UNCONSTRAINED FLOOD EVENT DETECTION USING ADVERSARIAL DATA AUGMENTATION

Samira Pouyanfar[*], *Yudong Tao*[†], *Saad Sadiq*[†], *Haiman Tian*[*], *Yuexuan Tu*[†], *Tianyi Wang*[*], *Shu-Ching Chen*[*], *Mei-Ling Shyu*[†]

[*]School of Computing and Information Sciences, Florida International University, Miami, FL, USA
[†]Department of Electrical and Computer Engineering, University of Miami, Coral Gables, FL, USA

ABSTRACT

Nowadays, the world faces extreme climate changes, resulting in an increase of natural disaster events and their severities. In these conditions, the necessity of disaster information management systems has become more imperative. Specifically, in this paper, the problem of flood event detection from images with real-world conditions is addressed. That is, the images may be taken in several conditions, including day, night, blurry, clear, foggy, rainy, different lighting conditions, etc. All these abnormal scenarios significantly reduce the performance of the learning algorithms. In addition, many existing image classification methods use datasets that usually include high-resolution images without considering real-world noise. In this paper, we propose a new image classification framework based on adversarial data augmentation and deep learning algorithms to address the aforementioned problems. We validate the performance of the flood event detection framework on a real-world noisy visual dataset collected from social networks.

***Index Terms*—** Flood event detection, deep learning, style transfer, generative adversarial networks

1. INTRODUCTION

Natural disasters such as flood, earthquake, and hurricane have caused devastating losses on human lives, environment, and economy. In the last decade, disasters have made more than $10 billion in losses [1]. In particular, flood, one of the most severe natural disasters, has caused tremendous damages on roads, houses, and agriculture, and has been the main topic in numerous studies [2, 3, 4]. Meanwhile, disaster information management systems have grown significantly in recent few years due to the latest advances in data collection, analysis, and visualization [5].

On the other hand, machine learning and deep learning have achieved substantial progresses in image classification. However, there are only very few methods that leverage deep learning for real-world disaster detection and management [6, 7]. This is mainly due to the limited annotated data available in this domain. Existing work usually collects the data from Web/social media and annotates them manually [7]. Nevertheless, the variability of images in such datasets may not be sufficient to create a robust model which can be used in different real-world situations. For example, many flood images in social media were taken during the day, most users posted clear images without significant noise, etc. The model trained on such data cannot easily detect a specific disaster from the real-world noisy images (e.g., blurry or night images). Fig. 1(a) shows several noisy flood images that cannot be detected by a deep learning model trained on normal flood images (Fig. 1(b)).

(a)

(b)

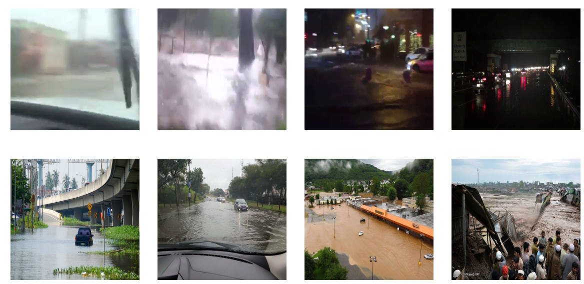

Fig. 1: Samples of (a) noisy and (b) normal flooding images

In this study, a new disaster detection framework based on state-of-the-art deep neural networks is proposed to address the aforementioned challenges. Specifically, we focus on flood detection and collect flood-based images from social media sources such as Twitter and YouTube. We also consider three sets of real-world styles including “night”, “blurry”, and “rainy”. Since collecting and annotating these sets of images are difficult and tedious, recent photorealistic style transfer techniques are utilized to transfer images between two different domains (normal to style) in an unsupervised manner. Specifically, a new data augmentation method based on Cycle-Consistent Generative Adversarial Networks (CycleGANs) is proposed. For each set of styles, a CycleGAN is trained to transfer the images from regular flood to the styled flood (e.g., to night-flood, rainy-flood, and blurry-flood.). These images are later utilized in the data augmentation step to train a Convolutional Neural Network

(CNN). To the best of our knowledge, this is the first work that applies style transfer to flood event detection. In addition, this is the first flood detection framework that can detect unusual flood images without seeing such irregular images in the training set.

The remaining of this paper is organized as follows. Section 2 discusses the related work in flood detection and image style transfer using deep learning. Section 3 presents the proposed framework and the experimental results are given in Section 4. Finally, Section 5 concludes this paper.

2. RELATED WORK

2.1. Flood Detection

Traditional flood detection techniques based on machine learning mainly focus on extracting representative and discriminative features from remote sensing and aerial imagery data to train the learning models [3, 8, 9]. In [3], texture, color, and fractal features extracted from the Unmanned Aerial Vehicle (UAV) surveillance images were used to classify the flood and non-flood geographic areas. In another work, Support Vector Machine with different kernels on TerraSAR-X satellite imagery was proposed [8].

Recently, deep learning techniques such as CNNs have also been utilized in flood detection due to their great performance in image classification and object recognition [4, 10]. In [4], CNNs with dilated convolutions and deconvolution layers were proposed for flood detection using satellite imagery. A multi-modal deep learning model that utilizes the image and textual data from social media for flood detection was developed [10].

2.2. Image Style Transfer using Deep Learning

Recent advances in deep learning enable the transformation of styles from one domain (source) to another domain (target) [11, 12, 13, 14, 15, 16]. Based on this approach, it is possible to generate the minority-class samples by combining the main object in the content images and the artistic style (color, local structures, etc.) from the style images. In [11], separation and recombination of content with neural representations were used to transfer the style. A generalized framework that combines untied weight sharing, discriminative modeling, and a GAN loss was proposed for visual style transformation [12]. Also, an unsupervised pixel-level domain adaptation method without the need of source and target domain pairs was developed to learn the style transformation [13]. The Multi-Style Generative Network (MSGNet) [16] uses a CoMatch Layer approach that learns to match the lower order statistics of content image with the style images. Different from the existing work, we propose a novel framework that applies adversarial style transfer data augmentation to flood event detection.

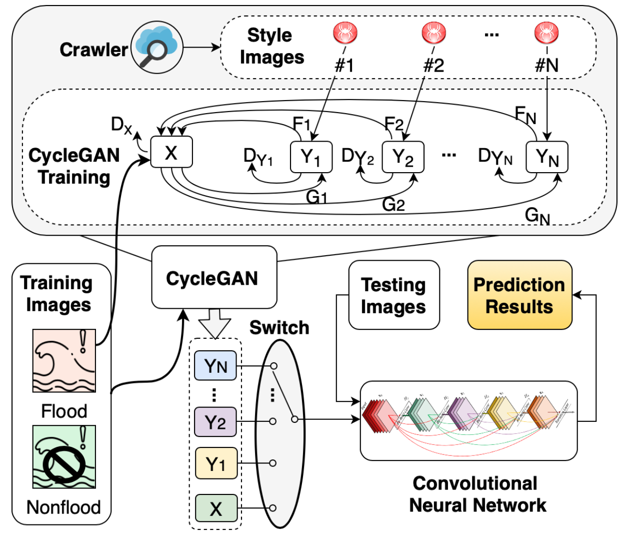

Fig. 2: The proposed deep learning framework

3. THE PROPOSED FRAMEWORK

The proposed framework aims to train a CNN model for image classification which is robust to various contexts (styles), denoted as $Y_1, Y_2, \ldots, Y_N$. For the application of flood event detection, the training images are classified into two categories: flood and non-flood. Therefore, a training dataset $\mathcal{I} = \{I_1^0, I_2^0, \ldots, I_M^0\}$, containing both flood-related and non-flood images, is collected to train the model. Fig. 2 depicts the training process of our proposed deep learning framework in which CycleGAN is utilized as data augmentation to enhance the CNN classifier for flood event detection.

3.1. CycleGAN Data Augmentation

Data augmentation is a common way to enhance the training dataset and improve the performance of the CNN models and its generalization capability. The conventional approaches performing data augmentation include flipping, scaling, cropping, rotation, etc., which manipulate the pixel values in a simple manner. However, the patterns of the images can significantly change in various contexts with a complicated transformation. In the case of flood event detection, the images show different visual characteristics in day or night, in rainy or sunny weather, and when the camera is moving or not. GAN has shown powerful performance to learn the patterns/styles of contexts regardless of the objects in the images, and thus in this paper, we propose a novel approach of data augmentation by utilizing CycleGAN [17] to perform carefully curated style transfer for flood in different contexts. We first define the most common context as the regular context X. All the original images in the training dataset are from X. Then, for each stylized target context Y_i, a CycleGAN model is trained to translate a given flood image from X to

Y_i without any paired image samples. The goal is to learn a set of functions $G_i : X \rightarrow Y_i, \forall i$ so that the learned transformation of images after applying $G_i(X)$ are indistinguishable from the style references Y_i by using an adversarial loss. The adversarial loss (L_{Adv}) is applied to the mapping functions $G_i(X)$ as follows.

$$\begin{aligned}\mathcal{L}_{Adv}(G_i, D_{Y_i}, X, Y_i) = \mathbb{E}_{y_i \sim p(y_i)}[\log D_{Y_i}(y_i)] \\ + \mathbb{E}_{x \sim p(x)}[\log(1 - D_{Y_i}(G_i(x))]\end{aligned} \tag{1}$$

where G_i generates images $G_i(x)$ or $\widehat{y_i}$, and D_{Y_i} discriminates the training sample $G_i(x)$ from the real target y_i. A similar loss is applied to the inverse mapping $F_i : Y_i \rightarrow X$ and its discriminator D_X. Since these adversarial mapping functions are under-determined and prone to overfitting, further reduction of mapping functions is achieved through cycle-consistency, i.e., $x \rightarrow G_i(x) \rightarrow F_i(G_i(x)) \approx x$. Similarly, another inverse cycle-consistency is introduced that learns the transformation back $y_i \rightarrow F_i(y_i) \rightarrow G_i(F_i(y_i)) \approx y_i$. This is achieved by using a cycle consistency loss, defined as

$$\begin{aligned}\mathcal{L}_{cyc}(G_i, F_i) = \mathbb{E}_{x \sim p(x)}[||F_i(G_i(x)) - x||_1] \\ + \mathbb{E}_{y_i \sim p(y_i)}[||G_i(F_i(y_i)) - y_i||_1]\end{aligned} \tag{2}$$

The aforementioned generative model is trained with images of regular floods as well as other stylized contexts. At the end, the reconstructed images $F_i(G_i(x))$ closely match the input images x. These stylized images $G_i(x)$ are then used in the training of the CNN flood detection model.

3.2. Image Classification

Given all the CycleGAN models, each training image I_i^0 can be transferred into N types of contexts. The synthetic images can be represented by $I_i^1, I_i^2, \ldots, I_i^N$, respectively. In each training epoch, the switch randomly selects one of the transferred or original contexts for each image in the training dataset. Then, it feeds the selected images into the CNN model and updates the model parameters accordingly. In other words, in each epoch, a proxy dataset $\mathcal{I}_k = \{I_1^{n_{1,k}}, I_2^{n_{2,k}}, \ldots, I_M^{n_{M,k}}\}$ is generated to train the CNN model, where k is the epoch number and $n_{i,k} \in \{0, 1, 2, \ldots, N\}$ is the selected context of image I_i^0, randomly generated by the uniform distribution.

For the image classification, ResNet50 is applied, where the last layer is replaced by a fully connected layer with sigmoid activation. After the CNN model is trained, the test images are directly used to compute the prediction results, without using any CycleGAN model to transfer the style.

4. EXPERIMENTS AND ANALYSIS

4.1. Experimental Setup

Datasets. We collected flood-related and non-flood images from YouTube and Twitter with the corresponding keywords and tags. First, we used the keyword "Harvey" which was a major hurricane that occurred in the United States in 2017 with a severe inland flooding to search flood-related videos on YouTube. Meanwhile, Twitter is leveraged to collect flood-related visual data via Twitter API [18]. Both images and videos are collected from the tweets with hashtags "flooding" and "flood". The training set included all the data collected from YouTube, however it randomly selected 30% of Twitter's data. The remaining data from Twitter are used as the testing set. All the images are manually labeled as flood and non-flood for training and evaluation purposes, while the flood-related images are tagged as "night", "rainy", and "blurry" for evaluation purposes only, i.e., the proposed framework is blind to these tags. Each image, if applicable, can have more than one tag. Meanwhile, we also collected style images from Google Images to train CycleGAN models with the corresponding keywords. The number of images of each context is shown in Table 1.

Table 1: The size of the collected datasets

		Training	**Test**	**Style**
Non-Flood	Total	1866	1364	-
Flood	Total	8645	5072	21000
	Regular	-	3627	-
	Night	-	294	7000
	Rainy	-	799	7000
	Blurry	-	434	7000

Computing Environment. An NVIDIA Tesla P100 GPU with 16GB of GPU device memory is used to deploy the proposed framework, including ResNet50 and all the CycleGAN models in the experiment.

Hyperparameters. The ResNet50 model [19] pre-trained on ImageNet [20] is used as the image classifier. Adam solver [21] with learning rate=1e-3 and decay=1e-6 is applied to train the image classifier for 100 epochs. We implement CycleGAN to transfer the image style and train the models with Adam solver with learning rate=2e-4 for 150 epochs.

4.2. Experimental Results

Fig. 3 shows several samples generated by Cycle-GAN style transfer model on our dataset. The first row of images includes original images from the collected dataset, while the subsequent rows demonstrate the synthetic images generated from each Cycle-GAN for blurry, rainy, and night contexts, respectively. Although some of the generated images are not realistic (e.g., blurry), it can still help the model to generalize well and detect real-world flood images.

To demonstrate the effectiveness of the proposed framework in transferring the style for each context (night, rainy, and blurry), its performance (Recall) is compared with the original CNN without style transfer (please refer to Table 2). Recall or true positive rate is selected to show the number of

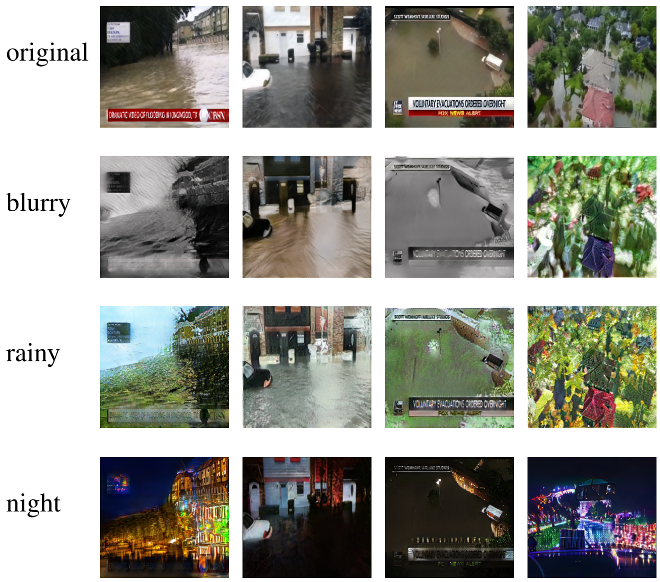

Fig. 3: Cycle-GAN style transfer samples on the flood dataset

Table 2: Recall scores on the flood dataset separated by style

Method	night	rainy	blurry	avg. noisy	flood (total)
CNN	0.785	0.795	0.791	0.790	0.898
Proposed framework	0.831	0.927	0.873	0.877	0.936

correctly classified images for each context. It can be seen from the table that the proposed framework significantly enhances the performance in all categories (the average recall in three categories is increased by more than 8% and the total flood recall reaches 0.94). In other words, the proposed model is able to accurately detect noisy and abnormal flooding samples compared to the conventional CNNs.

Now the question is why Cycle-GAN is utilized in this framework rather than other style transfer techniques. To answer this question, we compare the proposed framework with two other relevant style transferred methods, namely MSGNet [16] and neural style transfer by Gatys [11], both described in Section 2. Table 3 shows the comparison results between these three style transfer techniques and the CNN model. As can be inferred from the table, CNN has the highest precision compared to other techniques, meaning it can detect non-flood images better than other methods. However, its recall value is the lowest among the others, which is already discussed in Section 2. On the contrary, both style transfer methods can achieve very high recall but significantly lower precision. Conclusively, the proposed framework beats all the benchmarks regarding the F1 score (the weighted average of precision and recall) and accuracy.

Finally, Fig. 4 depicts several noisy flood samples that are

Table 3: Comparison results between different style transfer techniques and the baseline

Method	Precision	Recall	F1	Accuracy
CNN	**0.916**	0.898	0.907	0.855
MSGNet [16]	0.855	0.970	0.909	0.847
Style transfer [11]	0.853	**0.978**	0.911	0.849
Proposed framework	0.896	0.936	**0.916**	**0.864**

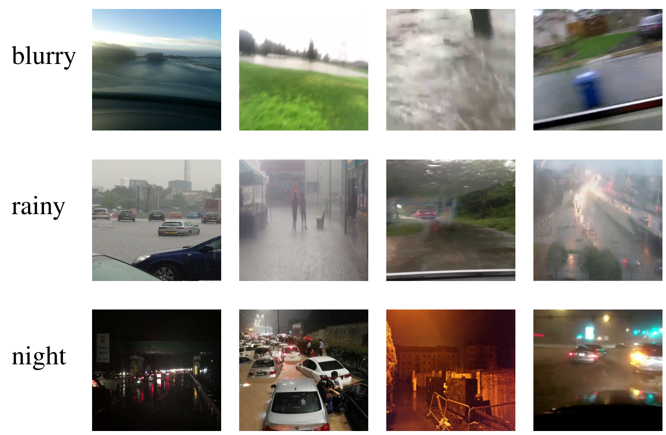

Fig. 4: Correctly classified samples by the proposed framework for each style category on the flood dataset

correctly classified by our framework, whereas the regular CNN cannot detect any of them. These results are evident that the proposed adversarial augmentation model can significantly improve the existing disaster management systems.

5. CONCLUSION

This paper presents a new deep learning framework for real-world flood event detection. CycleGAN is employed as part of the proposed data augmentation to transfer various styles (blurry, night, rainy, etc.) to regular flood images. This technique increases the generalization in deep learning, improves the detection of real-world flood images collected from social media, and reduces the need to have a large-scale annotated dataset. Another important application of this framework is to detect disasters from real-time images collected from public cameras (e.g., network cameras) and assist the community for better decision-making and faster and more reliable emergency responses during a disastrous event.

6. ACKNOWLEDGMENT

This project is supported in part by NSF CNS-1461926 and the Dissertation Year Fellowship award from Florida International University's Graduate School.

7. REFERENCES

[1] Tao Li, Ning Xie, Chunqiu Zeng, Wubai Zhou, Li Zheng, Yexi Jiang, Yimin Yang, Hsin-Yu Ha, Wei Xue, Yue Huang, Shu-Ching Chen, Jainendra Navlakha, and S. S. Iyengar, "Data-driven techniques in disaster information management," *ACM Computing Surveys*, vol. 50, no. 1, pp. 1–45, 2017.

[2] Lamovec Peter, Mikoš Matjaž, and Oštir Krištof, "Detection of flooded areas using machine learning techniques: Case study of the ljubljana moor floods in 2010," *Disaster Advances*, vol. 6, no. 7, pp. 4–11, 2013.

[3] Dan Popescu, Loretta Ichim, and Traian Caramihale, "Flood areas detection based on uav surveillance system," in *IEEE International Conference on System Theory, Control and Computing*, 2015, pp. 753–758.

[4] Keiller Nogueira, Samuel G Fadel, Ícaro C Dourado, Rafael de O Werneck, Javier AV Muñoz, Otávio AB Penatti, Rodrigo T Calumby, Lin Tzy Li, Jefersson A dos Santos, and Ricardo da S Torres, "Exploiting ConvNet diversity for flooding identification," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 9, pp. 1446–1450, 2018.

[5] Maria E Presa Reyes, Samira Pouyanfar, Hector Cen Zheng, Hsin-Yu Ha, and Shu-Ching Chen, "Multimedia data management for disaster situation awareness," in *International Symposium on Sensor Networks, Systems and Security*, 2017, pp. 137–146.

[6] Jinzhuo Wang, Wenmin Wang, and Wen Gao, "Multiscale deep alternative neural network for large-scale video classification," *IEEE Transactions on Multimedia*, vol. 20, no. 10, pp. 2578–2592, 2018.

[7] Samira Pouyanfar, Yudong Tao, Haiman Tian, Shu-Ching Chen, and Mei-Ling Shyu, "Multimodal deep learning based on multiple correspondence analysis for disaster management," *World Wide Web*.

[8] Mahyat Shafapour Tehrany, Biswajeet Pradhan, Shattri Mansor, and Noordin Ahmad, "Flood susceptibility assessment using gis-based support vector machine model with different kernel types," *CATENA*, vol. 125, pp. 91–101, 2015.

[9] Gareth Ireland, Michele Volpi, and George P Petropoulos, "Examining the capability of supervised machine learning classifiers in extracting flooded areas from landsat tm imagery: a case study from a mediterranean flood," *Remote Sensing*, vol. 7, no. 3, pp. 3372–3399, 2015.

[10] Laura Lopez-Fuentes, Joost van de Weijer, Marc Bolanos, and Harald Skinnemoen, "Multi-modal deep learning approach for flood detection," in *Proc. of the MediaEval Workshop*, 2017.

[11] Leon A Gatys, Alexander S Ecker, and Matthias Bethge, "A neural algorithm of artistic style," *CoRR*, vol. abs/1508.06576, 2015.

[12] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell, "Adversarial discriminative domain adaptation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2962–2971.

[13] Konstantinos Bousmalis, Nathan Silberman, David Dohan, Dumitru Erhan, and Dilip Krishnan, "Unsupervised pixel-level domain adaptation with generative adversarial networks," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 95–104.

[14] Judy Hoffman, Eric Tzeng, Taesung Park, Jun-Yan Zhu, Phillip Isola, Kate Saenko, Alexei A Efros, and Trevor Darrell, "CyCADA: Cycle-consistent adversarial domain adaptation," in *International Conference on Machine Learning*, 2018, pp. 1994–2003.

[15] Ashish Shrivastava, Tomas Pfister, Oncel Tuzel, Joshua Susskind, Wenda Wang, and Russell Webb, "Learning from simulated and unsupervised images through adversarial training," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2242–2251.

[16] Hang Zhang and Kristin Dana, "Multi-style generative network for real-time transfer," in *European Conference on Computer Vision*, 2018, pp. 349–365.

[17] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *IEEE International Conference on Computer Vision*, 2017, pp. 2242–2251.

[18] Twitter Inc., "Twitter developer api," https://developer.twitter.com/en/docs.html, 2019.

[19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[20] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[21] Diederik P. Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.