https://doi.org/10.2352/ISSN.2470-1173.2020.8.IMAWM-270

PSO and Genetic modeling of Deep Features for Road Passibility Analysis during Floods

Naina Said[1], Aysha Nayab[1], Kashif Ahmad[2], Muhib Ullah[3], Touqir Gohar[1] and Ala Al-Fuqaha[2]
[1] University of Engineering and Technology, Peshawar, Pakistan
[2] Hamad Bin Khalifa University
[3] Norwegian University of Science and Technology, Norway

Abstract

In recent years, social media outlets have been widely exploited for disaster analysis and retrieving relevant information. Social media information can help in several ways, such as finding the mostly affected areas and information on casualties and scope of the damage etc. In this paper, we tackle a specific facet of social media in natural disasters, namely the identification of passable routs in a flooded region. In detail, we propose several solutions for two relevant tasks, namely (i) identification of flooded and non-flooded images in a collection of images retrieved from social media, and (ii) identification of passable roads in a flooded region. To this aim, we mainly rely on existing deep models pre-trained on ImageNet and Places dataset, where the models pre-trained on ImageNet extract object specific and the ones pre-trained on places dataset extract scene-level features. In order to properly utilize the object and scene-level features, we rely on different fusion methods including Particle Swarm Optimization (PSO) and Genetic Modeling of the deep features in a late fusion manner. The evaluation of the proposed methods are carried out on the large-scale datasets provided for MediaEval-2018 benchmarking competition on Multimedia and Satellites. The results demonstrate significant improvement in the performance over the baselines.

Introduction

Natural disasters may have devastating effects on both human lives and infrastructure. Due to the unpredictability of natural disasters, quick response might be a key to save human lives and prevent/mitigate other damages. The notion of immediate response to a natural disaster strongly depends upon the relevant information. For instance, the place of the disaster, number of injuries, number of buildings damaged and the scale of the event etc. It is a difficult and challenging task to obtain such information in emergency situations.In such events, most of the time, due to the lack of unavailability of reporters from news agencies or relevant authorities in the affected areas, a substantial amount of delay occurs [1]. In this modern world, the role of social media in decimation of information cannot be blind eyed. In recent years, Social media has emerged as an effective source of valuable information in such adverse events. However, several challenges are associated with collecting and processing of data shared in social media.

Realizing this fact, researchers are trying to develop systems ensuring an effective use of data available in social media outlets for different purposes and applications in emergency situations. One such application of social media in disasters has been proposed in MediaEval 2018 benchmark competition namely Multimedia Satellite Task: Emergency Response for Flooding Events [2], where social media outlets have been crawled and analyzed for the identification of passable routs in flooded regions. In detail, two different tasks, namely (i) Flood Classification for Social Multimedia (FCSM) and (ii) Flood Detection in Satellite Images (FDSI), have been proposed in the challenge. This work focuses on FCSM task, which has been further divided into two sub-tasks. The first sub-task, aims to identify whether an image provides evidence for road passability while in the second sub-task images showing passable and non-passable roads have to be differentiated.

In this work, we mainly rely on visual contents, and treat both tasks as a classification problem. Firstly, each image is analyzed for evidence of passability by classifying the images in flooded and non-flooded images. In the second task, the images providing evidences of passability are further analyzed and classified into passable and non-passable. It is important to mention that road passability depict the water-level and the surrounding context. For both tasks, we rely on a pool of deep models jointly used in a late fusion fashion, where two different weight optimization techniques, namely Particle Swarm Optimization (PSO) and Genetic Algorithm (GA), have been employed. We also evaluate the performances of several deep models pre-trained on both ImageNet [3] and Places datasets [4] in both tasks, individually. The motivation for deep features based approach is based on it its proven performances in different application domains [5, 6].

The main contributions of the work can be summarized as:

- We analyze how important are object and scene-level features in disaster analysis in still images.
- We demonstrate that ensemble of deep models significantly outperforms individual models in disaster analysis. Moreover, weights should be assigned to models on merit basis instead of treating them equally.

The rest of the paper is organized as follows: Section 2 provides detailed overview of the literature; Section 3 details the methodology while experimental setup, results and analysis are provided in Section 4. Finally, Section 5 concludes.

Related Work

Over the last few years, data shared in different social media outlets have gained significant attention of the research community in a diversified set of applications, such as healthcare [7], education [8] and crime management [9]. More recently great attention has been noticed in its applications in disaster analysis [1].

Research is being carried out in recognition of social media contents to ensure the development of techniques that can be used to infer information in emergency situations [10, 11].

In literature various techniques have been proposed to analyze disaster-related multimedia content, available in form of images, videos and text, for different purposes [1]. In this regards, Twitter has been widely exploited to extract and share information during a disaster event. The multi-modal (i.e., text, visual and audio) nature of the information available in Twitter makes it a better choice for disaster analysis. A better example can be found in the detailed and comprehensive study [12] analyzing the use and importance of social media during 2010 floods in Pakistan. A similar kind of study has been conducted in [13], where the authors target a typhoon occurred in Philippines as a case study and analyze the before and after usage of twitter during the event. In another work, to separate tweets relevant to storm and hurricane from the rest, Troung et al. [14] used a Bayesian approach. The system relies on a set of features containing nine different types of features for training a Bayesian classifier. As compared to the feature set obtained from BOW model comprising of more than 3000 features, feature set used for this study was comparatively very small, making it a more suitable choice for low powered hand held devices. Another study [15] addresses the short comings of conventional approaches used for detection of disasters in social media content. In this study, a crisis mapping system was proposed where for classification of disaster related content, SVMs classifier was trained on different linguistic features. In addition, the location of origin of a tweet was determined using a novel geo parsing technique.

A vast majority of the literature also exploits images shared in different outlets of social media for disaster analysis [1]. For instance, Ahmad et al. [16] collected and analyzed images from different sources of social media for disaster analysis. In this regard, the additional information available in different platforms of social media with images have been proved very effective. These information are either utilized seperately or in combination with other sources/types of information. For instance, Bischke et al. [17] found the additional information very useful in their multi-modal analysis of disaster-related content from social media. The multi-modal information are exploited both ways i.e., separately as well as in combination with visual information. Alam et al. [18] also utilize multi-modal information including images and meta-data by incorporating human experts' knowledge and machine learning algorithms for disasters analysis.

Disaster analysis in social media contents has also been part of a benchmark competition namely MediaEval for three years with slight changes in the task each year. In MediaEval 2017, the task aimed the identification of flooded regions in social media data and satellite imagery [19]. Majority of the approaches proposed for the task relied on multi-modal information with particular focus on existing models pre-trained on ImageNet and Places datasets for visual information. In MediaEval-2018, the participants were asked for identification of passable roads in a flooded region. Similar to previous year, majority of the approaches relied on deep architectures for classification of images showing evidence of passability, and classifying them into passable and non-passable. This paper target the challenge proposed in MediaEval-2018 relying on visual information, only.

Methodology

As mentioned earlier, in this work we treat both tasks as classification problem. Figure 1 provides the block diagram of the proposed methodology. In both tasks, as a first step, deep features are extracted via several deep models pre-trained on both ImageNet and Places datasets, followed by Support Vectors Machines (SVMs) based classification. In the final step, classification scores of all classifiers trained on the features extracted through each model are combined in a late fusion manner, where PSO and GA based optimization techniques are employed to assign weights to each classifier. The next subsections provide details of these steps.

Feature Extraction and Classification

The motivation for using pre-trained CNN models as feature descriptors comes from the fact that building a custom deep model from scratch requires heavy computation resources as well as large amount of data. To this aim, we mainly used five different CNN models form three different deep architectures, namely Alexnet [20], Resnet-50 [21] and vggnet19 [20]. In the case of AlexNet and VggNet, we used both versions i.e., the one pre-trained on ImageNet and the other pre-trained on places dataset. By using models pre-trained on ImageNet, we try to capture object-level features whereas the Places dataset is useful to extract scene-level features. These pre-trained deep models are only used as feature descriptors without any fine-tuning, and features are extracted from the last fully connected layers of each model. Once the features have been extracted, individual SVMs classifiers are trained on features extracted through each model. The selection of SVMs for classification purposes is based on its proven good performance for similar applications [22].

Late Fusion with PSO and GA Algorithms for Weight Optimization

In order to improve the classification accuracy obtained with individual models, we use two different types of weight optimization techniques in late fusion where classification scores obtained with the five models are combined using Equation 1. Here, S_c is the final score while $w(n)$ and p_n represent the weight (optimized via GA and PSO based techniques) and the probability vectors obtained with the classifier trained on the features extracted with n^{th} model, respectively.

$$S_c = w(1) * p_1 + w(2) * p_2 + ... + w(n) * p_n \tag{1}$$

The use of PSO and GA based optimization techniques for the weights to be assigned to each model is motivated by our previous experiences in event [23] and food recognition [24] domains.

Experiments and Results

In this section, we provide description of the dataset used for evaluation, the experimental setup, the conducted experiments, and a detailed analysis of the experimental results and comparisons against state-of-the-art.

Dataset

For the evaluation of the proposed solutions for identification of passable roads in flooded regions, we used the dataset provided for MediaEval 2018 challenge on Mulimedia and Satellite task

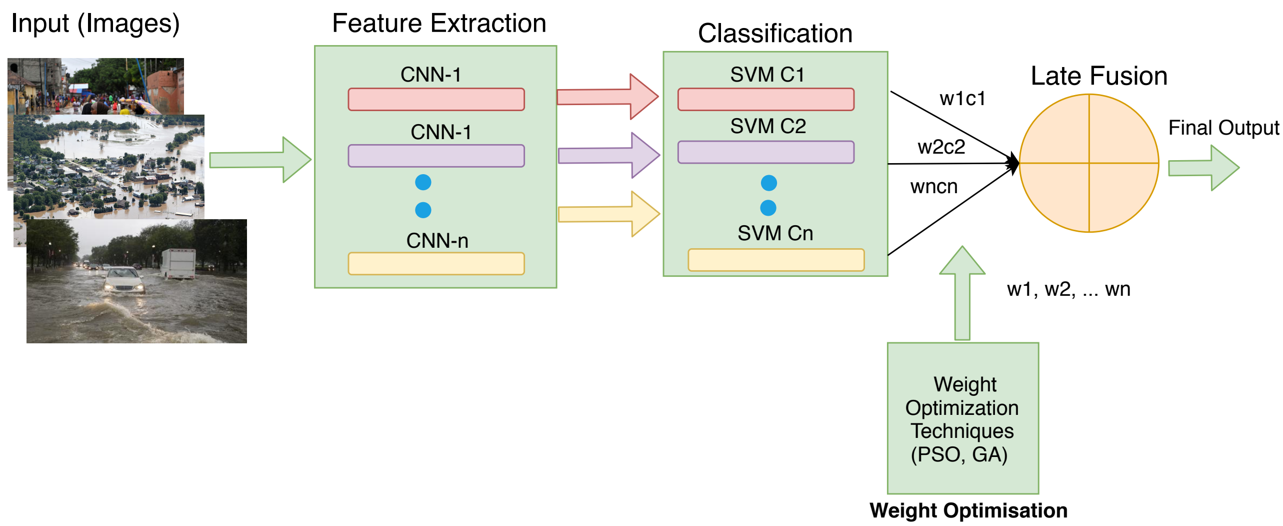

Figure 1. *Block diagram of the proposed methodology.*

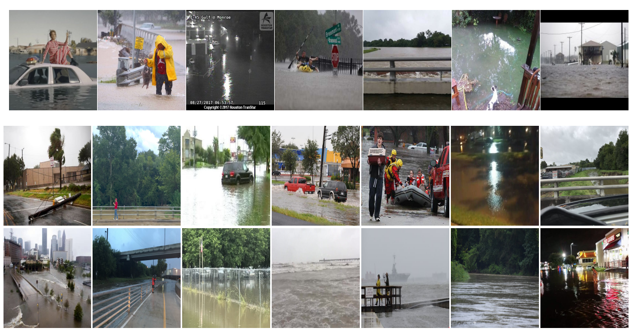

Figure 2. *Sample images from the dataset.*

[2]. The dataset provides large collections of images for both sub-tasks in terms of development and test sets containing 7,387 and 3,683 tweets and associated images, respectively. Ground-truths are provided for both tasks, Separately. Figure 2 provides some sample images from the dataset.

Experimental Setup

The objectives of the work is manifold. On one side, we want to evaluate the performance of the individual deep architectures. On the other-side, we want to analyze the contribution of object and scene-level features in both ways: individually and when jointly utilized. We also want to analyze the performances of different optimization techniques for assigning weights to the classifiers trained on features extracted through each model. In order to achieve these goals, we performed two different experiments. In the first experiment, we analyze the performance of the individual models while in the second experiment, these models are combined in two different late fusion techniques.

It is important to mention that during the experiments, we kept same experimental setup (i.e., SVMs parameters, validation set and number of iterations for calculation of the error rate in the fitness functions etc.,).

Experiments and Results

Table provides the results of our first experiment where we analyzed the performance of the individual models on both sub-tasks. Generally the accuracy is on the higher side for the first task (i.e., evidence) showing less complexity for the task compared to the second task (i.e., passability).

Models	Performance (accuracy in %)	
	Evidence sub-task	Passability sub-task
AlexNet (places)	88.29	64.71
VggNet (places)	88.43	65.00
AlexNet (ImageNet)	87.12	65.67
VggNet (ImageNet)	87.43	65.29
ResNet (ImageNet)	88.43	67.57

PSO Based Fusion 76.00

Considering the contributions of the individual architectures/models, overall better results have been observed for ResNet-50 on both tasks compared to its counterparts AlexNet and VggNet. As far as the contribution/importance of object and scene-level features is concerned, mixed response has been observed. On evidence task, models pre-trained on places dataset performed slightly better compared to when pre-trained on ImageNet. However, opposite trend has been observed on the passability task, though there is no significant difference in the performances.

Models	Performance (accuracy in %)	
	Evidence sub-task	Passability sub-task
GA based Fusion	91.86	76.88
PSO based Fusion	91.71	76.00

Evaluation of the weight optimization techniques in terms of accuracy.

Table 1 provides the results of our fusion techniques in terms of accuracy on both tasks. A significant improvement can be observed in the results compared to the performance of individual models on both tasks showing the efficacy of joint use of deep models for the tasks. As far as the evaluation of the weight optimization techniques is concerned, there is no significant differences in the performance of both techniques with slight improvement for GA based fusion on both tasks.

We also provide comparison against baselines [25] in Table 2. Our first baseline 1 is based on early fusion while the second one treats the models equally in a late fusion fashion. As can be seen, our merit based fusion techniques has better performances on the evidence task while comparative results have been obtained on the passability task.

Models	Performance (accuracy in %)	
	Evidence sub-task	Passability sub-task
Our method 1 (GA)	91.86	76.88
Our Method 2 (PSO)	91.71	76.00
Baseline 1 [25]	88.81	76.7
Baseline 1 [25]	90.36	76.0

Comparisons against baseline methods.

Conclusion

In this paper, we assess whether a road is passable or not in flooded regions after flood evidences have been found in the region. We also assess the performances of several deep models both individually and when jointly utilized. We also tried to analyze the importance of object and scene-level features. During the experiments, we found scene-level information more useful as compared to object-level information in classification of a scene as flooded or not flooded. However, to detect whether a road is passable or not, object-level information have a slight advantage over scene-level features. We also used PSO and GA optimization techniques to find optimal weights for fusion of different models. We found these optimization techniques useful in assigning merit based weights to models in the fusion process leading to an improvement in the classification accuracy compared to baseline of simply averaging the classification score as well as early fusion of the extracted features.

References

[1] N. Said, K. Ahmad, M. Riegler, *et al.*, “Natural disasters detection in social media and satellite imagery: a survey,” *Multimedia Tools and Applications* **78**, 31267–31302 (2019).

[2] B. Bischke, P. Helber, Z. Zhao, *et al.*, “The multimedia satellite task at mediaeval 2018: Emergency response for flooding events,” in *2018 Working Notes Proceedings of the MediaEval Workshop, MediaEval 2018*, 1–3, CEUR-WS. org (2018).

[3] J. Deng, W. Dong, R. Socher, *et al.*, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*, 248–255, Ieee (2009).

[4] K. He, X. Zhang, S. Ren, *et al.*, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778 (2016).

[5] M. Ullah and F. A. Cheikh, “Deep feature based end-to-end transportation network for multi-target tracking,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 3738–3742, IEEE (2018).

[6] K. Pogorelov, O. Ostroukhova, A. Petlund, *et al.*, “Deep learning and handcrafted feature based approaches for automatic detection of angiectasia,” in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 365–368, IEEE (2018).

[7] K. Courtney *et al.*, “The use of social media in healthcare: organizational, clinical, and patient perspectives,” *Enabling health and healthcare through ICT: available, tailored and closer* **183**, 244 (2013).

[8] C. Greenhow and C. Lewin, “Social media and education: Reconceptualizing the boundaries of formal and informal learning,” *Learning, media and technology* **41**(1), 6–30 (2016).

[9] M. Hamilton, F. Salim, E. Cheng, *et al.*, “Transafe: a crowdsourced mobile platform for crime and safety perception management,” in *2011 IEEE International Symposium on Technology and Society (ISTAS)*, 1–6, IEEE (2011).

[10] K. Ahmad, K. Pogorelov, M. Riegler, *et al.*, “Social media and satellites,” *Multimedia Tools and Applications* **78**(3), 2837–2875 (2019).

[11] H. Shekhar and S. Setty, “Disaster analysis through tweets,” in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1719–1723, IEEE (2015).

[12] D. Murthy and S. A. Longwell, “Twitter and disasters: The uses of twitter during the 2010 pakistan floods,” *Information, Communication & Society* **16**(6), 837–855 (2013).

[13] B. Takahashi, E. C. Tandoc Jr, and C. Carmichael, “Communicating on twitter during a disaster: An analysis of tweets during typhoon haiyan in the philippines,” *Computers in Human Behavior* **50**, 392–398 (2015).

[14] B. Truong, C. Caragea, A. Squicciarini, *et al.*, “Identifying valuable information from twitter during natural disasters,” *Proceedings of the American Society for Information Science and Technology* **51**(1), 1–4 (2014).

[15] S. Cresci, A. Cimino, F. Dell’Orletta, *et al.*, “Crisis mapping during natural disasters via text analysis of social media messages,” in *International Conference on Web Information Systems Engineering*, 250–258, Springer (2015).

[16] K. Ahmad, A. Sohail, N. Conci, *et al.*, “A comparative study of global and deep features for the analysis of user-generated natural disaster related images,” in *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, 1–5, IEEE (2018).

[17] B. Bischke, P. Bhardwaj, A. Gautam, *et al.*, “Detection of flooding events in social multimedia and satellite imagery using deep neural networks,” in *Working Notes Proceedings MediaEval Workshop*, 2 (2017).

[18] F. Alam, F. Ofli, and M. Imran, “Processing social media images by combining human and machine computing during crises,” *International Journal of Human–Computer Interaction* **34**(4), 311–327 (2018).

[19] B. Bischke, P. Helber, C. Schulze, *et al.*, “The multimedia satellite task at mediaeval 2017: Emergence response for flooding events,” in *Proceedings of the MediaEval 2017 Workshop (Sept. 13-15, 2017). Dublin, Ireland*, (2017).

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 1097–1105 (2012).

[21] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint*

arXiv:1409.1556 (2014).

[22] K. Ahmad and N. Conci, "How deep features have improved event recognition in multimedia: A survey," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **15**(2), 39 (2019).

[23] K. Ahmad, M. L. Mekhalfi, N. Conci, *et al.*, "Ensemble of deep models for event recognition," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **14**(2), 51 (2018).

[24] S. Khan, K. Ahmad, T. Ahmad, *et al.*, "Food items detection and recognition via multiple deep models," *Journal of Electronic Imaging* **28**(1), 013020 (2019).

[25] K. Ahmad, K. Pogorelov, M. Riegler, *et al.*, "Automatic detection of passable roads after floods in remote sensed and social media data," *Signal Processing: Image Communication* **74**, 110–118 (2019).