown in a dimmer line for better readability. The overall downward trend of the curves shows the adequate nature of the network modeling flood detection, which is a sign for the model learning the inherent pattern of flooding in the images.



a) b)

Fig. 3. (a) Multi-task and (b) bounding box loss graphs

The loss graphs indicate that the model successfully learns the properties of each image. Fig. 3 and Fig. 4 show the same decreasing nature.



a) b)

Fig. 4. (a) Classification and (b) segmentation mask loss graphs

The validation loss is another metric like training loss, but it is not used to update the weights. It is used for monitoring if the model generalizes well. Checking validation loss helps to avoid overfitting. If the validation loss does not improve for a long time, model training will need to be stopped early. The total validation loss is shown in Fig. 5. for 30 epochs. After 20 epochs, the loss tends to increase, indicating the potential for overfitting.

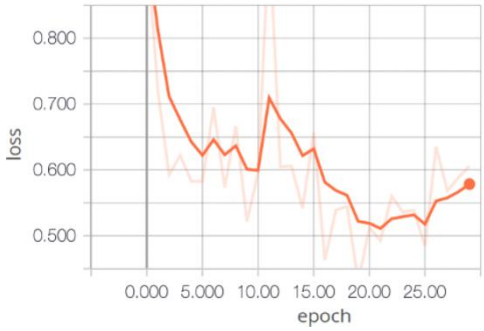
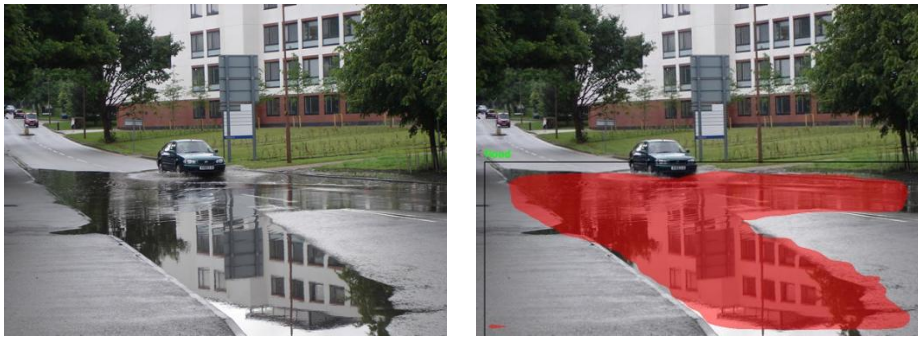


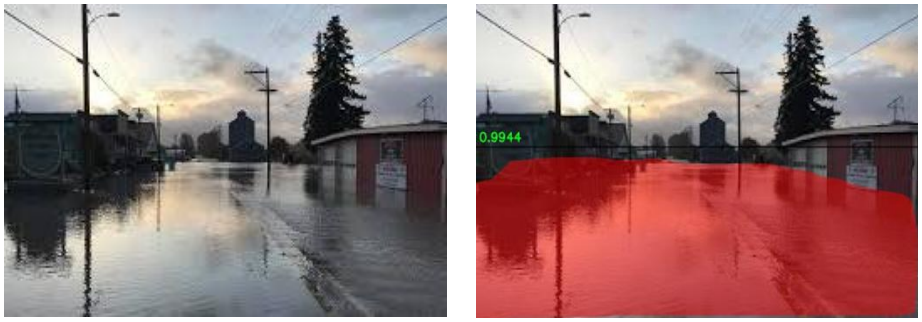
Fig. 5. Total validation loss graph for 30 epochs

Fig. 6 to Fig. 10 depict flood segmentation and original images side by side. The predicted region, as floodwater by the Mask-R-CNN, is marked in red. Since the model first determines a bounding box and then segments within this box, there may be straight lines demarcating the floodwater boundaries in the scene. Overall, the model captures the floodwater boundaries reasonably well.



(a) (b)

Fig. 6. (a) Original and (b) segmented image of a suburban scene. Segmented flood regions are overlaid red



(a) (b)

Fig. 7. (a) Original and (b) segmented image of a suburban scene. Segmented flood regions are overlaid red

(a) (b)

Fig. 8. (a) Original and (b) segmented image of a roadway with a red overlay

(a) (b)

Fig. 9. (a) Original and (b) segmented image of an urban scene with a red overlay

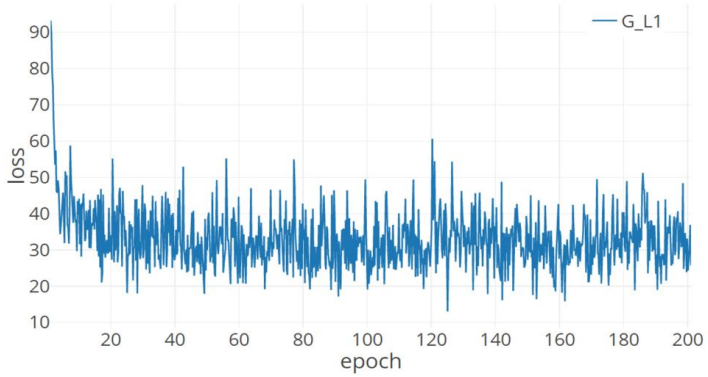
(a) (b)

Fig. 10. (a) Original and (b) segmented image of a natural scene with a red overlay

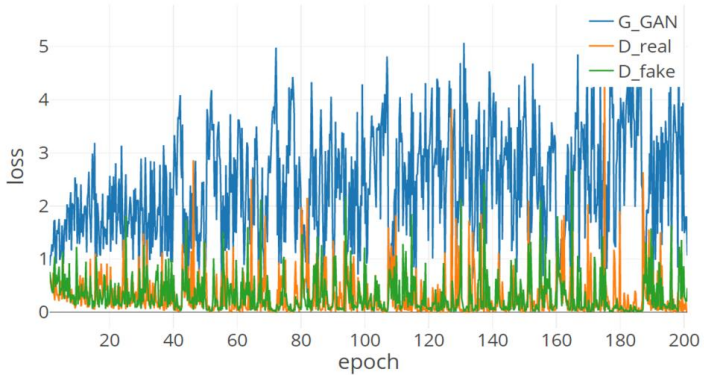
5.2. Results of cGAN-based model

The cGAN-based model has distinctive training characteristics as a different framework (PyTorch) was used in comparison to the Mask-R-CNN-based model (Tensorflow). There are four different loss functions that are used in the GAN model, D_real, D_fake, G_GAN, and G_L1. D_real and D_fake loss functions for real and fake samples directly update the discriminator network. Generator network weights are also updated with two loss functions, adversarial loss (G_GAN) from discriminator network and L1 loss (G_L1) [37]. The training of discriminator and generator networks is a zero-sum game and causes a non-converging problem [38]. Therefore, the training progress cannot be understood by the loss graph alone. The loss graphs of cGAN based model extracted from the visdom virtualization package are indicated in Fig. 11.

The segmentation accuracy of the cGAN based model is 67%. The performance of the cGAN model is low in contrast to previous studies and the proposed Mask-R-CNN model regarding the accuracy parameter. This is an expected result as GAN models need more images to converge. Sample images from the test set are shown in Figs. 12-16.

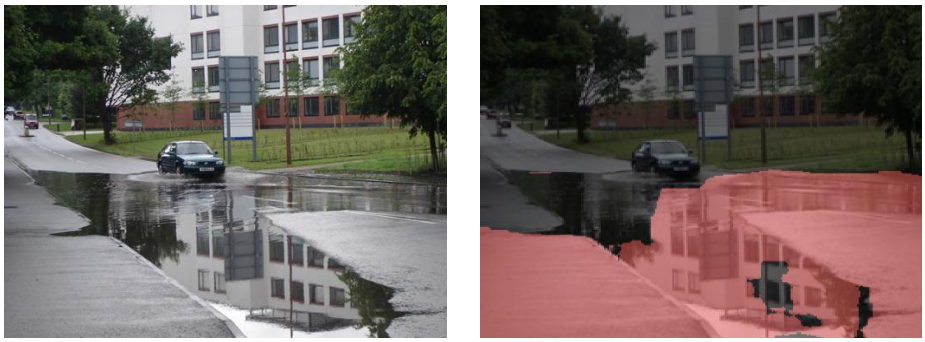


(a)



(b)

Fig. 11. Graph of L1 loss (a) and other three different loss functions (b)



(a) (b)

Fig. 12. (a) Original and (b) segmented image of a suburban scene. Segmented flood regions are overlaid red

(a) (b)

Fig. 13. (a) Original and (b) segmented image of a suburban scene. Segmented flood regions are overlaid red

(a) (b)

Fig. 14. (a) Original and (b) segmented image of a roadway with a red overlay.

(a) (b)

Fig. 15. (a) Original and (b) segmented image of an urban scene with a red overlay.

(a) (b)

Fig. 16. (a) Original and (b) segmented image of a natural scene with a red overlay.

5.3. Comparison of the models

Table 1 compares a recent study [39] with the proposed Mask-R-CNN and GAN models. Our Mask-R-CNN model has better precision, recall, and F1-score results over the Fully Convolutional Neural Network (FCN) based on a pre-trained VGG-16 network [2].

Table 1. Model comparison with a previous study

Model	Precision	Recall	F1-score
Mask-R-CNN	0.96	0.96	0.96
GAN	0.88	0.73	0.80
FCN [2]	0.92	0.90	0.91

Results indicate that Mask-R-CNN achieves better accuracy over cGAN based model and previously proposed models [2]. Object reflections have adverse effects on segmentation accuracy. Training our model on a larger dataset with more reflection cases will increase the model's accuracy. The downscaling of images, as explained previously, to reduce computational complexity may also decrease segmentation accuracy. Instance segmentation on downscaled images provides faster results, which is essential for autonomous vehicles.

The Mask-R-CNN performs segmentation on a rescaled (smaller-28x28) version of the original image (385x512). This has the side effect that the segmentation pixel resolution is very coarse. Figs. 8 and 9 show this artifact very well. The headlights of the car are labeled as a flood in Fig. 8. This could be improved by modifying the Mask-R-CNN to work on a similar size to the original image. However, the downside of increased image resolution is that learning will only converge with a much bigger image dataset because the number of parameters would increase exponentially.

cGAN-based model suffers from a class imbalance problem as it could segment the flooded areas in the lower half of the image better. When the same image in Fig. 9 and

Fig. 15 are compared, cGAN based model achieves better segmentation for the flooded areas around the car. We envision that the cGAN-based model has significant room for improvement if a sufficient dataset is provided.

6. Conclusion

This paper presents state-of-the-art deep learning algorithms, Mask-R-CNN and cGAN, to the segmentation of inundated sections of the roadways, on an image dataset consisting of urban, suburban, and natural scenes. The models are trained to detect floodwater in these images and delineate its boundaries using segmentation techniques. The proposed Mask-R-CNN model achieves 99% and 93% accuracy for floodwater detection and segmentation tasks, respectively. The graphs of the loss functions of the three encapsulated networks, which are classification, bounding box detection, and segmentation mask networks, indicate the efficient training of the model.

The proposed cGAN based model underperforms in comparison to Mask-R-CNN and previous FCN based models. The main reason for the low performance of the cGAN based model is the limited dataset size, which was insufficient to gather data distribution of the training set by the GAN algorithm. However, its straightforward implementation, fast, and apparent structure provide significant potential. With the increasing importance of data collection, the underlying dataset size problem could be resolved. Mask-R-CNN also has many benefits, but it is expensive in terms of time and memory consumption.

As a future study, the floodwater depth prediction will be investigated. The flood level will be determined using nearby objects, such as cars and traffic light poles. In addition to flood depth prediction, the speed of flood detection algorithms will be investigated as well. The dataset size will be increased to get a better result. Finally, water reflections will be studied to improve flood segmentation accuracy.

References

1. W. V. Sweet and J. Park, “From the extreme to the mean: Acceleration and tipping points of coastal inundation from sea-level rise,” Earth's Future, vol. 2, no. 12, pp. 579–600, Dec. 2014.
2. C. Sazara, M. Cetin, and K. M. Iftekharuddin, "Detecting floodwater on roadways from image data with handcrafted features and deep transfer learning," 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 2019, pp. 804-809.
3. N. M. Robertson and T. Chan, "Aerial image segmentation for flood risk analysis," 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, 2009, pp. 597-600.
4. Klemas V. 2014. Remote sensing of floods and flood-prone areas. J. Coastal Res. 31(4):1005–1013.
5. Kang, Wenchao, Yuming Xiang, Feng Wang, Ling Wan, and Hongjian You. "Flood Detection in Gaofen-3 SAR Images via Fully Convolutional Networks." Sensors 18.9 (2018): 2018. Web.

6. Mousa, M., Xiangliang Zhang, & Claudel. (2016). Flash Flood Detection in Urban Cities Using Ultrasonic and Infrared Sensors. IEEE Sensors Journal, 16(19), 7204-7216.
7. Megan A. Witherow, Cem Sazara, Irina M. Winter-Arboleda, Mohamed I. Elbakary, Mecit Cetin & Khan M. Iftekharuddin (2019) Floodwater detection on roadways from crowdsourced images, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 7:5-6, 529-540, DOI: 10.1080/21681163.2018.1488223
8. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014. Generative adversarial networks. arXiv preprint arXiv:1406.2661.
9. He, Kaiming, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. "Mask R-CNN." IEEE Transactions on Pattern Analysis and Machine Intelligence PP.99 (2018): 1. Web.
10. Manisha, M. and Mithra, K.S., Various Image Segmentation Techniques: A Review.
11. Tongbram, S., Shimray, B.A., Singh, L.S. and Dhanachandra, N., 2021. A novel image segmentation approach using fcm and whale optimization algorithm. Journal of Ambient Intelligence and Humanized Computing, pp.1-15.
12. Weng, L., Object Detection for Dummies Part 3: R-CNN Family. 2017. http://lilianweng.github.io/lil-log/2017/12/31/object-recognition-for-dummies-part-3.html
13. Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." ArXiv.org (2014): ArXiv.org, Oct 22, 2014. Web.
14. Girshick, Ross. "Fast R-CNN." ArXiv.org (2015): 27. Web.
15. Shaoqing Ren, R., Kaiming He, Girshick, and Jian Sun. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." IEEE Transactions on Pattern Analysis and Machine Intelligence 39.6 (2017): 1137-149. Web.
16. Radford, A., Metz, L. and Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.
17. Isola, P., Zhu, J.Y., Zhou, T. and Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).
18. Y. Zhou and T. L. Berg. Learning temporal transformations from time-lapse videos. In ECCV, 2016.
19. J.-Y. Zhu, P. Kra ̈henbu ̈hl, E. Shechtman, and A. A. Efros. Generative visual manipulation on the natural image mani-fold. In ECCV, 2016.
20. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Aitken, A. Te- jani, J. Totz, Z. Wang, and W. Shi. Photo-realistic single im- age super-resolution using a generative adversarial network. In CVPR, 2017.
21. D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros. Context encoders: Feature learning by inpainting. In CVPR, 2016.
22. C. Li and M. Wand. Precomputed real-time texture synthesis with markovian generative adversarial networks. ECCV, 2016.
23. Johnson, J.W., 2018. Adapting mask-rcnn for automatic nucleus segmentation. arXiv preprint arXiv:1805.00500.
24. Cheng, B., Wei, Y., Shi, H., Feris, R., Xiong, J., and Huang, T., 2018. Revisiting rcnn: On awakening the classification power of faster rcnn. In Proceedings of the European conference on computer vision (ECCV) (pp. 453-468).
25. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S., 2017. Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125).
26. He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
27. Zimmermann, R.S. and Siems, J.N., 2019. Faster training of Mask R-CNN by focusing on instance boundaries. Computer Vision and Image Understanding, 188, p.102795

28. Huang, H., Yu, P.S. and Wang, C., 2018. An introduction to image synthesis with generative adversarial nets. arXiv preprint arXiv:1803.04469.
29. Ge, H., Xia, Y., Chen, X., Berry, R., and Wu, Y., 2018. Fictitious gan: Training gans with historical models. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 119-134).
30. Sarp, S., Kuzlu, M., Wilson, E. and Guler, O., 2021. WG2AN: Synthetic wound image generation using generative adversarial network. The Journal of Engineering, 2021(5), pp.286-294.
31. Bakkay, M. C., Rashwan, H. A., Salmane, H., Khoudour, L., Puig D., and Ruichek, Y., "BSCGAN: Deep Background Subtraction with Conditional Generative Adversarial Networks," 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, 2018, pp. 4018-4022.
32. Ronneberger, O., Fischer, P. and Brox, T., 2015, October. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
33. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X., 2016. Improved techniques for training gans. In Advances in neural information processing systems (pp. 2234-2242).
34. Sazara, Cem; Cetin, Mecit; Iftekharuddin, Khan (2019), "Image Dataset for Roadway Flooding," Mendeley Data, v1http://dx.doi.org/10.17632/t395bwcvbw.1
35. W. Abdulla. Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow. 2017. Accessed on: February 20, 2020. [Online]. Available: https://github.com/matterport/Mask_RCNN.
36. 36. Sarp, S., Kuzlu, M., Wilson, E., Cali, U. and Guler, O., 2021. The Enlightening Role of Explainable Artificial Intelligence in Chronic Wound Classification. Electronics, 10(12), p.1406.
37. Sarp, S., Kuzlu, M., Pipattanasomporn, M. and Guler, O., 2021. Simultaneous wound border segmentation and tissue classification using a conditional generative adversarial network. The Journal of Engineering, 2021(3), pp.125-134.
38. Goodfellow, I., 2016. NIPS 2016 tutorial: Generative adversarial networks. arXiv preprint arXiv:1701.00160.
39. Sarp, S., Kuzlu, M., Cetin, M., Sazara, C. and Guler, O., 2020, August. Detecting Floodwater on Roadways from Image Data Using Mask-R-CNN. In 2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)(pp. 1-6). IEEE.

Salih Sarp (Student Member IEEE) received the B.S degree from Dogus University, Istanbul, Turkey, in 2014, and the M.Sc. degree in electrical and computer engineering from George Washington University, Washington DC, USA, in 2018. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Virginia Commonwealth University, USA. His research interests include machine learning, embedded systems and internet of things.

Murat Kuzlu (Senior Member IEEE) joined Old Dominion University (ODU) of Electrical Engineering Technology Department as an Assistant Professor in 2018. He received his B.Sc., M.Sc., and Ph.D. degrees in Electronics and Telecommunications Engineering in 2001, 2004, and 2010, respectively. From 2005 to 2006, he worked as a Global Network Product Support Engineer at the Nortel Networks, Turkey. In 2006, he joined the Energy Institute of TUBITAK-MAM (Scientific and Technological Research Council of Turkey-Marmara Research Center), where he worked as a senior researcher.

Before joining ODU, he worked as a Research Assistant Professor at Virginia Tech's Advanced Research Institute. His research interests include smart grid, demand response, smart metering systems (AMR, AMI, AMM), home and building energy management system, co-simulation, wireless communication and embedded systems.

Yanxiao Zhao (Senior Member, IEEE) received the Ph.D. degree from the Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, VA, USA, in 2012. She is currently an Associate Professor with the Electrical and Computer Engineering Department, Virginia Commonwealth University. Her research interests include, but not limited to: the Internet of Things (IoT), 5/6G communications, cyber security, wireless networks, including cognitive radio networks, vehicular networks, wireless autonomous networks, wireless body area networks, software-defined networks, device-to-device (D2D) communications, wireless energy harvesting, and power management and communications in smart grid. Her research has been supported by NSF, NASA, and Air Force. She has published over 70 papers in prestigious journals and international conferences. She has been actively organizing international conferences by serving as a TPC chair, publicity chair, and TPC member. She was a recipient of the Best Paper Award for three international conferences: WASA2009, ChinaCom2016, and ICMIC2019.

Mecit Cetin earned his M.S. degree in Civil Engineering and Ph.D. degree in Transportation Engineering from Rensselaer Polytechnic Institute (RPI), Troy, NY, in 1999 and 2002, respectively. Dr. Cetin has joined the Civil and Environmental Engineering Department at Old Dominion University (ODU) as an assistant professor in August 2008. Prior to that, he had worked as an assistant professor for four years in the Department of Civil and Environmental Engineering at the University of South Carolina (USC), Columbia, SC. Dr. Cetin conducts research in various areas including mining big transportation data, modeling and simulation of traffic operations, congestion pricing, freight transportation, sustainable transportation, traffic signal control, probe vehicle technologies, and system state estimation in transportation networks. Dr. Cetin is currently directing various research projects as the principle investigator totaling more than $1.5M. Dr. Cetin has been working with traffic data from various sensor types, including loops, radar, vehicle classification sensors, Bluetooth, accelerometers within mobile devices, video cameras, GPS, weigh-in-motion sensors, etc., to predict network conditions both on freeways and arterials. His publications include 41 articles in archived journals and 72 refereed international conference proceedings.

Özgür Güler is an imaging scientist specializing in 3D wound imaging and computer vision. Prior to eKare, he was a researcher at the Sheikh Zayed Institute (SZI) for Pediatric Surgical Innovation in Washington DC, where he developed the segmentation and classification algorithms that laid the groundwork of the eKare inSight system. Dr. Güler received his PhD from the Medical University Innsbruck in Austria with focus on image-guided diagnosis and therapy, MS in Computer Science with focus on image-guided surgery and BS in Computer Science from Leopold-Franzens University Innsbruck.

Received: March 13, 2021; Accepted: August 31, 2021.