Detection of Disaster Affected Regions based on Change Detection using Deep Architecture

Naga Pavan Srivathsav C1, Anitha K2, Anvitha K3, Maneesha B4, Sagar Imambi S5

Abstract: Natural disasters pose a serious threat to national economy, human lives and can disturb the social fabric of the society, although we can not entirely prevent these natural disasters from happening but with the advancements in the satellite imagery, remote sensing and machine learning it has become possible to minimize the damage caused by them. Satellite images are very useful because they can give you a huge amount of information from a single picture. Since it is becoming easy to get these satellite images the climate and environmental detection systems are in high demand. In this paper, we propose a post disaster system which we have named, Automatic Disaster Detection System (ADDS) which is designed to detect the disaster affected areas and help in the relief operations. The existing methods for detection of disaster affected regions are mostly dependent on manpower where people use the drone technology to see which area is affected by flying that drone over a large area which takes a lot of time. A new approach of Convolution Neural Network towards detection of disaster affected areas through their satellite images is examined in this paper which is comparatively better than previous image processing techniques. This method is based on deep learning which has been a widely popular technique for image processing in recent past. This technique can help save lives by reducing the response time and increasing the efficiency of the relief operations.

Index Terms: Convolution Neural Network, Deep Learning, Image Processing, Flood Detection, Satellite Imagery, Remote Sensing, Machine Learning.

I. INTRODUCTION

Flood occurs when the water overflows from the rivers, oceans due to heavy rainfall and it can happen during anytime of the year. Floods cause huge damage to both life an-d property and also make the rescue efforts very difficult. Floods also cause one of the major problems of cutting the transportation which makes the rescue efforts even tougher. During floods usually a very large area is affected and it becomes very difficult for the Relief Forces to identify the areas which are most affected and reach them.

Disaster affected Region detection system has been one of the major areas of research in the recent past since saving human lives is the number one priority after the disaster has occurred. An immediate fast response is required in case of deadly disasters like floods so that human lives can be saved. There have been several systems based on machine learning techniques which have shown good results. Prior systems are mostly based on sensors, and they are not that efficient due to the unavailability of a good number of sensors. The range of the disaster affected region detection system is also limited and this system cannot handle huge amount of satellite imagery data and detect the disaster in a short period of time. Identifying the flood affected areas or regions automatically can lead to great decrease in the response time and loss of life. The present system focuses more on looking drone images which could require a lot of manpower since during floods large amounts of areas are affected. Hence, this paper focuses on creating an Automatic Disaster Detection System which takes the satellite images and detects the occurrence of disaster and shows the area affected by it. Automatic Disaster Detection System can be very useful in the domain of human services where there hasn't been much automation in the recent times. When an area is hit by floods then identification of a particular area where there has been huge damage becomes difficult and leads to increase in response time by the authorities. By using the ADDS we can see what are the areas most affected by it and then prioritize our rescue efforts so that areas which require most immediate attention are rescued first. The ADDS uses the Google Earth images to compare two images, the images after the disaster has struck and the other image before the disaster has struck.

II. RELATED WORKS

The paper "Change Detection using a Local Similarity Measure" [1], presented by M. Jahari, S. Khairunniza Bejo introduced a technique of local mutual information and image thresholding with a limitation mentioning that 2-D entropies can't reflect information of different image accurately and can't acquire optimal threshold in the year 2008. The paper which is presented by W.A.C. Fernando, C.N. Canagarajah "A unified approach to scene change detection in uncompressed and compressed video" [2] in the year 2016 by the algorithm that novel unified algorithm

for scene change detection in uncompressed and MPEG-2 compressed video sequences using statistical features of images and also proposed the future work as can be extended the for camera movements detection within the same framework. The review presented by Shiva Kumar B R in the year 2016 through a paper "Change detection using image differencing" [4] tells that Image differencing is a method of subtracting the DN (Digital Number) value of one data with the other one of the same pixel for the same band which results in new image, having a limitation that no forest approachment has been observed over the study area. The review presented by Suresh Kumar, Margaret Anouncia, Sebastian Johnson in the year 2012 through a paper "Agriculture change detection model using remote sensing images and GIS" [5] along with the method of SVM having a limitation that the results obtained for their re- search was limited to the district of Vellore. The paper "Change detection methodology based on region classification fusion" [6] presented by Ting Liu, George Gigli, and George A. Lampropoulos introduced the system through thresholding, Fuzzy-C-Means (FCM) and decision trees in the year 2014. The paper "An application of image change detection Urbanization" [7] proposed by Jovit Reno. A, Beulah David. D introduced a method that is proposed for finding changes between images of same set occurred at various time intervals may be between years or various dates by telling the limitation that Land changes often due to the seasonal changes of land covers , deforestation natural disasters and many other factors in the year 2015. The paper "Unsupervised Image Change Detection Algorithm Based on 2-D Histogram" [8], presented Wenbang Sun, Haiyan Tang, Hexin Chen, Guang Yu. introduced a 2-D histogram is ascertained by using Fisher Criterion by observing the that the system should have little control over image classes, detailed spectral knowledge of surface may be necessary in the year 2010.

I. PROPOSED METHOD

A. CHANGE DETECTION

Change detection is a process which is used to find out the change in the landscape or characteristics of an area between different time periods. It often involves comparing the satellite images of the area. It has been used to understand the change in the characteristics of the land to determine forestation, growth of population, impact of the floods etc. The task identifying the changes that occurred after the disaster has hit a particular area can be considered a change detection task i.e. the changes that have occurred when compared with the image of the same area at a different timeline.

B. LEARNING ALGORITHM

Usually when we use CNN for image processing tasks such as Object detection, Image Classification, Digit Recognition, Pattern Recognition we use specific kind of images. Training the model from the scratch would require a huge amount of images for thoroughly training

the CNN. Hence we use a very different approach while training our model. Our model is specially designed for using a smaller size data-set when compared to other models and increasing the accuracy of the model. We cut a single image into 1024 image patches and use the ground truth images for labeling each image patch as disaster hit or not hit. This makes the training process easy. Then we use VGG-16 architecture of CNN which identifies the features of the image patches. According to my survey, CNN algorithm is the one which satisfies all the conditions by omitting the limitations which are in the above related techniques and papers i.e. papers mentioned in the Related Works section. The advantage of CNN over other image processing technique is that, Convolution neural networks (CNN s) are a specialized kind of neural network for processing input data that has an inherent grid-like topology. Generally, the input data to a CNN will have natural structure to it such that nearby entries are correlated. Examples of this type of data are 1-D audio time series data and 2-D images. A CNN attempts to find the most relevant patterns that help determine how to accomplish the given task. Whereas the disadvantages of image processing are, initial cost can be high depending on the system used such as CCD or digital cameras that are used for digital image processing have disadvantages like: Memory card problems, higher cost, and battery consumption. We also considered the size of the data-set which we have used and made sure that there is no over fitting of the data-set. We used the Loss and Accuracy Trends to make sure that there is no over-fitting of the data.

II. OUR APPROACH

A. DATA-SET GENERATION

For the purpose of the ADDS we need a data-set which consists of images which show the region of interest (ROI) before and after the disaster has hit. This kind of data-set wasn't available online, so we used Google Earth to prepare the data-set. This data-set consists of set images (i.e. each image is a set of two which shows the region before and after disaster has hit) of several places of Joso City, Japan. All the images in the data-set are of same resolution. This data-set has about 50 sets of images. These images have a lot of features, and we found a way to minimize the number of features in the that to make it simple for the CNN to learn the important features that we require for the ADDS. There- fore we used the Ground Truth images. The ground truth images are nothing but the images which has the flooded areas in the white color and the unaffected areas in black color.These ground truth are usually made by domain experts, but we had to make these ground truth images ourselves since the ground truth images vary from place to place and it is very difficult to find a data-set consisting of both the normal images along with their ground truth images. We prepared these ground truth images using Adobe Photoshop.

We have about 50 ground truth images for all the 50 set of images in our data-set. We have used 75% of the images of the dataset for the purpose of training and the other 25% for the purpose of testing our model.

Figure 1: Image from Data-set before disaster has struck.

Figure 2: Image from Data-set after disaster has struck.

B. PRE-PROCESSING OF DATA-SET

The pre-processing step is very crucial for making the model efficient so that model has good detection rate. In the pre-processing step we need to cut the 50 sets of images into smaller patches each of size 32*32. Hence, we will get 1024 patches for each big satellite image. We repeated the same process also for the ground truth images. After the preparation of patches is over we need to label those patches for efficient training. Labeling of the image patches is one of the important steps before we proceed to the process of training our model. Here based upon the ground truth images we label the image patches as disaster hit or not hit. This is important as the CNN needs to learn the features of the disaster image patches. We have used the concept of finding out the percentage of the white pixels from the ground truth patches and if the percentage is more than 10% then we label the corresponding true image as 1 (disaster hit) or else if it less than 10% we label it as 0 (not disaster hit). The 10% which we used here is internationally acceptable standard. We have used the opencv and numpy packages for the labeling of the images.

C. ARCHITECTURE OF CNN

For our experiments, we have used the python libraries like OpenCv and Tensorflow for training and implementing our model. There are several pre-trained models available on the internet like LeNet(1998),

AlexNet(2012), ZFNet(2013), GoogleNet(2013), VGGNet(2014), etc. The VGG net consists of 16 convolution layers and has a very uniform architecture. It has 3*3 convolutions but has more filters. It is one of the most widely used models for feature extraction. Hence, we have chosen the VGGNet in our model for the purpose of training the model.

The pooling layers don't count as they do not learn anything. To reduce the number of parameters in such very deep networks, small 3x3 filters are used in all convolutional layers with the convolution stride set to 1. At the end of the net- work are three fully-connected layers. The VGG networks use multiple 3x3 convolutional layers to represent complex features. Therefore, here we chose VGGNet architecture similar to AlexNet for our network which is going to produce minimum error rate and is in top 5 error rate architectures ILSVRC (ImageNet Large Scale Visual Recognition Competition) 2014 competition.

D. TRAINING OF THE CNN

As mentioned above we have used the ground truth images to label the true images and now after labeling the true images we use the true images to train our model. The data-set is randomly divided into the training and testing set. Our ADDS model was trained for 10 epochs on our Flood data-set. This was trained done on the Nvidia Tesla K80. This was done in the Google Colab. The training of the CNN can be increased from 10 to more epochs depending on the availability of more computation power. By training the CNN the CNN learns the features of

the disaster affected regions and would easily distinguish between disaster hit and not hit areas in an image.
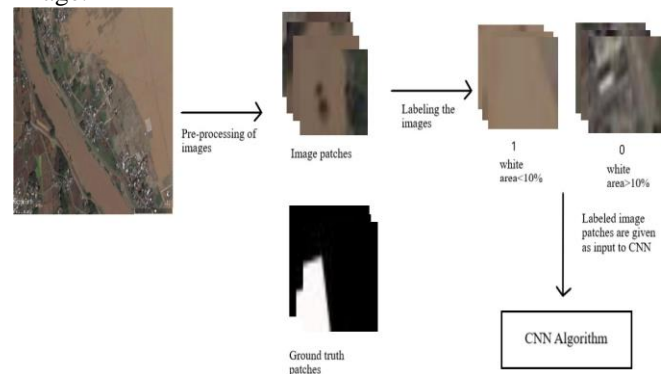
Figure 3: Image of the training process.

E. DETECTION OF FLOOD HIT REGIONS

We use the above trained model for the prediction of the disaster hit image patches from the test image patches.

Here we have taken a satellite image from the disaster hit area and we cut it into 1024 image patches, now these image patches are given as input to the trained machine learning model.

Whenever the model detects a image patch as flood hit then it draws a red border around the image patch. Now we will have a folder which will consist of image patches in which the flood hit image patches have a red border. Now we have to resize all the flood hit image patches since its size gets increased due to addition of the border around the image. Now after resizing all the images we have to merge those images again together using their filenames as their filenames will be in a ordered fashion. Now after we merge all the image patches we again get a whole image which will have all the disaster hit pixels inside a red box and the normal will be outside the red box.
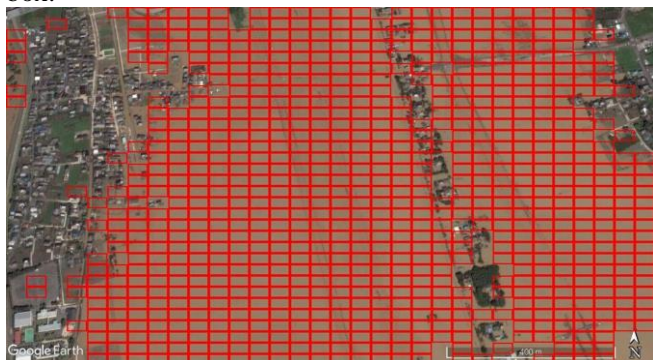

Figure 4: Detected flood hit regions inside red boxes

III. RESULTS AND DISCUSSION

In this paper, one of the major advantages of the proposed model is use of small data-set for the purpose of detection of the flooded region. In the model presented by Aparna R Joshi, Isha Tarte, Sreeja Suresh and Shashidhar G Koolagudi in the paper "Damage Identification and Assessment using Image Processing on Post Disaster Satellite Imagery" [3] published in 2017 they used 1500 images or data regions compared to a single image which we used for the purpose of training. Even if we want a well-trained model we will need about 10 images which will give us 10240 image patches. Hence, ADDS model is more suitable and apt when we have a limited number of images.

To recognize and detect the areas as disaster hit or not hit, we should look at the accuracy of the VGG-16 model which is trained by using the data-set generated for this purpose. We have trained our model using 770 images, and we used the rest 254 images as a validation set. The validation accuracy of the model was approximately 70% (refer Figure 5), though our data-set is small the accuracy of the model is encouraging. We used a small data-set during execution due to several constraints like computation cost etc. If the size of the data-set is increased we may be able to get 90% accuracy for the model..

Despite the perfect score, the result might be attributed to the over-fitting on the data-set. Therefore, after pre-processing, we must analyze the loss and accuracy trends of our model (ADDS) by changing the batch and epochs size. Through these above graphs (Figure 4, Figure 5) we came to a conclusion that our model, ADDS is perfectly fitted. But, we need to

improve our accuracy and decrease the loss values which can be done by increasing the batch size of the model from 16 to 32. The Figure-6 and Figure-7 shows the change in the loss and accuracy trends when the batch size is changed from 16 to 32.

A. LOSS AND ACCURACY TRENDS

There are basically 2 cases in deep learning regarding loss and accuracy values. Loss trends are generally analyzed with respect to Training Loss (TL) and Validation Loss (VL). Accuracy trends are generally analyzed with respect to Training Accuracy (TA) and Validation Accuracy (VA).

1. Under-fitting

The model is said to be under-fitted only in the case when training loss is greater than validation loss (**TL>VL**), and when training accuracy is much lower than validation accuracy (**TA<<VA**).

2. Over-fitting

Over-fitting generally refers to a model that models the training data too well. It can be recognized when the training loss is much lower than the validation loss (**TL<<VL**), and when the training accuracy is slightly greater than the validation accuracy (**TA>VA**). Therefore, we can say from above images that the model got over-fitted as **TL<VL** and **TA>VA**.

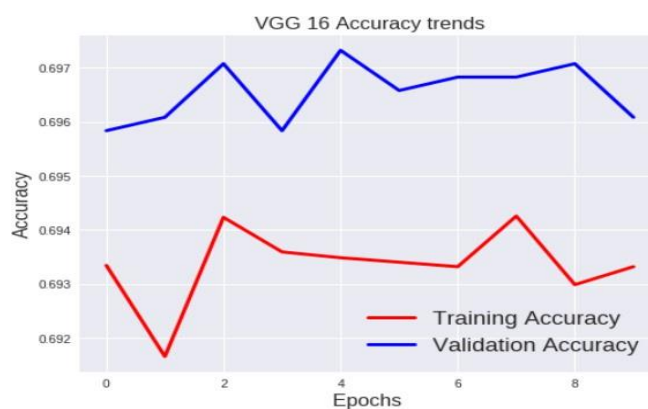**Figure 5: VGG-16 Loss trends when batch size is 16.**

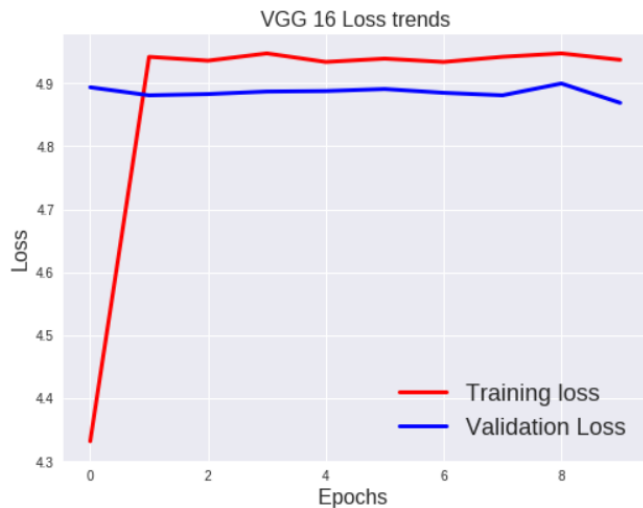**Figure 6: VGG-16 Accuracy trends when batch size is 16.**

Figure 7: VGG-16 Loss trends when batch size is 32. Figure 4 is the final output of our model. In the Figure 4 we can clearly see the areas which are affected by floods inside the red box and the areas which are not affected by the floods are outside the red box. In this way we can automate the process of detection of the disaster hit regions.
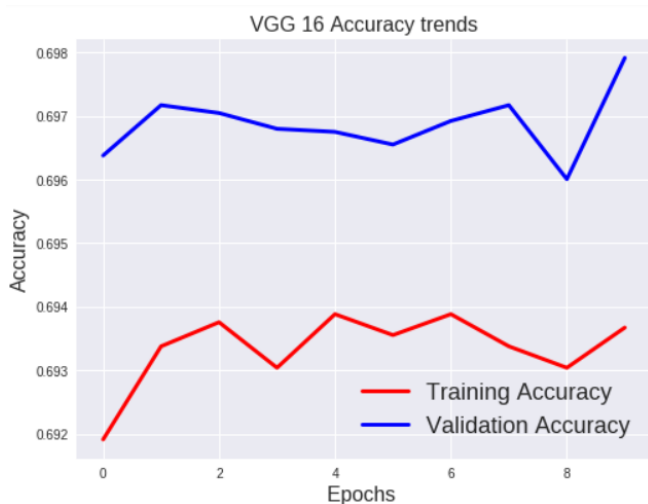

Figure 8: VGG-16 Accuracy trends when batch size is 32.

IV. CONCLUSION

A disaster is a deadly natural occurrence which causes huge damage to both life and property. After the disaster has hit, disaster relief forces start the rescue operations and try to save as many lives as possible. At present the disaster relief forces have to rely on phone calls and drone images during this process. This is where we wanted to change things and make that process automated so that a lot of time can be saved since in those cases time is of essence. ADDS is a model which automates the process of detection of flood hit regions. For the purpose of training our model we built a dataset which consists of 50 pairs of images (post and pre disaster images).

Our ADDS model achieved an accuracy of 70% after running the model for 10 epochs. This accuracy can be improved with better training of model using good computation resources. We have trained ADDS using two batch sizes 16 and 32 to check and remove the over-fitting of the data. The result of ADDS model are the modified satellite images where the disaster hit regions are highlighted and are present inside a red box. The results are very easy to interpret even for a layman.

In this model (ADDS) we focused on detecting the flood only from the land areas and didn't consider the river paths. The proposed model has been trained using the satellite images of the Joso City, Japan which was hit by floods during the year of 2015. But this model can be used for detecting floods in any area. The flood affected areas have similar features and thus this model can be generalized to for any area. The detection of the flood affected areas is very quick and easy since we use a pre-trained model every time, and we also don't need ground truth images for the purpose of detection.

FUTURE SCOPE

Future research may include considering the river paths also during the training phase. We would also like to make a more generalized model which could be used for any disaster since in the present model we mostly focused on flood affected areas. We would also like to make the model more user-friendly like developing an app using this model by using technologies like TensorFlow Lite.

ACKNOWLEDGEMENTS

REFERENCES

1. M. Jahari, S. Khairunniza-Bejo and A.R.M. Shariff (Eds.) "Change detection using a local similarity measure", in IEEE Conference on Innovative Technologies in Intelligent Systems and Industrial Applications, 12-13 July, 2008 Multimedia University, Cyberjaya, Malaysia.
2. W. Fernando, C. Canagarajah and D.R. Bull, "A unified approach to scene change detection in uncompressed and compressed video",in Digest of Technical Papers, International Conference on Consumer Electronics, Nineteenth in the Series (Cat. No.00CH37102), 13-15 June 2000, Los Angles, CA, USA, USA.
3. A.R. Joshi, I. Tarte, S. Suresh, S.G. Koolagudi, "Damage identification and assessment using image processing on post-disaster satellite imagery", in IEEE Global Humanitarian Technology Conference (GHTC), 19-22 October. 2017, San Jose, CA, USA.
4. Karthik, B.R. Shivakumar, "Change detection using image differencing: A study over area surrounding Kumta, India", in Second International Conference on Electrical, Computer and Communication Technologies (ICECCT), 22-24 Feb. 2017.
5. S. Kumar, M. Anouncia, S. Johnson, "Agriculture change detection model using remote sensing images and GIS", in International Conference on Radar, Communication and Computing (ICRCC), 21-22 Dec. 2012, Tiruvannamalai, India.
6. T. Liu, G. Gigli, G.A. Lampropoulos, "Change detection methodology based on region classification fusion", in 10th International Conference on Information Fusion, 9-12 July 2007, Quebec, Canada.
7. A.J. Reno, D.B. David,"An application of image change detectionurbanization", in International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015], 19-20 March 2015, Nager-coil, India.
8. W. Sun, H. Tang, H. Chen, G. Yu, "Research of unsupervised image change detection algorithm based on 2-D histogram", in IEEE 10th INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING PROCEEDINGS, 24-28 Oct. 2010, Beijing, China.

Retrieval Number: E2890038519/19©BEIESP
Journal Website: www.ijitee.org

128

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.