CrisisFlow: Multimodal Representation Learning Workflow for Crisis Computing

Patrycja Krawczuk, Shubham Nagarkar, Ewa Deelman
{krawczuk, sinagark, deelman}usc.edu
Department of Computer Science, University of Southern California

Abstract

An increasing number of people use social media (SM) platforms like Twitter to report critical emergencies or disaster events. Multimodal data shared on these platforms often contain useful information about the scale of the event, victims, and infrastructure damage. The data can provide local authorities and humanitarian organizations with a big-picture understanding of the emergency (situational awareness). Moreover, it can be used to effectively and timely plan relief responses.

In our project, we aim to address the challenge of finding relevant information among the vast amount of published SM posts. Specifically, we use deep learning algorithms to produce embeddings that encode the informativeness of multimodal SM data in the context of disaster events. Our method improves the state-of-the-art performance on the informative vs. non-informative classification task for the CrisisMMD dataset. To ensure the reliability and scalability of our solution in real-world scenarios, we implement the resulting crisis computing workflow in the Pegasus Workflow Management System (WMS).

Motivation

Fig. 1. There are many real-world examples of the use of Twitter data to save lives during crisis events. The news articles report use of SM for targeted emergency responses during the 2011 Tohoku earthquake and tsunami in Japan, and the 2018 Hurricane Harvey in Huston [1].

Global warming exacerbates the frequency and scales of extreme weather events such as floods, hurricanes, wildfires, cyclones, and blizzards. Natural disasters require quick and targeted emergency responses to save human lives and mitigate infrastructure damage. Increasingly, SM platforms serve as indispensable tools to acquire data needed to locate victims and plan efficient emergency reliefs, especially in the early hours after a disaster's onset.

One of the biggest obstacles in the use of SM content for crisis response is handling the information overload. As millions of SM posts are published at any time, there is a need to develop a system capable of processing a massive amount of data in near real-time. Such a system should be able to filter through thousands of short, informal messages and large amounts of images to extract the SM posts that are credible and build situational awareness during a disaster event.

Objective

- To improve the accuracy of the existing deep learning methods on informative vs. non-informative classification tasks on the CrisisMMD [2, 3] dataset.
- To develop CrisisFlow, a computational workflow capable of processing and classifying thousands of multimodal SM posts.

Methodology

Supervised Contrastive Learning on Images

Contrastive learning methods learn embeddings by pulling the representations of similar images together and pushing dissimilar ones apart from each other in the latent space.

Sentence Embeddings on Texts

Sentence embeddings encode the semantics and structure of a sentence into a vector. We fine-tuned the DistilBERT [5] model to create objective-specific embeddings for our downstream classification task.

Multimodal Learning: Early Fusion

Early fusion is a technique that allows merging data from different sources and modalities (image, text) at a vector level. The image and text embeddings are concatenated and passed through a neural network that classifies the SM post as informative or non-informative.

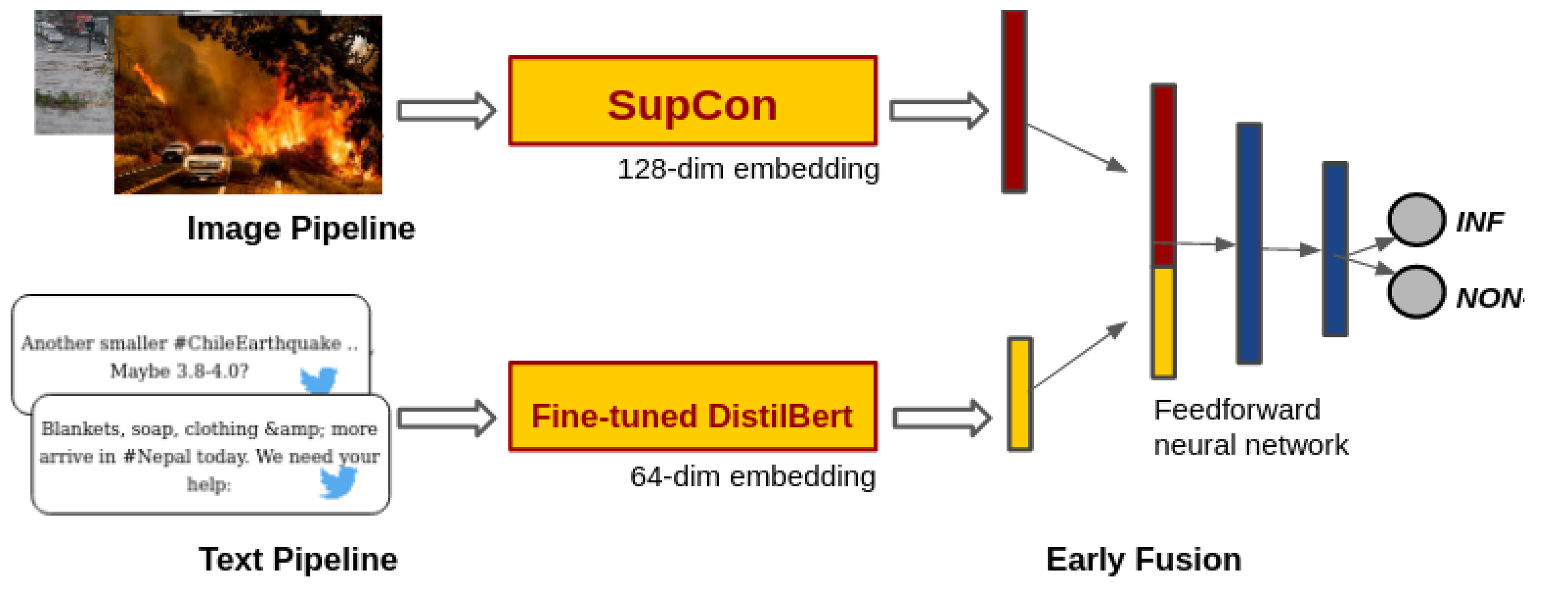

Fig. 2. Our architecture consists of two pipelines that ingest pictorial and textual parts of SM posts, respectively. The vector representations produced by the pipelines are then concatenated and fed into a neural network that classifies the post.

Experimental Results

Images: To produce meaningful, low-dimensional representations of the data, we train the SupCon [4] model for 150 epochs with a batch size of 4, SGD-Momentum optimizer and evaluate the quality of the representations with a kNN method.

Text: The DistilBERT is pre-trained on the GloVe Twitter 27 B embeddings. We fine-tune the model for 4 epochs with a batch size of 16 and Adam optimizer with a weight decay of 0.01 and a custom weighted loss function.

Images and Text: The image embeddings of size 1x128 and their corresponding text representations of size 1x64 are merged into a single vector. We train the network for 30 epochs with a learning rate of 0.001, batch size of 32, and Adam optimizer.

CrisisFlow: To ensure the reliability and scalability of our solution in real-world scenarios, we implement our method, CrisisFlow, in Pegasus WMS [6]. The directed acyclic graph (DAG) of the CrisisFlow where the nodes represent individual tasks while edges represent dependencies between them, is automatically mapped by pegasus on available resources and executed in parallel.

Table 1. Comparison of the results between Olfi et al. and CrisisFlow

Data Modality	Method	
	Olfi et al.	*CrisisFlow*
Text	0.808	**0.84**
Image	0.833	**0.89**
Image + Text	0.844	**0.91**

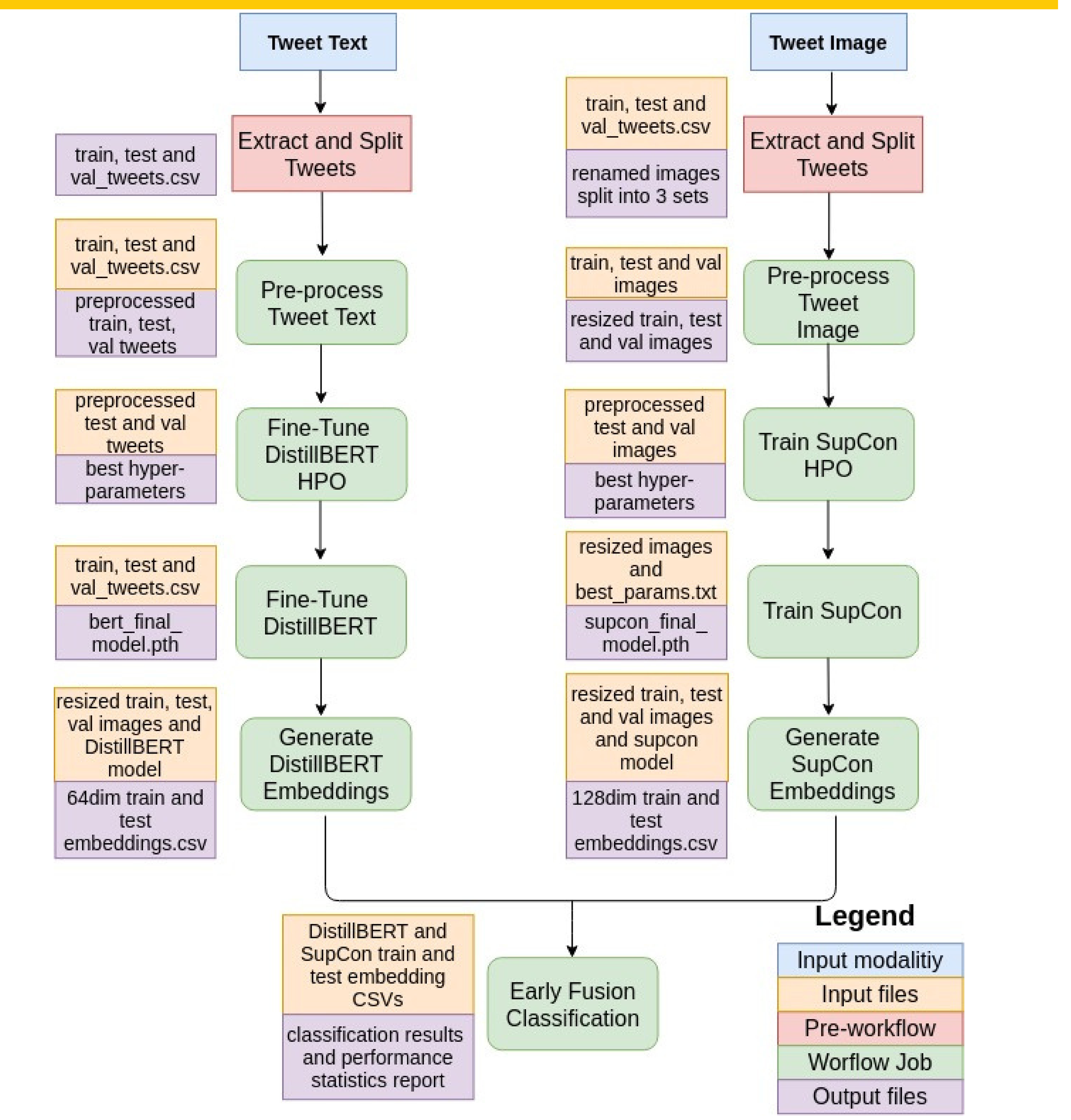

Fig. 3. The CrisisFlow graph describes job dependencies, input and output files.

Conclusion

The use of SM content presents an opportunity to provide timely and actionable information to first responders and local officials during a disaster event. However, there are many challenges before we can take full advantage of the data to build situational awareness during a crisis. One of them is the issue of information overload. In this work, we implement and train CrisisFlow, a scalable workflow that can process large amounts of multimodal SM data and classify them as informative or non-informative. In the future, we aim to generate a coherent summary of a disaster event and further improve the performance of our method.

Related Work

1. M. Koren, "Using Twitter To Save A Newborn From A Flood". TheAtlantic, 2017
2. F. Alam, F. Ofli, M. Imran, "CrisisMMD: Multimodal Twitter Datasetsfrom Natural Disasters". Proceedings of the 12th International AAAI Conference on Web and Social Media (ICWSM), 2018
3. F. Ofli, F. Alam, M. Imran, "Analysis of Social Media Data using Multimodal Deep Learning for Disaster Response". 17th International Conference on Information Systems for Crisis Response and Manage-ment (ISCRAM), 2020
4. P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A.Maschinot, C. Liu, D. Krishnan, "Supervised Contrastive Learning".Conference on Neural Information Processing Systems (NeurIPS), 2020
5. V. Sanh, L. Debut, J. Chaumond, T. Wolf, "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter", 2020
6. E. Deelman, K. Vahi, G. Juve, M. Rynge, S. Callaghan, P. Maechling, P. Mayani, W. Chen, R. Silva, M. Livny, K. Wenger, "Pegasus: a Workflow Management System for Science Automation". Future GenerationComputer Systems, 2015, pp. 17-35

Acknowledgements

This work is funded by NSF contract 1664162, "SI2-SSI:Pegasus: Automating Compute and Data Intensive Science"