International Journal of Applied Science and Engineering

Disaster management using deep learning on social media

Bhavana, Parvathi Ramasubramanian*

School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India

ABSTRACT

The main goal of this proposed work is to provide solutions for disaster management using deep learning algorithms on social media images. The MNIST dataset was used to initially build the deep learning models. The images were trained using LeNet5, VGG13, VGG 16 and LSTM deep learning models. Later a dataset containing 3460 images were taken from social media. The labels earthquake, wildfire and floods were used to achieve classification results. The images were trained and validated using LSTM, VGG13 and VGG16. The performance of the algorithms is compared and the disaster response technique is generated based on the image classification and disaster management strategies are provided based on classification.

Keywords: Deep learning, VGG, CNN, MNIST, Disaster management.

1. INTRODUCTION

During natural disasters, a swift and immediate response must take place in order to minimize damage and save lives. For this, governments often employ search and rescue, field survey teams and local governing bodies to assess the damages and the nature of the disaster. These survey methods are highly risky and damage assessment is often done much later, which may also delay timely relief efforts also there is insufficient information for rescue operators and for citizens to know the severity of disaster and reach safety. In these scenarios, information from social media can be used as it is cheap and timely. However the extremely prevalent nature of social media offers a solution to this problem. As it allows local governments and citizens to share distress messages and damage information. Also the ease of use of social media ensures a large volume of participation during natural disasters.

Social media allows sharing of data in multiple formats, including text, images and video. These media can be useful for assessing damages during disasters; the concise nature of social media posts makes it easy to gather essential information quickly. For quick classification of algorithms. Early works (Takahashi et al., 2015; Sit et al., 2019; MacEachren et al., 2011; Kongthon, 2011), in this area have studied the role of social media during disasters by measuring social media activity with respect to geo-spatial relations and reliability of social media during disasters. Other recent works (Affonso et al., 2017; Pedamonti et al., 2018; Kumar et al., 2019; Yu et al., 2019; Yang et al., 2019; Barz et al., 2019) have studied tweets using CNN to classify disasters and damages.

Deep learning is a subdivision of machine learning where the algorithms use multiple layers to form a neural network which extracts features from data. In image classification problems, deep learning has been the most used and successful approach. Hence in the proposed work where a dataset containing images is used and the performance of multiple deep learning algorithms is compared, will be greatly beneficial in tackling disaster management and assessing the performance of the deep learning algorithms.

OPEN ACCESS

Received: January 27, 2021

Accepted: March 22, 2021

Corresponding Author:
Parvathi Ramasubramanian
parvathi.r@vit.ac.in

Publisher:
Chaoyang University of Technology
ISSN: 1727-2394 (Print)
ISSN: 1727-7841 (Online)

Neural networks can be classified into three types, convolution neural network (CNN), artificial neural network (ANN), recurrent neural network (RNN). CNN consists of layers known as convolutional layers, pooling layers, fully connected layers, and normalization layers and is most commonly used in computer vision applications such as image processing and classification, examples of CNN models are LeNet5, VGG16, etc. ANN is a neural network which is inspired by the biological brain; an artificial neural network consists of a collection of neurons. Each neuron is a node which is connected to other nodes to form connections. Each link has a weight, which is used for generating predictions, e.g. perceptron. Recurrent neural networks (RNN) are a type of neural network where output from previous steps is taken into account and is used as input for the subsequent steps. Whereas in other types of neural networks this does not occur and all the inputs and outputs are independent of each other. An example of RNN's would be the long short-term memory (LSTM) algorithm.

To further improve the performance of the deep learning models by reducing overfitting, techniques such as data augmentation, adding additional dropout layers and early stopping were used.

Data augmentation refers to the modification of the dataset by creating additional images from the existing image pool. This is done by changing the rotation, orientation, by cropping, etc. of an image to create a new image. This technique is very beneficial when the sample size is limited for model training. Often during the training phase of a model, peak accuracy is reached earlier than the total number of epochs a model is trained for, this overtraining causes overfitting which is not desired. Early stopping is used to combat this; in early stopping, the validation accuracy is measured regularly during training. If there is a stagnation or fall in the validation accuracy values, the training period is then stopped.

In this paper, VGG 13, VGG16 which are types of CNN deep learning models and LSTM which is a type of RNN deep learning model are trained on disaster images taken from social media and the modified national institute of standards and technology (MNIST) dataset and their performances are measured and compared based on accuracy and a disaster management technique is suggested.

2. LITERATURE SURVEY AND EXISTING MODEL

During the sudden onset of natural disasters, disaster managers and responders depend upon timely and accurate information about the disaster situations (e.g. damages) in order to generate effective disaster management strategies and make quick response decisions. In a study performed by Xiao et al. (2015), they performed analysis on social media data by analysing geo-spatial trends and by performing an empirical analysis of tweets about Hurricane Sandy in New York City in particular and how information about casualties and damage, donation efforts, and alerts are more likely to be used and extracted to improve recovery response during a time-critical event.

Later geographically grounded situational awareness was also studied by MacEachren et al. (2011). In an article by Steiglitz et al. (2014), various methodologies and frameworks of social media analytics are explored along with its role in big data analytics. Charalabidis et al. (2014), worked on using social media as a tool for communication during disasters, their results demonstrate that social media significantly improves communications during disasters. Sit et al. (2019) worked on the identification of disaster related tweets and analyse in their context using deep learning.

Due to the increase in use of social media by various organizations and emergency response teams to determine mass public opinions and trends and to gauge reactions to an event, Roche et al. (2013) studied the directional flow of communication and how social media efficiently facilitates response and recovery efforts. Chowdhury et al. (2013) extracted and performed classification of tweets generated during a natural disaster based on factors such as, if the tweet included caution and advice or if it provided information regarding the casualty and damage along with the information source to improve our assessment and knowledge of a disaster situation. Yu et al. (2019) examined sampled tweets generated during the 2012 Hurricane Sandy using CNN, and improved upon the previous study by creating more classification categories such as infrastructure and resource.

To study the problem of information reliability of twitter Yang et al. (2019) used hurricane harvey as a use-case to create a twitter credibility framework. Social media usage was studied by Kaigo (2012); Al-Saggaf and Simmons (2015); Bird et al. (2012); Kongthon et al. (2014); Takahashi et al. (2015) by analysing social media during various natural disasters.

Barz et al. (2019) worked on improving the usage of social media for disaster response by improving and working on algorithms to retrieve flood images from social media. Similarly in 2014, Ashktorab et al. (2014) created Tweedr which worked by mining twitter to find information for disaster relief workers during natural disasters. It performed classification using sLDA, SVM, and logistic regression.

The use and performance of deep learning algorithms on image datasets was studied by Affonso et al. (2017) for biological images using CNN, Razzak et al. (2018) for medical images using many deep learning models, Li et al. (2016) worked on plankton image classification using deep residual networks in 2016. Later, Pedamonti (2018) compared various non-linear activation functions for deep learning networks on the MNIST classification dataset.

The following paper is organized as follows: section 3 explains the proposed model; section 4 contains the implementation details; section 5 contains results and discussion, section 6 concludes the proposed work and section 7 lists the scope for improvement.

3. PROPOSED MODEL

3.1 Architecture

The proposed model can be divided into 2 steps, building and analysing the neural networks using the MNIST dataset and implementation of the better performing models on social media data and suggesting disaster management techniques based on classification. The flowchart of the proposed model can be seen in Fig.1.

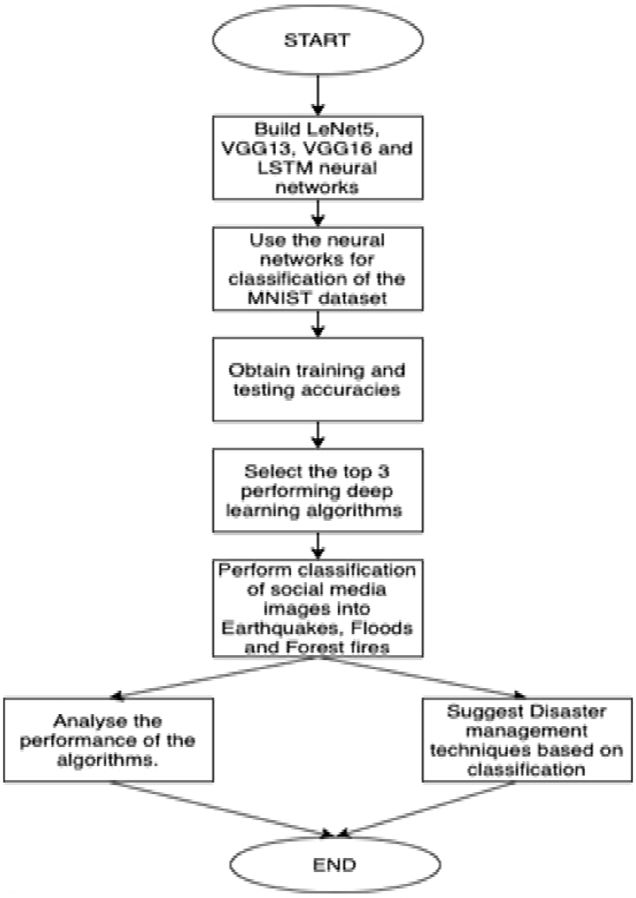

Fig. 1. Flowchart of the proposed model

3.2 Methodology

Social media which are often taken from various angles, alignments, etc. Due to this, it is difficult for simple algorithms to achieve high accuracy. Convolution neural networks (CNN) and recurrent neural networks (RNN) have been shown to have the best performance for image classification and thus will be greatly beneficial for accurately classifying images taken from social media.

First to build the deep learning models and generate a baseline of performance, the MNIST database of handwritten digits is used. For this, LeNet 5, VGG13, VGG16 and LSTM models were used.

LeNet5 is a simple feed-forward CNN. LeNet5 consists of convolution layers (using 5x5 kernel and with sigmoid activation), average pooling layers (2x2 kernel). It is a simple neural network that is quick and easy to train.

VGG is a deep convolution layer consisting of blocks which was very successful in the ILSVRC 2014 competition for classifying imagenet images. A VGG block consists of a convolutional layer, ReLU activation and a max pooling layer. The VGG algorithm is built using multiple blocks, in the original VGG paper, the convolution layers used 3x3 kernels and the pooling layers used 2x2 kernels .

LSTM is an artificial recurrent neural network (RNN) architecture which uses feedback connections. Therefore it can also process a long series of data values such as videos. LSTMs have three gates; input gate, forget gate, and output gate which control the movement of data during the execution of the algorithm. LSTMs are designed to handle and solve the vanishing gradient problem which is found in other RNN's.

Then based on the performance of the models on the MNIST data, models were chosen to be trained on the disaster images dataset. The disaster images dataset consists of images classified with the following labels, earthquake, wildfire and floods. After achieving good classification accuracy, using guidelines and frameworks, disaster management techniques and solutions are suggested.

3.3 Implementation Details

First to build the deep learning models and generate a baseline of performance, the MNIST database of handwritten digits is used. The MNIST database consists of 60,000 training images and 10,000 testing images and was directly imported using Keras and thus required no image pre-processing. The training period was for 20 epochs. For this, LeNet5, VGG13, VGG16 and LSTM models were built using Keras and Tensorflow packages on python3.6.

Then based on performance of the MNIST data were selected and were trained on the disaster images dataset which was compiled from various social media sites. The disaster images dataset was created by Kumar (2019). It was built to create an automated natural disaster detection system using CNN. A sample image of the disaster images dataset can be seen in Fig 2. The dataset consists of 3460 images classified with the following labels, earthquake, wildfire and floods. Image pre-processing methods such as resizing, conversion to grayscale, etc. were implemented based on the input requirements of the deep learning model. The test-train split was 80:20 and therefore 2768 images were used for training and 692 images were used for testing. The training period was for 450 epochs, whereas for the MNIST dataset only 20 epochs.

After monitoring and comparing the training and testing performance, overfitting which can be attributed to the noise in the data. To resolve this issue, the deep learning models and the dataset were modified.

First, data augmentation was performed. For this new images were created by changing the rotation, padding orientation of the existing images randomly. This allowed us to increase the dataset size from 3460 to 7000 images. Then, two additional dropout layers were added into the neural network. For the VGG16 and VGG13 algorithms, the dropout layers were added after the max-pooling layer in the 2nd and the 4th blocks and in the LSTM algorithm, an input dropout of 40% has been implemented. To further reduce overfitting, early stopping was implemented during the model training phase by checking for drop of plateauing of accuracy values during validation.

After achieving good classification accuracy, emergency disaster management techniques and solutions are suggested using national guidelines and frameworks.

4. RESULTS AND DISCUSSION

In the proposed work, a model for classifying disaster images has been created. Initially the deep learning models (LeNet5, VGG13, VGG 16, and LSTM) were built and tested on the MNIST database of handwritten digits. The MNIST database contains 60,000 training images and 10,000 testing images and was directly imported using Keras and thus required no image pre-processing. The training period was for 20 epochs. The resulting accuracy values are as listed below in Table 1 and can be visualized using Fig 3, that LSTM has the best testing accuracy at 98.73%, followed by VGG 16 at 98.2%, VGG 13 at 97.99% and LeNet5 at 97.70%.

Due to comparatively better accuracy values, the deep learning models VGG13, VGG16 and LSTM were chosen and used for the disaster images dataset. The disaster images dataset consists of 3,460 images classified with the following labels, earthquake, wildfire and floods. Pre-processing methods such as resizing, conversion to grayscale, etc. were implemented. The test-train split was 80:20 and therefore 2,768 images were used for training and 692 images were used for testing. The training period was for 450 epochs, which was required to reach the peak accuracy value whereas for the MNIST dataset only 20 epochs were required to reach peak accuracy values. This is due to the larger size of the MNIST dataset and lower number of features extracted from the MNIST dataset. The accuracy values are as listed below in Table 2. It can be seen that VGG 16 at 76.1% followed by VGG 13 at 70.03% and LSTM at 62.97%. When we compare the training accuracy values of the deep learning models between the MNIST dataset and the disaster images dataset, a large difference is present. This is due to the much smaller size of the disaster images dataset. It is also observed that the testing accuracy values are also relatively low compared to the training accuracy values; this suggests that over-fitting is taking place. Classification by the models can be seen in Fig. 4 – Fig. 6 and the accuracy comparison can be visualized in Fig 7.

Fig. 2. Sample image from disaster images dataset

Table 1. Training and testing accuracy values of deep learning models with MNIST dataset

Accuracy model	Training accuracy	Validation accuracy	Testing accuracy
LeNet5	98.23%	97.70%	97.70%
VGG 13	98.8%	97.99%	97.99%
VGG 16	98.96%	98.2%	98.2%
LSTM	99.32%	98.73%	98.73%

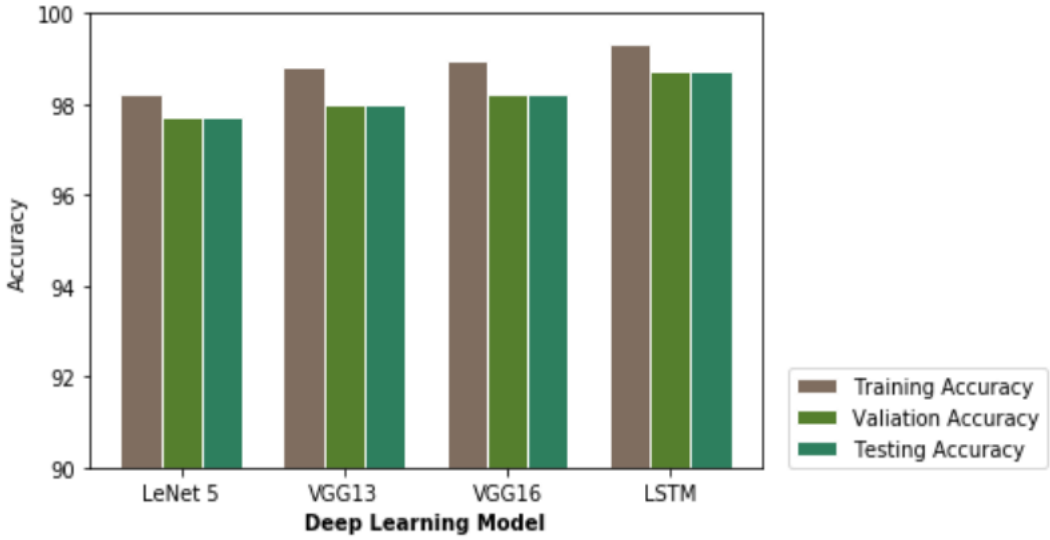

Fig. 3. Comparison of training, validation and testing accuracies of the deep learning models for MNIST dataset

Fig. 4. Image classification of wildfire using VGG16

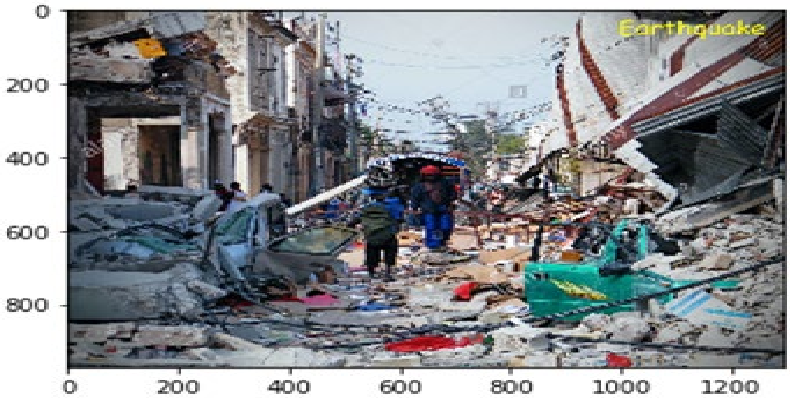

Fig. 5. Image classification using LSTM

https://doi.org/10.6703/IJASE.202106_18(2).014

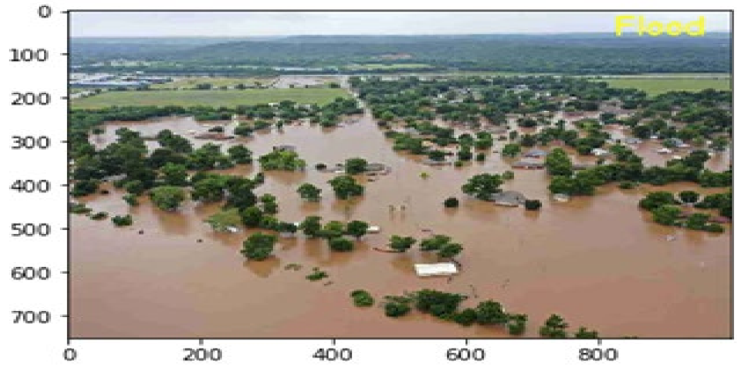

Fig. 6. Image classification using VGG13

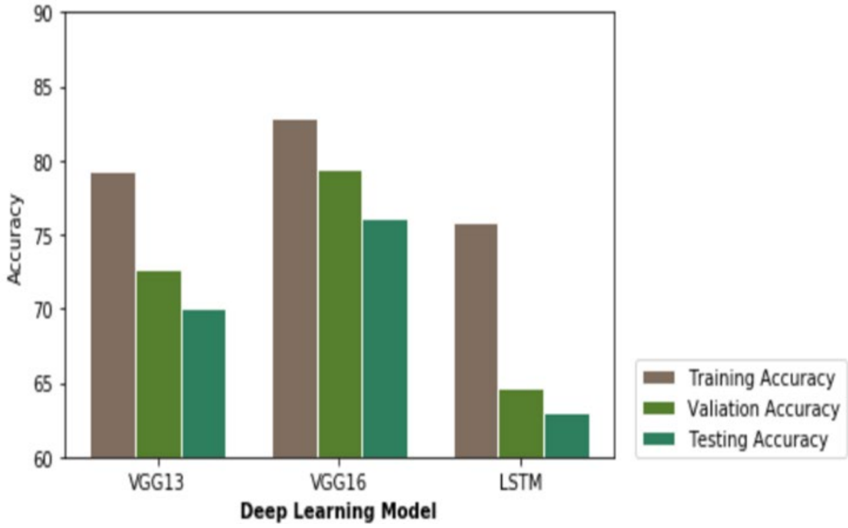

Fig. 7. Comparison of training, validation and testing accuracies of the deep learning models for disaster images dataset

Table 2. Training and testing accuracy values of deep learning models with disaster images dataset

Accuracy model	Training accuracy	Validation accuracy	Testing accuracy
VGG 13	79.34%	72.6%	70.03%
VGG 16	82.93%	79.47%	76.1%
LSTM	75.8%	64.6%	62.97%

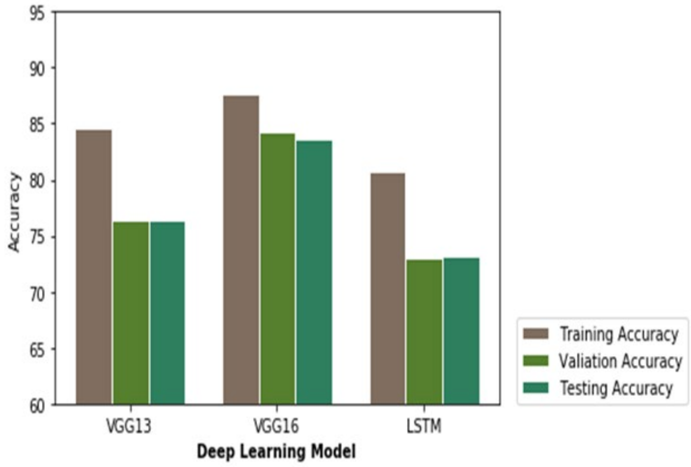

Fig. 8. Comparison of training, validation and testing accuracies of the modified deep learning models for disaster images dataset with data augmentation

Table 3. Training and testing accuracy values of deep learning models with disaster images dataset with data augmentation

Accuracy model	Training accuracy	Validation accuracy	Testing accuracy
VGG 13	84.52%	76.39%	76.39%
VGG 16	87.69%	84.3%	83.6%
LSTM	80.7%	73.02%	73.14%

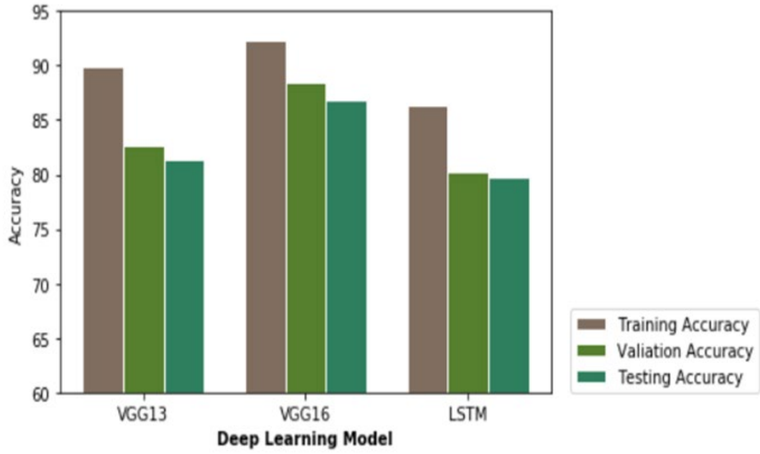

Fig. 9. Comparison of training, validation and testing accuracies of the modified deep learning models for disaster images dataset with data augmentation and reduced overfitting

Table 4. Training and testing accuracy values with data augmentation and modified deep learning models on the disaster images dataset

Accuracy model	Training accuracy	Validation accuracy	Testing accuracy
VGG 13	89.94%	82.7%	81.33%
VGG 16	92.4%	88.43%	86.825%
LSTM	86.37%	80.2%	79.82%

To improve training accuracy values and to avoid over fitting, additional steps were implemented. First, the dataset size was increased to 7000 using data augmentation. For this new angles and orientations of existing images were created. We can see the improvement in the model performance after performing data augmentation in Table 3 and Fig 8. We can note that the testing accuracy values are now 76.39%, 83.6% and 73.14% for VGG13, VGG16 and LSTM respectively. This is an improvement of 9%, 9.86% and 16.15% for VGG13, VGG16 and LSTM respectively.

Then new dropout layers were added to reduce overfitting. To further reduce overfitting, early stopping was implemented during model training and the model was trained on the new dataset which used data augmentation; these modifications improved the accuracy values of VGG13, VGG16 and LSTM to 81.33%, 86.825% and 79.82% respectively as shown in Table 4 and can be visualized in Fig 9. An improvement of 16%, 14% and 26% in the testing accuracy values VGG 13, VGG 16 and LSTM respectively when compared to the first iteration where data augmentation and other algorithmic modifications were not implemented.

From the above images and tables, it can be inferred that VGG 16 has the best performance (accuracy) followed by VGG 13 and LSTM. In conclusion, first the deep learning models were built and were implemented for the MNIST dataset followed by the disaster images dataset. Based on the performance of the models, fine-tuning methods such as data augmentation, early stopping, etc. were added. This has improved the performance of the deep learning models.

5. CONCLUSION

The main goal of this proposed work was achieved and solutions for disaster management using deep learning algorithms on social media images were provided. The MNIST dataset was used to initially build the deep learning models. The images were trained using LeNet5, VGG13, VGG 16 and LSTM deep learning models. Later a dataset containing 3460 images were taken from social media. The labels earthquake, wildfire and floods were used to achieve classification results. The images were trained and validated using LSTM, VGG13 and VGG16. The performance of the algorithms is compared and the disaster response technique is generated based on the image classification. Technical issues and issues of overfitting were successfully resolved using techniques of data augmentation, early stopping and dropout layers. The final testing accuracy values for LSTM, VGG13 and VGG16 were 79.82%, 81.33% and 86.825% respectively. It can be inferred that VGG 16 comparatively has the better performance (accuracy) followed by VGG 13 and LSTM for disaster images. Finally after classification of the images, disaster management techniques were suggested. The proposed work was successful in demonstrating the performance of various deep learning algorithms on social media disaster images.

6. SCOPE FOR IMPROVEMENT

The proposed work uses a dataset of images downloaded from social media, the project can be further developed by improving the processing power of the computational device and process large volume of social media images directly during the natural disaster. Furthermore using a larger dataset from more varied sources can be used to develop a better deep learning model.

REFERENCES

Affonso, C., Rossi, A.L.D., Vieira, F.H.A., de Leon Ferreira, A.C.P. 2017. Deep learning for biological image classification. Expert Systems with Applications, 85, 114–122.

Al-Saggaf, Y., Simmons, P. 2015. Social media in Saudi Arabia: Exploring its use during two natural disasters. Technological Forecasting and Social Change, 95, 3–15.

Ashktorab, Z., Brown, C., Nandi, M., Culotta, A. 2014. Tweedr: Mining twitter to inform disaster response. 11th Proceedings of the International Conference on Information Systems for Crisis Response and Management, 269–272.

Barz, B., Schröter, K., Münch, M., Yang, B., Unger, A., Dransch, D., Denzler, J. 2019. Enhancing flood impact analysis using interactive retrieval of social media images, Archives of Data Science Series A, 5, 1–21.

Bird, D., Ling, M., Haynes, K. 2012. Flooding Facebook-the use of social media during the Queensland and Victorian floods. Australian Journal of Emergency Management, 27, 27–33.

Charalabidis, Y., Loukis, E.N., Androutsopoulou, A., Karkaletsis, V., Triantafillou, A. 2014. Passive crowdsourcing in government using social media. Transforming Government: People, Process and Policy. 8, 283–308.

Chowdhury, S.R., Imran, M., Asghar, M.R., Amer-Yahia, S., Castillo, C. 2013. Tweet4act: Using incident-specific profiles for classifying crisis-related messages. 10th Proceedings of the International Conference on Information Systems for Crisis Response and Management, 34, 1–5.

Kaigo, M. 2012. Social media usage during disasters and social capital: Twitter and the Great East Japan earthquake. Keio Communication Review, 34, 19–35.

Kongthon, A., Haruechaiyasak, C., Pailai, J., Kongyoung, S. 2014. The role of social media during a natural disaster: A case study of the 2011 Thai Flood. International Journal of Innovation and Technology Management, 11, 1440012.

Kumar, G. 2019. [Online], https://towardsdatascience.com/average-rolling-based-real-time-calamity-detection-using-deep-learning-ae51a2ffd8d2, (Accessed December 22 2019).

Li, X., Cui, Z. 2016. Deep residual networks for plankton classification. OCEANS 2016 MTS/IEEE Monterey, Monterey, CA, USA, 1–4.

MacEachren, A.M., Robinson, A.C., Jaiswal, A., Pezanowski, S., Savelyev, A., Blanford, J., Mitra, P. 2011. Geo-twitter analytics: Applications in crisis management. 25th International Cartographic Conference, 3–8.

Pedamonti, D. 2018. Comparison of non-linear activation functions for deep neural networks on MNIST classification task. Available: http://arxiv.org/abs/1804.02763

Razzak, M.I., Naz, S., Zaib, A. 2018. Deep learning for medical image processing: Overview, challenges and the future. In: Dey N., Ashour A., Borra S. (eds) Classification in BioApps. Lecture Notes in Computational Vision and Biomechanics, 26. Springer, Cham.

Roche, S., Propeck-Zimmermann, E., Mericskay, B. 2013. GeoWeb and crisis management: Issues and perspectives of volunteered geographic information. GeoJournal, 78, 21–40.

Sit, M.A., Koylu, C., Demir, I. 2019. Identifying disaster-related tweets and their semantic, spatial and temporal context using deep learning, natural language processing and spatial analysis: a case study of Hurricane Irma. International Journal of Digital Earth, 1205–1229.

Stieglitz, S., Mirbabaie, M., Ross, B., Neuberger, C. 2018. Social media analytics–Challenges in topic discovery, data collection, and data preparation. International Journal of Information Management, 39, 156–168.

Takahashi, B., Tandoc Jr, E.C., Carmichael, C. 2015. Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines. Computers in Human Behavior, 50, 392–398.

Xiao, Y., Huang, Q., Wu, K. 2015. Understanding social media data for disaster management. Natural Hazards, 79, 1663–1679.

Yang, J., Yu, M., Qin, H., Lu, M., Yang, C. 2019. A twitter data credibility framework—Hurricane harvey as a use case. International Journal of Geo-Information, 8, 1–21.

Yu, M., Huang, Q., Qin, H., Scheele, C., Yang, C. 2019. Deep learning for real-time social media text classification for situation awareness–using Hurricanes Sandy, Harvey, and Irma as case studies. International Journal of Digital Earth, 12, 1230–1247.