MediaEval2019: Flood Detection in Time Sequence Satellite Images

Pallavi Jain, Bianca Schoen-Phelan, Robert Ross
Technological University Dublin
Dublin, Ireland
{pallavi.jain,bianca.schoenphelan,robert.ross}@tudublin.ie

ABSTRACT

In this work, we present a flood detection technique from time series satellite images for the City-centered satellite sequences (CCSS) task in the MediaEval 2019 competition [1]. This work utilises a three channel feature indexing technique [13] along with a VGG16 pre-trained model for automatic detection of floods. We also compared our result with RGB images and a modified NDWI technique by Mishra et al, 2015 [15]. The result shows that the three channel feature indexing technique performed the best with VGG16 and is a promising approach to detect floods from time series satellite images.

1 INTRODUCTION

Flooding is the most common natural disaster event, which affects people every year all around the world. In most cases, it directly impacts human life and damages properties. In recent years, many techniques have been developed to organise rescue operations in such events in more efficient ways. Flood mapping through satellite images is one such area where a lot of research has been conducted aiming to monitor floods and perform timely risk analysis [2, 3, 5, 18].

Sentinel-2 provides high resolution multi-spectral images, with 13 bands for emergency services, which can also be useful to monitor and analyse the flooding situation. Each of these bands highlights a certain geological features like water, land or clouds. Each band offers a different reflectance and absorbance property which can be exploited for flood detection and monitoring.

Among the 12 bands, visible range bands Red, Green and Blue create a true colour image. These images can map floods and standing water but often suffer from cloud or building shadows which prevents accurate mapping. For that reason several water index techniques have been proposed in order to reduce the effects of shadows and expose appropriate water values. The near infrared (NIR) band highly absorbs water reflectance and reflects vegetation. This property of NIR has made it a popular choice in the past in order to extract water bodies from images. For that reason the normalised difference water index (NDWI) was introduced [14], which leverages NIR and the green band as shown in equation 1. NDWI maximises water features and minimises all other features. Leveraging this particular water indexing technique resulted in the development of many improvements in recent years [6, 20].

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (1)$$

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

NDWI struggles to separate built up areas from water bodies, as NDWI and built-up falls in same range [20] of reflectance values. Considering the built-up area issue and incapability of shallow water detection using NDWI, a combination of two indices has been proposed [15]. The combination of the NDWI water index along with an index using Blue and NIR bands to highlight shallow water along with water bodies. Similarly, Li et al.,2017 [13] proposed a three channel feature index for supervised learning. In this work they leveraged the three indexes being NDVI, NDWI, and RE-NDWI and combined them to create 3 channel images instead of one like RGB [10]. All these indexing techniques are capable of mapping water bodies. Consequently, we assume that these can also be useful in flood water mapping. This could be helpful to rescue teams and provide an improved understanding of disaster situations and areas. As these processes are mostly manual, automating them can be hugely helpful in order to have accurate information in a timely manner.

Lately, Deep Convolutional Neural Networks (CNNs) such as AlexNet [11], VGG16 [17], have performed very well in many domains such as speech recognition, image classification and natural language processing. Remote sensing has also become a widely popular area where deep CNNs have shown good performance [16]. However, in order to train the CNN models with a large number of layers, a significant amount of data is required. This is one of the main challenges in the domain of flood detection. At the same time it has been shown that transfer learning or pre-trained deep CNNs can be a strong option for automating flood detection [8]. Among the deep CNNs, VGG16 has shown great performance previously in many image classification tasks like object detection, image segmentation and scene classification [7].

Flood water is mostly a shallow water body, and difficult to detect due to built-up area or cloud shadows. In this work we propose that if each type of feature such as vegetation, water or clouds are separated efficiently, it can be trained using a pre-trained deep CNN, which is capable to automate the process of flood detection in time series satellite imagery.

2 APPROACH

2.1 Image Processing

2.1.1 Run 1. As shallow water is difficult to map in remote sensing images due to built-up areas, a combination of water index techniques has been proposed in the past [15]. In this approach NDWI is used along with Blue and NIR band indexing as shown in equation 2

$$ModNDWI = \frac{Green - NIR}{Green + NIR} + \frac{Blue - NIR}{Blue + NIR} \quad (2)$$

2.1.2 Run 2. For this run we used true colour images, that is three channel RGB composite images with Red, Green and Blue bands.

2.1.3 Run 3. For this run we leveraged the three-channel index feature space approach [13]. The images are processed to NDVI [eq. 3] that uses NIR and Red bands, NDWI [eq. 1], and Red Edge NDWI (RE-NDWI) [eq. 4], which uses green and red edge (RE) vegetation band. All three of them are then combined horizontally to create a three-channel images like RGB. This approach highlights the individual properties of vegetation, water and clouds.

$$NDVI = \frac{Red - NIR}{Red + NIR} \tag{3}$$

$$RE\_NDWI = \frac{Green - RE}{Green + RE} \tag{4}$$

2.2 Model

The VGG16 network is one of the most popular deep CNN's for image classification and object detection [7, 8, 12]. It consists of 13 convolutional layers and 3 fully-connected layers. We leveraged the pre-trained VGG16 network, which is trained on the ImageNet dataset [4]. Initial layers only extract the general features, and task specific features are extracted by the later layers. We froze the initial 4 blocks and leveraged the last block for our task.

2.3 Experiment

The 12 band data was provided by MediaEval 2019 under subtask City-centered satellite sequences (CCSS) of the multimedia satellite task [1]. It consists of 267 sets of sequences in the development dataset and 68 sets in the test dataset. For the training and testing of the model we split the development dataset into 80% training set, 10% validation set, and another 10% development test set. Data had imbalance class, so we used stratified sampling by class to split the data into train, test, and validation datasets. We also used an image augmentation technique for training datasets by shifting, rotating and flipping the images and achieved the boost of approximately 2-4%.

VGG16 was originally trained to work on 3-channel image data like RGB. However, Mod-NDWI creates a single-channel images like greyscale, which we consequently converted into a 3-channel image assuming identical values for each input channel.

For processing the time series image, we used a pixel based technique. For that, we created the average image of each set of sequence images after individual image processing and fed those to the VGG16 model. The average image modifies only the changed values due to change in image while keeping unchanged values the same. The changed values in average image possibly be influenced by cloud coverage or atmospheric changes. But as each changed value is due to the different features, it might be distinguishable from change due to water values.

These averaged images are then fed to the VGG16 model with frozen 4 blocks and unfrozen last block for our task. The VGG16 network is then followed by a flatten layer, dense layer of 128 unit, and softmax layer. We also used a dropout [19] of 0.5 to avoid over-fitting and the ReLU activation function. The Adam optimiser [9] with learning rate of 5e-6 has been used with binary cross entropy loss function. The model is trained for approx. 30 epochs depending on best performance of each processed images.

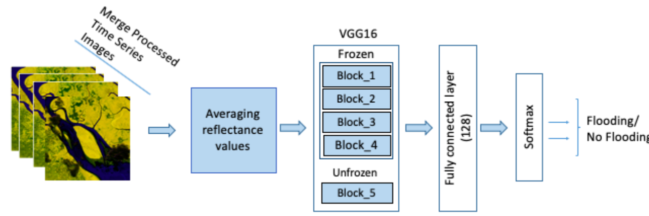

Figure 1: Model Architecture

3 RESULTS

For the evaluation of the model we used micro average F1 Score, as mentioned in the competition evaluation task [1]. Also, image data had imbalance classes, due to which accuracy measure can be misleading, for that reason F1 score is an appropriate evaluation metric as it provides balance score of precision and recall.

The result shown in table 1, which clearly show that the averaging of images can provide good performance in order to detect if a city is flooded. Additionally, the 3-dimensional feature indexing technique outperforms the true colour RGB and Mod-NDWI [15] by approximately 3% in both development and test results.

Table 1: Development and Test Results

Run	Dev F1	Test F1
Run 1	0.963	0.897
Run 2	0.963	0.941
Run 3	**1.00**	**0.970**

4 CONCLUSION

In this work, we explored the automatic detection of floods in an area for sequence of time series images. We used a pixel based averaging approach on RGB, Modified NDWI and a three-channel feature indexing technique along with deep CNNs model VGG16. The results pointed towards significant improvements in flood detection when using a three-channel feature index. Furthermore, it appears that the averaging technique is efficient in detection of flood in the city over the time period.

REFERENCES

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

[2] Miles A Clement, CG Kilsby, and P Moore. 2018. Multi-temporal synthetic aperture radar flood mapping using change detection. *Journal of Flood Risk Management* 11, 2 (2018), 152–168.

[3] Roberto Cossu, Elisabeth Schoepfer, Philippe Bally, and Luigi Fusco. 2009. Near real-time SAR-based processing to support flood monitoring. *Journal of Real-Time Image Processing* 4, 3 (2009), 205–218.

[4] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 248–255.

[5] Dieu Anh Dinh, B Elmahrad, Patrick Leinenkugel, and Alice Newton. 2019. Time series of flood mapping in the Mekong Delta using high resolution satellite images. In *IOP Conference Series: Earth and Environmental Science*, Vol. 266. IOP Publishing, 012011.

[6] Gudina L Feyisa, Henrik Meilby, Rasmus Fensholt, and Simon R Proud. 2014. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment* 140 (2014), 23–35.

[7] Gang Fu, Changjun Liu, Rong Zhou, Tao Sun, and Qijian Zhang. 2017. Classification for high resolution remote sensing imagery using a fully convolutional network. *Remote Sensing* 9, 5 (2017), 498.

[8] Fan Hu, Gui-Song Xia, Jingwen Hu, and Liangpei Zhang. 2015. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing* 7, 11 (2015), 14680–14707.

[9] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[10] Sascha Klemenjak, Björn Waske, Silvia Valero, and Jocelyn Chanussot. 2012. Unsupervised river detection in RapidEye data. In *2012 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 6860–6863.

[11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*. 1097–1105.

[12] Erzhu Li, Junshi Xia, Peijun Du, Cong Lin, and Alim Samat. 2017. Integrating multilayer features of convolutional neural networks for remote sensing scene classification. *IEEE Transactions on Geoscience and Remote Sensing* 55, 10 (2017), 5653–5665.

[13] Na Li, Arnaud Martin, and Rémi Estival. 2017. An automatic water detection approach based on Dempster-Shafer theory for multi-spectral images. In *2017 20th International Conference on Information Fusion (Fusion)*. IEEE, 1–8.

[14] Stuart K McFeeters. 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International journal of remote sensing* 17, 7 (1996), 1425–1432.

[15] Kshitij Mishra and P Prasad. 2015. Automatic extraction of water bodies from Landsat imagery using perceptron model. *Journal of Computational Environmental Sciences* 2015 (2015).

[16] Keiller Nogueira, Waner O Miranda, and Jefersson A Dos Santos. 2015. Improving spatial feature representation from aerial scenes by using convolutional networks. In *2015 28th SIBGRAPI Conference on Graphics, Patterns and Images*. IEEE, 289–296.

[17] Karen Simonyan and Andrew Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015*.

[18] Sergii Skakun, Nataliia Kussul, Andrii Shelestov, and Olga Kussul. 2014. Flood hazard and flood risk assessment using a time series of satellite images: A case study in Namibia. *Risk Analysis* 34, 8 (2014), 1521–1537.

[19] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* 15, 1 (2014), 1929–1958.

[20] Hanqiu Xu. 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing* 27, 14 (2006), 3025–3033.