Detection of passable roads using Ensemble of Global and Local Features

Muhammad Hanif, Muhammad Atif Tahir, Muhammad Rafi
Smart Video Surveillance Lab, School of Computer Science
National University of Computer and Emerging Sciences, Karachi Campus, Pakistan
{hanif.soomro,atif.tahir,muhammad.rafi}@nu.edu.pk

ABSTRACT

Disastrous situations can be better managed by availability of timely and relevant information. Social media plays very important role in providing information of disastrous events. The working paper is based on "Multimedia satellite task: Emergency response for flooding events", as a part of MediaEval, 2018. The dataset was provided with tweets and their respective images, which may or may not acquire evidence of roads and its status of pass-ability. An ensemble approach is followed in this paper by combining local features and global features of images. Text contents of the tweets were processed by their TF-IDF scores. Moreover, two level classification is performed by applying Spectral Regression based Kernel Discriminant Analysis (SRKDA) on individual feature categories as well as ensemble of different feature types. It is observed that F_1 score produced by visual, text and ensemble of both text and visual features for evidence of road remain 74.58%, 58.30% and 76.61% respectively. The average F_1 score for evidence of road and its status of passability remain 45.04%, 31.15% and 45.56% for visual, text and ensemble of both visual and text features.

1 INTRODUCTION

Disastrous situations require effective and timely information. Data could be collected with the help of sensors, but it is expensive to configure sensors on every critical location. Social media could be the best way to collect information in various forms including text, images and videos. Social media data can be utilized in managing disastrous situations including earthquake, fire and flooding situations. Data collected from social media could assist disaster response organizations, so that timely action may be taken.

2 RELATED WORK

Response in emergency situations, has recently been helped by enormous amount of data collected through social media. Various studies have been performed to collect data from social media and its significance has been proved in spreading timely awareness in disastrous situations. There are various situations where social media data can produce fruitful results, including fire, earthquake and flooding events. Tweets can produce helpful results for fire detection, when analysed with Deep Neural Network and Support Vector Machine (SVM) [5]. Social media data can also provide early awareness in flooding situations, by providing relevant text and images [8]. Another flood related case study revealed that people living in vicinity of affected areas, usually tend to send more tweets regarding the event [6]. Moreover, Hierarchical Disaster Image Classification (HDIC) framework has been proposed, which utilized visual and text data to classify images according to particular category of disaster. The framework analyzed major disaster categories as well as their respective affects [11]. Similarly, various other studies has also utilized social media data for awareness of disastrous situations [10] [3].

MediaEval'18, 29-31 October 2018, Sophia Antipolis, France

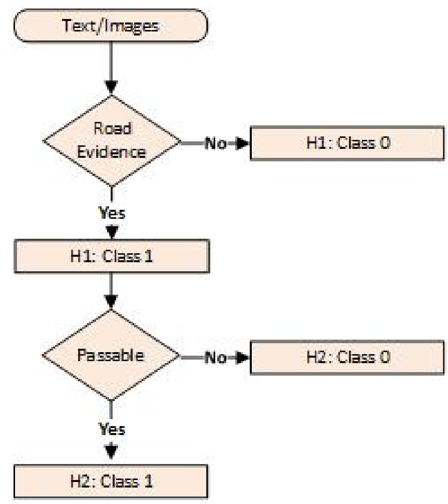

Figure 1: Data Flow showing the general processing steps for text and visual features.

3 APPROACH

Generally, tweets can contain two modalities such as (visual and textual), however visual information is not always present. Both categories of information are processed in two-level hierarchy of classification. At first level, evidence of the pathway is examined, while the status that road is actually passable or not is determined in second level of hierarchy, as shown in Figure 1. Various classifiers were implemented on training dataset, including Support Vector Machine (SVM), Multinomial Naive-Bayes, and Random Forest. However, best classification results were produced by Spectral Regression in combination with Kernel Discriminant Analysis (SRKDA) [9] [4] , which was also used to predict instances of test set.

3.1 Text

Text part of social media dataset was cleaned by eliminating hyperlinks, punctuations and symbols. Stop words were also removed from each of the tweet. User tags for each tweet were also extracted

Figure 2: Data Flow showing the general processing of text of tweets.

and combined with text of tweet. Then, TF-IDF [2] for each of word is calculated. The collection of TFIDF for each tweet is used for classification. Finally, Spectral Regression based Kernel Discriminant Analysis (SRKDA) is used to perform classification. The process is visualized in Figure 2.

3.2 Visual

Images are processed by using conventional visual descriptors provided by organizers of MediaEval, 2018 [1]. These descriptors were combined for the processing of each images, which includes Color and Edge Descriptor (CEDD), Color Layout (CL), Fuzzy Color and Texture Histogram (FCTH), Edge Histogram (EH), Joint Composite Descriptor (JCD) and Scalable Color (SC). Moreover, additional visual features for each image were extracted with the help of Deep Local Features (DELF) [7]. Finally, conventional visual descriptors and Deep Local Features were combined to create ensemble of visual features and processed through SRKDA for classification, as shown in Figure 3.

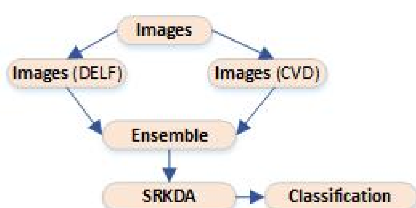

Figure 3: Data flow diagram showing method for image classification.

Deep Local Features (DELF): DELF is CNN-based local feature descriptor, specifically used for large-scale image recognition. It provides semantically local features, which are helpful in image retrieval. Furthermore, keypoint selection is also performed by proposed attention mechanism. The pre-trained model based on Convolutional Neural Network is also released, which is optimized for landmark recognition. Delf desriptor is extracted for each image. Delf descriptor is further processed with Gaussian Mixture Model (GMM) and fisher-vector having length of 2560 is generated for each image.

3.3 Ensemble

The ensemble is created by DELF and conventional visual descriptors, along with TFIDF of textual data. The prediction on ensemble of social media data is performed by the use of SRKDA followed by Nearest Neighbour classifier for prediction.

4 RESULTS AND ANALYSIS

Results are extracted for three different runs, which includes text, visual and ensemble of both text and visual features. First we will discuss results on validation set which is obtained from training data followed by discussion on test data. Table 1 shows the results on validation data which is obtained by taking 10% of training data. It is observed that DELF features did contribute in improving the accuracy using visual information. For the classification of road passability, around 80% F_1 score is obtained. However, for another scenario which is whether road is good enough to be used, score of around 0.60 is obtained.

Moreover, results were extracted by using visual, text and ensemble of both of them by using test dataset. The F_1 scores for road evidence classification has produced 74.58%, 58.30% and 76.61% for visual, text and ensemble of both. While, the average F_1 scores for road evidence and status of passability has produced 45.04%, 31.15% and 45.56% scores for visual, text and ensemble of both text and visual features, as shown in Table 2.

It is observed that better outcome has produced by using visual features along with SRKDA. It is part of future work to improve the results by deep analysis of local and global features.

Table 1: F_1 score on validation data for various scenarios.

Scenario	Global Features	Local Feature	Global& Local
Road Evidence	79.08	78.89	83.34
Passable	60.05	57.88	61.47

5 CHALLENGES AND FUTURE WORK

It has been observed that results produced for evidence of road are much better than its status of passable or not-passable. It is observed that less training data is available for classifier to check the status of passable roads. In future, data augmentation could be used to increase the quantity of visual training data. Moreover, TFIDF based sparse representation of text could be improved by the use of chi-square feature selection with probabilistic weighting scheme, as Chi-Square features are more discriminating about the model.

6 CONCLUSION

The paper presented results produced for the "Multimedia satellite task: Emergency response for flooding events". Different methods of feature extraction and classifiers are used to produce results. It is observed that an ensemble of conventional visual descriptors and DELF features with SRKDA classifier, has provided best outcome.

Table 2: F_1 score for road evidence and average F_1 score for road evidence and road passability.

Detail	Road Evidence(F1)	Average(F1)
Images	74.58	45.04
Text	58.30	31.15
Ensemble (Text and Images)	76.61	45.56

REFERENCES

[1] Benjamin Bischke, Patrick Helber, Zhengyu Zhao, Jens de Bruijn, and Damian Borth. The Multimedia Satellite Task at MediaEval 2018: Emergency Response for Flooding Events. In *Proc. of the MediaEval 2018 Workshop* (Oct. 29-31, 2018). Sophia-Antipolis, France.

[2] Max Bramer. 2007. the book" Principles of Data Mining. *Printed on acid-free paper© Springer-Verlag London Limited* 9003 (2007).

[3] Tom Brouwer, Dirk Eilander, Arnejan Van Loenen, Martijn J Booij, Kathelijne M Wijnberg, Jan S Verkade, and Jurjen Wagemaker. 2017. Probabilistic flood extent estimates from social media flood observations. *Natural Hazards & Earth System Sciences* 17, 5 (2017).

[4] Deng Cai, Xiaofei He, and Jiawei Han. 2011. Speed up kernel discriminant analysis. *The VLDB JournalâĂŤThe International Journal on Very Large Data Bases* 20, 1 (2011), 21–33.

[5] Ryan Lagerstrom, Yulia Arzhaeva, Piotr Szul, Oliver Obst, Robert Power, Bella Robinson, and Tomasz Bednarz. 2016. Image classification to support emergency situation awareness. *Frontiers in Robotics and AI* 3 (2016), 54.

[6] Zhenlong Li, Cuizhen Wang, Christopher T Emrich, and Diansheng Guo. 2018. A novel approach to leveraging social media for rapid flood mapping: a case study of the 2015 South Carolina floods. *Cartography and Geographic Information Science* 45, 2 (2018), 97–110.

[7] Hyeonwoo Noh, Andre Araujo, Jack Sim, Tobias Weyand, and Bohyung Han. 2017. Largescale image retrieval with attentive deep local features. In *Proceedings of the IEEE International Conference on Computer Vision*. 3456–3465.

[8] Robin Peters and João Porto de Albuquerque. 2015. Investigating images as indicators for relevant social media messages in disaster management.. In *ISCRAM*.

[9] Muhammad Atif Tahir, Josef Kittler, and Ahmed Bouridane. 2016. Multi-label classification using stacked spectral kernel discriminant analysis. *Neurocomputing* 171 (2016), 127–137.

[10] Nataliya Tkachenko, Stephen Jarvis, and Rob Procter. 2017. Predicting floods with Flickr tags. *PloS one* 12, 2 (2017), e0172870.

[11] Yimin Yang, Hsin-Yu Ha, Fausto Fleites, Shu-Ching Chen, and Steven Luis. 2011. Hierarchical disaster image classification for situation report enhancement. In *Information Reuse and Integration (IRI), 2011 IEEE International Conference on*. IEEE, 181–186.