Flood Impact Level Detection from Multimedia Images using Transfer Learning with CNN Architectures

Thilini Weerasinghe
Department of Computing and Information Systems
Sabaragamuwa University of Sri Lanka
BelihulOya, Sri Lanka
idttweerasinghe@std.appsc.sab.ac.lk

Pubudu Jaysena
Department of Computing and Information Systems
Sabaragamuwa University of Sri Lanka
BelihulOya, Sri Lanka
pubudu.nuwanthiak@gmail.com

***Abstract* — Natural as well as man-made disasters are a global threat to all populations. Some of the natural disasters are floods, cyclones, landslides, earthquakes, and droughts. Terrorism and bomb explosions are some of the disasters caused by man. Among the natural disasters, flooding is a significant contributor to loss of human life, economic loss and property damage disturbances. Floods can occur with little notice or forecast. In this paper, we study deep neural networks with a transfer learning mechanism to identify flood impact levels using multimedia images.**

Keywords* — *deep neural network, transfer learning

I. Introduction

Flooding is a major natural disaster that contributes to disruption of human life, damage to property and the economy of a country. Governments have to pay special attention to manage these emergencies because it can interrupt the development of the country and can affect economic growth. Flooding incidents call for a quick response. Rescue and emergency teams should respond quickly to the locations impacted and get victims to safety in a timely way. Unfortunately, in terms of quick response, rescue teams may be affected by floods also. Automatic identification of flood risk levels helps to promote preparations that will minimize the effects of disasters. Figure 1 shows the global disaster types from 1998 to 2017. Figure 1 highlights that flooding is the most prevalent natural disaster type.

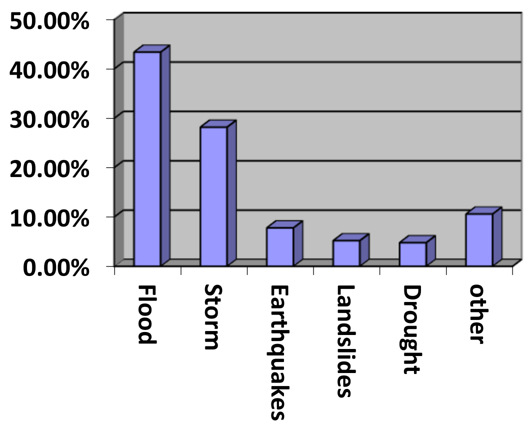

Fig. 1: Percentages of Global occurrences of natural disaster types from 1998-2017

By 2020, the amount of digital data produced will exceed 40 zettabytes for every man on the earth. The majority of these data will be produced by machines as they communicate with each other over the network [1]. With the huge development in information technology day by day, a massive set of data release to the internet daily. These data coming from different fields might contain very important information if we can sort them and understand the hidden relationships in those data. These developments change the activities on multimedia and produce both semi-structured and unstructured data on the internet. Therefore, multimedia data mining has become a very popular research area due to the huge amount of multimedia data on the internet as audio, images, video, graphics, text or sensor data, etc [2].

Big data analysis is an emerging technology that is used to analyze data with huge size which is beyond the ability of commonly used software tools to capture, manage, and process within a tolerable elapsed time[3]. When compared with traditional data sets, big data need instant analysis methods because of their unstructured data. We can get novel opportunities from analyzing big data such as realizing new values[4], to gather detailed knowledge about concealed values, and also acquire knowledge on possible ways to organize and manage multimedia datasets efficiently[5]. Visual data, sensor data, Geographical data, social media data, storing and sharing data, and satellite data are a few types of data that can be used to create a real-time flood emergency management system [6].

This study is based on the following two main research questions;

- What is the deep learning algorithm that can be used to manage flood emergencies?
- How to provide a comparative study to analyze those available approaches?

This paper includes the research objectives, proposed methodology, results, and then discussion and conclusion. In the conclusion section, recommendations for future work are also mentioned.

II. OBJECTIVES

The objectives of this study are to find available deep learning models to identify flood impact levels using images and create a comparative study to identify the most accurate model combination with high performance. The ultimate goal of this study is to support the flood emergency management in the country and contribute to the safe living and economic growth of the country.

ISSN 2756-9160 / November 2020.
International Conference on Advances in Computing and Technology (ICACT–2020) Proceedings

III. METHODOLOGY

The following figure shows the proposed methodology of the research study.

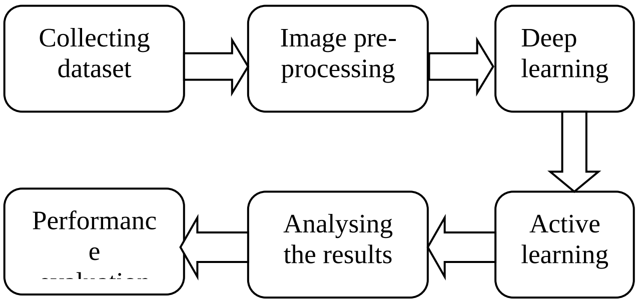

Fig. 2: Proposed methodology

A. Dataset

Dataset is compromised of 1543 social media images and has categorized into three categories. There are 553 “highrisk level” images, 607 “Medium-risk level” images, and 383 as “low-risk” level images. Dataset divides into parts as a training set (1243 images) and a validation set(300 images).

B. Image Pre-processing

Image pre-processing is an application of a wide array of activities ranging from altering images, applying artistic filters, enhancing the comprehensibility of the images, image segmenting, etc. In image pre-processing data augmentation is techniques use to increase the number of image data. Data augmentation is a process of creating new images by performing rotation or reflection of the original image, zooming in and out, and shifting, applying distortion, changing the color palette [7].

C. Deep Learning Models

In the image processing context, many advanced Conventional Neural network architectures have been used to feature extractions.VGG16, VGG19, InceptionV3, Mobilenet, DenseNet 169, DenseNet201, Denset 121 are some of models available to image feature extractions[8]. A Convolutional Neural network is formed with several layers. These layers perform image classification tasks. A CNN architecture generally consists of alternate layers of convolution layer and pooling layer and one or more fully connected layers at the end. For the transfer learning mechanism used in this study, we only choose 3 models, VGG16, VGG19, and MobieNet.

D. Transfer Learning

Transfer learning and domain adaptation refer to the situation where what has been learned in one setting is exploited to improve generalization in another setting. Transfer learning is an optimization that allows rapid progress or improved performance when modeling the second task. Simply, the model trains with a large amount of data and adjusts model weight and bias during training. These learned weights are transferred to solve another task. Another network model that is going to solve this new task starts with these pre-trained weights.

E. Confusion Matrix

The Confusion matrix is used for finding the accuracy and the performances of the model for the classification problem where the output is two or more types of classes.

- True Positive (TP) = Real label and predicted the label of the sample is positive.
- True Negative (TN) = Real label and predicted the label of the sample are negative.
- False Negative (FN) = Real label of the sample is positive, and the predicted label is negative.
- False Positive (FP) = Real label of the sample is negative, and the predicted label is positive.
- Real Positives (RP) = Real positives
- Real Negatives (RN) = Real Negatives
- Predictive Positives (PP) =Predicted Positives
- Predictive Negatives (PN) =Predicted Negatives

F1-score defines true positives are considered as twice important as the other samples.

$$\text{Accuracy} = \frac{TN+TP}{TN+FN+TP+FP} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1Score} = \frac{2TP}{P\ 2TP+FP+FN0} \quad (4)$$

IV. RESULTS AND DISCUSSION

Table 1: Summary result table

	VGG19	VGG16	MobileNet
Accuracy Epoch=15, batch-size=32	52%	57%	89%
Accuracy Epoch=15, batch-size=10	50%	53%	86%
Speed	Low	medium	high
Loss Epoch=15, batch-size=32	26%	22%	5%
Loss Epoch=15, batch-size=10	23%	15%	5%

After analyzing the results and selected the MobileNet model as the high accuracy feature extraction model. In the evaluation phase, we used the F-measure mechanism to validate the performance of the chosen model. In the analyzing part, it shows the **F-measure value** as **0.62450**. Figure 2 shows the confusion matrix for the model.

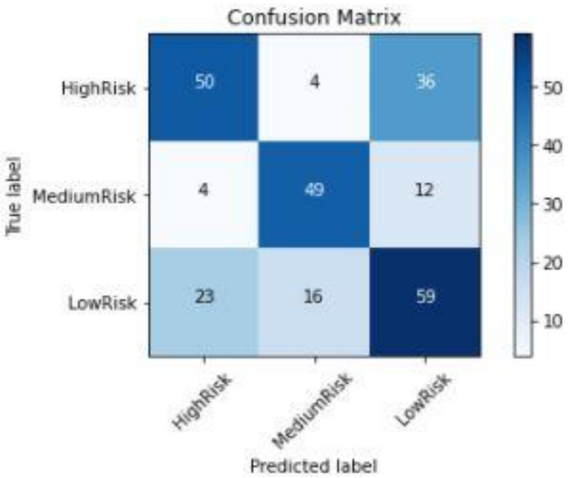

Fig. 3: Confusion Matrix for Mobile Net Model

For the flood impact level prediction, we choose images from social media that are not in the dataset and then go through the trained model and got the results. Here are the results of them.

Figure 4: Used images for prediction

Img 1 Img 2 Img 3

Table 2: Prediction results table

	High-Risk	Medium-Risk	Low-Risk	**Prediction level**
Img 1	0.935892	0.036322	0.027786	**High-Risk**
Img 2	0.067895	0.633542	0.298563	**Medium-Risk**
Img 3	0.006743	0.195617	0.79764	**Low-Risk**

V. CONCLUSION

In this study, we have studied the level based flood image classification and flood area segmentation and also provided a flood image dataset. We have proposed an image classification model according to their impact level. Further, it provides a comparative study between VGG16, VGG19, and MobileNet deep learning models. When comparing the results, MobileNet model has resulted in higher accuracy. MobileNet model predicts the impact level with over 89% accuracy. The other conclusion that can stated after analyzing the results is that the accuracies of the models increase when increasing the batch size of the training phase. As this study mostly used image data, we used the CNN algorithm to get high accuracy results. This automated approach minimizes manual work because it is difficult to define the severity of the flood without prior knowledge and it is a time-consuming process. As future work, we suggest to combine different types of multimedia data such as image, audio, and video data and to create a real-time platform for flood prediction.

REFERENCES

[1] Lucas Mearian, "By 2020, there will be 5,200 GB of data for every person on Earth." [Online]. Available: https://www.computerworld.com/article/2493701/by-2020--there-will-be-5-200-gb-of-data-for-every-person-on-earth.html.

[2] W. Zhu, C. Luo, J. Wang, and S. Li, "Multimedia cloud computing," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 59–69, 2011, doi: 10.1109/MSP.2011.940269.

[3] Z. Xu, Y. Liu, L. Mei, C. Hu, and L. Chen, "Semantic-based representing and organizing surveillance big data using video structural description technology," *Journal of Systems and Software*, vol. 102, pp. 217–225, 2015, doi: 10.1016/j.jss.2014.07.024.

[4] P. S. Mathew and A. S. Pillai, "Big Data solutions in Healthcare: Problems and perspectives," *ICIIECS 2015 - 2015 IEEE International Conference on Innovations in Information, Embedded and Communication Systems*, 2015, doi: 10.1109/ICIIECS.2015.7193211.

[5] S. T. N. kumar Sudeep Tanwar, *Multimedia Big Data Computing for IoT Applications*.

[6] Saparamadu, "E-Government for emergency management : use of multimedia big data platform for flood emergency management in Sri Lanka," *International Journal of Advanced Scientific Research*, vol. 4, no. 1, pp. 15–21, 2019.

[7] F. Agostinelli, M. Hoffman, P. Sadowski, and P. Baldi, "Learning activation functions to improve deep neural networks," *3rd International Conference on Learning Representations, ICLR 2015 - Workshop Track Proceedings*, no. 2013, pp. 1–9, 2015.

[8] D. Dias and U. Dias, "Flood detection from social multimedia and satellite images using ensemble and transfer learning with CNN architectures," *CEUR Workshop Proceedings*, vol. 2283, pp. 3–5, 2018.