Deep learning models for estimation of flood severity using Satellite and News Article Images

Hariny Ganapathy, Geetika Bandlamudi, Yamini L,
Bhuvana J, T.T. Mirnalinee
SSN College of Engineering, Tamilnadu, India
(hariny,geetika,yamini)@cse.ssn.edu.in,(bhuvanaj,mirnalineett)@ssn.edu.in

ABSTRACT

This paper addresses the Multimedia Satellite Task at MediaEval 2019. We have focussed on the challenge of extracting information present in satellite images. Satellite images provide variety of information like weather, how any event on land unfolds and hence they play an important role in disaster management. We present our approaches for three subtasks: (1) Image-based News Topic Disambiguation, (2) Multimodal Flood Level Estimation from news, (3) Classification of city-centered satellite sequences. All three tasks are related to classifying the images as flood related or not. We have discussed about the performance of proposed CNN models and pre-trained models in the context of binary classification of images for flood related data.

1 INTRODUCTION

Disaster Management is very important in order to help victims survive during natural or man-made disasters. Disasters are unpredictable most of the times. But with the technologies evolving and increased understanding of those technologies, the unpredictability of disasters can be reduced [3]. Flood is one of such major disasters that can be both man-made as well as natural. The general idea behind floods is overflowing of water into land regions. Floods may be caused due to incessant rains or dam breakage and also pose serious damage to humans, infrastructure and wildlife. It is also important to detect floods in areas that are not physically reachable [2]. This paper presents flood detection methods for Multi-modal and Satellite images.

Various classifiers such as vanilla Convolutional Neural Network (CNN), pre-trained models like VGG19 are used to classify the data into flooded or not. The first task is to build a binary classifier that classifies whether an image is related to flooding event or not. The second task involves building a binary classifier that classifies an image into two classes, based on whether the image has at least one person standing in knee-length water. The third task is about utilizing Sentinel-2 satellite images to classify whether the image has the area that is flooded or not [2].

2 RELATED WORK

Remote sensing has been proved to be an effective method to detect and monitor the physical characteristics of an area by measuring its reflected and emitted radiation at a distance from the targeted area, this provides an overview of modeling techniques used for forecasting the vulnerability to flooding of an area [7]. Convolutional neural networks have had shown great performance in various fields such as image classification, pattern recognition etc. Pre-trained networks like DeepSentiBank have also been used to detect floods with the help of social multimedia and satellite imagery and are proven to be giving good results to detect floods [1]. A study on detection of floods was carried out using pre-trained model FCN-16 with 4-fold cross validation and produced highly accurate classification [5]. More research is being carried out on other satellite images like TerraSAR-X images and a rule based classifier on this data produced a considerably low accuracy and Multispectral imagery produced a high accuracy for detection of floods [8]. An ensemble of CNN models trained for retrieving flood relevant tweets had their best model trained only on visual information and the reason stated was that nowadays people are likely to share photographs to address their current situation, rather than detailed textual descriptions [2].

3 APPROACH

In pre-processing, the images are resized without interpolation, to 256*256 to retain all the features when the images are fed to the model. The models are implemented using *Keras* framework and *TensorFlow* as back-end. The development data was divided into 2 sets as training and validation. Based on the validation accuracy, model parameters were altered to achieve better results. The tunnable model parameters used are listed in Table 1. Binary image classifiers were built for each task as explained below.

Table 1: Tunnable Hyper-parameters of the models used

Subtask	Run	No. of Epochs	Learning Rate	Optimizer
INTD	1	100	0.001	Adam [6]
	2	100	0.1	Adam
MFLE	1	30	0.001	Adam
CCSS	1	100	0.1	Adam

3.1 Image-based News Topic Disambiguation (INTD):

The development data had 2673 images with 82% images falling in the non-flooded category and 18% falling in the

MediaEval'19, 27-29 October 2019, Sophia Antipolis, France

flooded category. A validation set was taken to be having 1087 images with 50% of the images in the flooded category and 50% in the non-flooded category.

Run 1: In run 1, the proposed architecture has 6 convolutional layers, 7 activation layers, 5 pooling layers and 1 fully connected and dense layer along with 2 batch normalisation layers. Cross-entropy was the loss function used and the final activation function as softmax with a dropout rate of 0.35. We observed that few of the non-flooded images with lakes and rivers were being classified as flooded. To overcome this we modified our architecture by fine tuning a pre-trained model as Run2.

Run 2: A fine tuned VGG19 model, pre-trained on Imagenet was used, whose top layer is removed in order to customize this classifier for our dataset. Global Average Pooling was used followed by a dropout of 20% [4] which was found to be optimal. An additional fully connected layer is added with *softmax* activation function that can predict the class labels as as flooded or non-flooded.

3.2 Multimodal Flood Level Estimation from News (MFLE):

Run 1: The dataset for the second task included articles and images related to flooding events. In this run, we have analyzed only the image modality. The model was trained using 4770 development set and 950 validation samples. Using only the images, a CNN with 6 *conv* layers and *softmax* activation function was implemented with K fold cross validation with 10 folds. The validation accuracy obtained was 66.67%. For better results, we removed the K fold cross validation and modified the CNN to have *relu* activation function for all the layers and *softmax* in the last layer.

3.3 City-centered satellite sequences (CCSS):

The dataset used for this task includes the satellite images of sentinel-2 band. It contained 12 bands of each image and the pre-processing was done to merge the bands 2,3,4 [10] that represent the blue, green and red bands respectively to form an RGB image [3]. *T*he rasterio tool in python was used to merge the bands of the satellite images as shown in Figure 1. The best prediction value is obtained by sliding over 16 pixels of the merged images. The converted and scaled images were then trained on a VGG19 [9], with dynamic input sizes. Since the first input layer of VGG19 expects an input dimension three, we only pass the RGB information of the satellite data into the network.

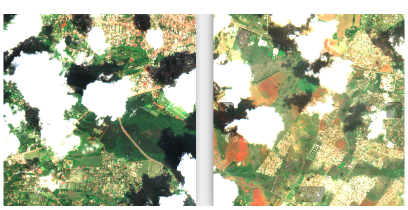

Figure 1: Sample RGB Image after fusion of selective bands

4 RESULTS AND ANALYSIS

Task 1 has 1087 test samples, Task 2 with 1216 and Task 3 has 60 [2]. The test results have been reported as their micro-averaged F1 score for Task1 and Task3 and macro-averaged F1 score for Task2 and have been reported in Table 2. Among two runs submitted for Task 1, the second run produced a better result. Task 2 produced a lower macro-average F1 score because of the model's inability to differentiate between an actual flooded region and a natural water body or a sand/mud coloured area, which led to a lot of non-flooded images to be classified as flooded. There are flooding images but only a subset of them belong to the class that we are interested in to (persons standing in water above knee level) and image-level descriptor/feature might not have been able to learn this. Another reason that contributed to this low score was that only the images were taken into account and not the articles associated with them. Also the model was poor at differentiating the flooded areas with muddy water and deserted region. The following observation was made during the validation phase - about 30% of the 950 sampled validation images were muddy areas/desert regions and less than 20% of those images were correctly classified. For Task 3, extracting the RGB bands from the satellite imagery lead to producing good results. This illustrates the importance of merging the bands of these satellite images.

Table 2: F1 Scores obtained for 3 Multimedia Satellite Tasks

Subtask	Run	Micro Averaged F1 score	Macro Averaged F1 score
INTD	1	0.6312	-
	2	0.8426	-
MFLE	1	-	0.4015
CCSS	1	0.7206	-

5 CONCLUSION

In this paper, we presented our approach for the Multimedia Satellite Task at MediaEval 2019. We have proposed approaches to classify images as flood related or not, for all the 3 subtasks. These images include satellite images as well. As an extension, this work can be applied on active radar data (Synthetic Aperture Radar). We plan to use the results of this work in the future for the monitoring and prediction of flooding events.

6 DISCUSSION AND OUTLOOK

The results of the pre-trained model VGG19 are found to be promising and adding more layers to the proposed CNN model may enhance the performance, since additional layers would extract more features. However the low performance of the model used for the second subtask is owing to the fact that only a part of the given data set i.e. images were used but not the articles. Also, the future work can focus on differentiating muddy/desert regions from actual flooded areas as the error percentage of this was high.

REFERENCES

[1] Benjamin Bischke, Prakriti Bhardwaj, Aman Gautam, Patrick Helber, Damian Borth, and Andreas Dengel. 2017. Detection of Flooding Events in Social Multimedia and Satellite Imagery using Deep Neural Networks. In *MediaEval* (Sept. 13-15, 2017). Dublin, Ireland.

[2] Benjamin Bischke, Patrick Helber, Simon Brugman, Erkan Basar, Zhengyu Zhao, and Martha Larson. The Multimedia Satellite Task at MediaEval 2019 Flood Severity Estimation. In *Proc. of the MediaEval 2019 Workshop* (Oct. 27-29, 2019). Sophia Antipolis, France.

[3] Yu Feng, Sergiy Shebotnov, Claus Brenner, and Monika Sester. 2019. Ensembled convolutional neural network models for retrieving flood relevant tweets. *Image* 10 (2019), 1.

[4] Yarin Gal, Jiri Hron, and Alex Kendall. 2017. Concrete dropout. In *Advances in Neural Information Processing Systems*. 3581–3590.

[5] A Gebrehiwot, L Hashemi-Beni, G Thompson, P Kordjamshidi, and TE Langan. 2019. Deep Convolutional Neural Network for Flood Extent Mapping Using Unmanned Aerial Vehicles Data. *Sensors (Basel, Switzerland)* 19, 7 (2019).

[6] Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

[7] Victor Klemas. 2014. Remote sensing of floods and flood-prone areas: an overview. *Journal of Coastal Research* 31, 4 (2014), 1005–1013.

[8] Biswajeet Pradhan, Mahyat Shafapour Tehrany, and Mustafa Neamah Jebur. 2016. A new semiautomated detection mapping of flood extent from TerraSAR-X satellite image using rule-based classification and taguchi optimization techniques. *IEEE Transactions on Geoscience and Remote Sensing* 54, 7 (2016), 4331–4342.

[9] K. Simonyan and A. Zisserman. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. *CoRR* abs/1409.1556 (2014).

[10] Qunming Wang, Wenzhong Shi, Zhongbin Li, and Peter M Atkinson. 2016. Fusion of Sentinel-2 images. *Remote sensing of environment* 187 (2016), 241–252.