Analyzing Social Media Texts and Images to Assess the Impact of Flash Floods in Cities

Bipendra Basnyat, P.E., Amrita Anam, Neha Singh, Aryya Gangopadhyay, Nirmalya Roy
Department of Information Systems
University of Maryland, Baltimore County
Email: {b125, amrita1, nehasingh1, gangopad, nroy}@umbc.edu

Abstract—Computer Vision and Image Processing are emerging research paradigms. The increasing popularity of social media, micro-blogging services and ubiquitous availability of high-resolution smartphone cameras with pervasive connectivity are propelling our digital footprints and cyber activities. Such online human footprints related with an event-of-interest, if mined appropriately, can provide meaningful information to analyze the current course and pre- and post-impact leading to the organizational planning of various real-time smart city applications. In this paper, we investigate the narrative (texts) and visual (images) components of Twitter feeds to improve the results of queries by exploiting the deep contexts of each data modality. We employ Latent Semantic Analysis (LSA)-based techniques to analyze the texts and Discrete Cosine Transformation (DCT) to analyze the images which help establish the cross-correlations between the textual and image dimensions of a query. While each of the data dimensions helps improve the results of a specific query on its own, the contributions from the dual modalities can potentially provide insights that are greater than what can be obtained from the individual modalities. We validate our proposed approach using real Twitter feeds from a recent devastating flash flood in Ellicott City near the University of Maryland campus. Our results show that the images and texts can be classified with 67% and 94% accuracies respectively.

Keywords—Social Sensors, Social Media Analytics, Computer Vision, Image Analysis

I. INTRODUCTION

Rapid advancement in mobile and ubiquitous computation have changed our way of information sharing. With the sophisticated phone in hand, the speed at which information are being shared across various social media is skyrocketing. Needless to say that, these new technologies have transformed our approach to personal and social aspects of our lives. We now ask "Google" if we need a raincoat or a jacket before leaving home; speak to the outer world via our Facebook status, tweets and Instagram post and image sharing.

The social media activity is, even more, prominent in the crisis situation. The phenomenon was apparent in the recent flash flood event that occurred in Ellicott City, a suburb old downtown in Howard County, Maryland. On Saturday, July 30th, 2016, torrential rainfall occurred around Ellicott City, resulting in the severe flash flood. The event led to two loss of lives, destruction to the historic city, and damages to hundreds of vehicles, restaurants and shopping stores. We posit this flash flood event as our testbed in this work and investigate the importance of Twitter to gather the contents and types of tweets which can be postulated and reciprocated with different data modalities during the crisis.

We envision the event and aftermath as a mini-disaster situation whose characters and the dire situations can be extrapolated to the longer and more severe natural disaster. We exemplify the Ellicott City flash flood event in this paper where people were enjoying their Sunday evening in this small downtown on the main street and an unprecedented storm swept through the area. According to the meteorologists, such massive rainfall was a rare 1-in-1,000-year event. The National Weather reported that the storm dumped 6.5 inches of rain on Ellicott City in only about 3 hours, with 5.5 inches falling in just 90 minutes. The storm killed two people and damaged many small scale commercial buildings and numerous vehicles parked on the road. Within minutes, the touristic street were turned into a raging river, carrying away cars and other debris and forcing dramatic rescues of people trapped in flood using human chain.

The surge in the storm and flooding outside trapped people in restaurants, thereby giving them an opportunity to share the event unfolding on real-time into many social media. Numerous real-time tweets/facebook with images and videos of the event "flooded" social media platform about the main street of Ellicott City. The commuters and visitors were acting just like the physical sensors on the ground, e.g., smartphone GPS sensors share location data, smartphone cameras provide real-time pictures of the unfolding event and human themselves share their personal observations on Twitter which help make data of multiple modalities available into the cyber world. The proliferation of sensors in the possession of the average individual, together with the popularity of social networks allow massive information dissemination and heralds an era of social sensing that brings about new research challenges and opportunities in this emerging field [4]. Motivated by this, in this work we investigate several computational techniques to perform independent analysis on different data modalities such as texts and images first and then postulate how the inferred knowledge from one modality can help boost the other. More specifically, we make the following research contributions.

- We investigate a Latent Semantic Analysis (LSA) technique to extract the meaningful features of textual data and cluster them based on the distance measures.
- We investigate a Discrete Cosine Transformation (DCT) technique to extract the relevant features of image data

and cluster them based on their 1-dimensional and 2-dimensional values of density functions.

- We establish the cross-correlation and relevancy metrics between the narrative (text) and visual (image) guided queries and postulate how a query on single data modality can be enhanced by extracting the deeper contextual meaning and percolating the semantic knowledge gain from the other relevant data modality.
- We perform the experiments and validate all our hypotheses on real data traces collected and processed from a life threatening flash flood event in a modern American city.

II. RELATED WORK

The text and image categorization in general is a well investigated research area, but only a few studies have been devoted to the environmental and smart city disaster management applications. The research on image analysis in this domain mostly focus on high resolution GIS specific images from advanced satellite for monitoring the different environmental contexts and situations. In recent time researchers have started to pay more attention to automatic categorization of medical images [8]. Twitter is a monumental social networking site with more than 313 million active users, which is heavily used for fast communication and broadcasting the message. There are various tools available for analyzing the Twitter data. Tweedr is a Twitter data analysis tool which extracts actionable information during natural disaster and inform relief workers using various machine learning techniques for classification and clustering [3]. TweetTracker is an analysis tool for humanitarian and disaster relief, capable of tracking, analyzing and monitoring the tweets in real time in order to get the situational awareness just after the disaster [7]. Twitcident is a web based framework for filtering, searching and analyzing the tweets about real-world incidents or crisis [1]. TweetXplorer is a system which uses little information about the event and effective visualization techniques to gain the better knowledge of an event [12].

Apart from the tools, various techniques have been employed for event detection on social media. [1], [13] propose the probabilistic methodology to monitor tweets and detect the target event using Kalman filter for location estimation. The authors in [9] apply Named Entity Recognizers successfully to extract locations from microblogs at geo-location and point-of-interest. [14] presents a novel algorithm for geo-spatial event detection that monitor all twitter posts of a given geographic area and recognize highly active places with high precision and recall. [4] provided an unsupervised algorithm for event detection by fusing multiple social network and integrates data from twitter and Instagram using expectation maximization approach that shows an improvement over the approach utilizing single social network data independently. [6] deals with characterization of micro-events via social sensing of image sharing and involving semantics of image content. They proposed EiM framework that uses generative model and statistical clustering based approach (ST+LDA) for micro-event detection. [2] presents a matrix factorization approach with ensemble based gradient descent algorithm to uncover the polarization in social network by separating the tweets of different polarities.

Many work has been done in data visualization using social media. [11] describes two tools for the unstructured tweets to be useful, TweetQL for programmers which provides SQL-like interface to the twitter API and TwitInfo shows interaction with end-users to make information more accessible. [5] proposes a novel statistical method and anomaly based method for event detection in twitter, which relies only on tweets and leverages the creation frequency of dynamic links. [10] proposes a novel visual analytics framework for twitter data consisting sentiment modelling and geo-graphic visualization to explore the underlying trends in positive and negative sentiments with respects to disaster and geographically related sentiments. [15] summarizes the evolution of combined event detection with visual analytics over 14 years and demonstrates an overview of the state-of-the-art methods. In this paper, we look into the multi-facet nature of the twitter feeds with respect to a specific query and analyze the knowledge gain from both text and image components independently and jointly.

III. METHODOLOGY

In this section we describe the goals and contributions of our proposed uni-modal and multi-modal approaches. Our main objective is to take advantage of using data from multiple modalities over single modalities in answering queries that cut across themes such that each theme is better expressed in a different data modality. Some examples of such queries are given below:
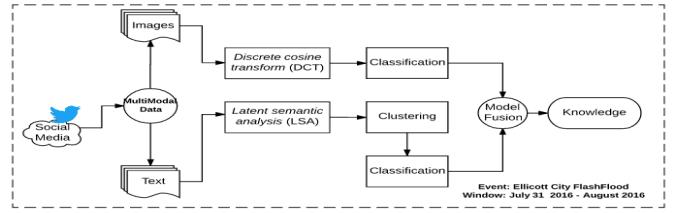
Fig. 1: Process Flow for Social Sensing Module

Q1: Find all tweets that are related to casualties about flash floods. In this query there are two themes: *casualties* and *flood*. Casualties are better described by text that depict the number of casualties, description of people affected, etc., whereas the flooding situation is better captured by images. Therefore, combining texts and images can convey the results more effectively than each by itself.

Q2: Find all tweets that describe damages caused in the aftermath of flash floods. In this query the damages can be visually described by images far better than texts. However, the details of the damage such as quantitative measures are better described in texts.

Q3: Find tweets related to business losses caused by flash floods. In queries such as this damages of stores and commercial properties can be graphically depicted by images whereas the extent of losses due to closures,

Fig. 2: Text Cluster

Fig. 3: Word Clouds w.r.t Cluster c0 help/rescue

Fig. 4: Word Clouds w.r.t Cluster c1 damage/aftermath

Fig. 5: Word Clouds w.r.t Cluster c2 casualty
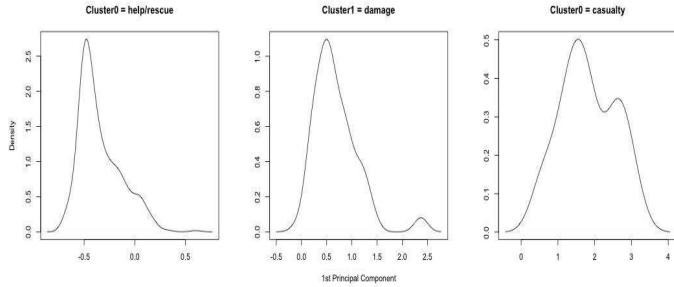
access problems, and losses of properties and inventory are better described in texts.

Fig. 1 depicts the model that we have devised to separate out the multimodal data (texts and images) into their own category and our implementation of machine learning algorithms to extract the knowledge that help to answer the above queries. Next, we present our text and image analysis techniques.

A. Tweet Text Analysis

The purpose of analyzing tweets on a natural disaster is to find the context of the conversation during and after the disaster has taken place. We applied unsupervised techniques on 596 tweets about Ellicott City flash flood to first learn about the context to categorize them and then train a model to classify unseen tweets. This process was followed by two steps: i) cluster the tweets to identify categories, and ii) train a model to classify unseen tweets to fit one of the categories.

1) Clustering: In the first step we used Latent Semantic Analysis to find tweet similarity. We used tf-idf weighing scheme to vectorize the tweets. We accepted terms that appeared in at least two tweets and ignored the terms that appeared in more than half of the tweets. Next, we applied Singular Value Decomposition for dimensionality reduction following the Equation 1.

$$X_k = U_k \sum_k V_k{}^T \tag{1}$$

We took the top k =150 singular values, which explained 90% of the variance. We computed a tweet-tweet cosine similarity matrix from the right-singular matrix, V. After that, we performed k-means clustering on the 1-cosine similarity matrix to group the tweets based on their similarity with other tweets. We identified three categories by analyzing terms with the highest tf-idf in the clusters. The three categories as we found are "help/rescue", "damage/aftermath" and "casualty". Fig. 2 shows the density plots of the first principal components of the *help*, *damage*, and *casualty* clusters, respectively.

2) Classification: After labeling all the tweets with thee classes, we separated them into training set (542 tweets) and test set (53 tweets). We vectorized the tweets by both tf-idf and SVD (k =100) to feed as our inputs. For classification, we used K nearest neighbor (KNN) algorithm on the vetorized tweets. We used k = 3 as our number of neighbors need for voting, brute force as the search algorithm and cosine similarity as the
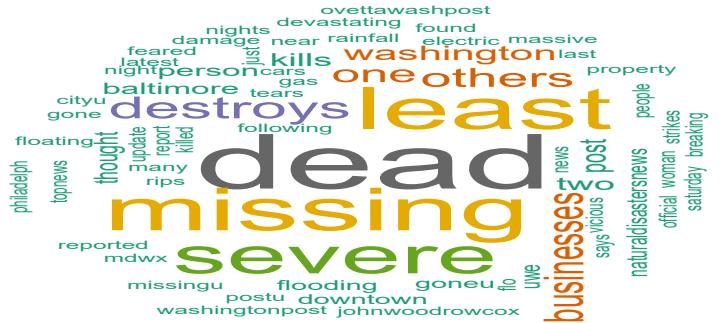
similarity metric. We observed an overall accuracy of 86.79% from tf-idf vectors and 92.45% from SVD. Our classifier can detect "help/rescue" with 92.6%, "damage/aftermath" with 90% and "casualty" with 100% accuracy.

Figs. 3, 4, 5 present the cluster of words from the tweets. We visualized the clusters using the first two principal components and word clouds generated from the tf-idf weights of each cluster. The red cluster (c0) comprises of majority of the tweets. By looking at the word cloud, which highlights the terms with the highest tf-idf weights, we selected the theme of the cluster c0 to be "help/rescue" as shown in Fig. 3. Note that the green (c1) and the blue (c2) clusters are smaller in size with more closely related texts as shown respectively in Fig. 4 and Fig. 5. Again with the help of the word cloud we came up with the themes to be "damage/aftermath" for c1 and "casualty" for c2.

TABLE I: 1D DCT values (Mean & SD) with and without Outliers

1D DCT	1D DCT values with Outliers			1D DCT values without Outliers		
	No Water	Others	Water	No Water	Others	Water
Mean	344	388	229	374	510	159
SD	137	190	110	106	132	61

B. Image Analysis

Our objective is to train a model that could classify a previously unseen image into one of three classes: flooding, damage in the aftermath of a flood, and other desiderata such as reports, news items, commercial for flood related products etc. The first step in this process is feature selection that will lead to the highest accuracy of classification for the images. We have developed two different approaches for automated analysis of the images, both based on the Discrete Cosine Transform (DCT). DCT transforms a signal vector or matrix into the frequency domain such that most of the energy is captured by a small number of elements in the transformed domain. DCT was chosen in favor of other techniques such as the Fast Fourier Transform (FFT) because it has strong energy compaction and all of the coefficients in the real number domain make it computationally easy to use and quantize. The two methods we referred both include one and two dimensional DCTs. In the 1D case, we vectorized each image before the transformation. In the 2D case, we applied the standard technique of analyzing the image using 8×8 blocks. Since each image has a large number of pixels and only a few of the coefficients capture most of the energy we kept the top k coefficients to capture 90% or more of the energy as shown in Equation 2, where D_i is the transformed value for the ith element and N is the number of elements in the image vector.

$$D_i = c_i \sum_{x=0}^{N-1} p_i \cos \frac{(2x+1)i\pi}{2N}$$
$$\text{where } c_i = \begin{cases} \sqrt{1/N} & \text{if } i = 0 \\ \sqrt{2/N} & \text{if } i > 0 \end{cases} \quad (2)$$

The images were first converted to grayscale and vectorized and then transformed using DCT. Both 1D and 2D DCT were performed on the images. About 97% of the energy were preserved in the first three DCT coefficients. The means of the DCTs were well separated for the three classes. However, the standard deviations were large indicating the fact that the DCT values for several images in each class overlapped with those of the other classes as shown in Table I. The pseudocode of our proposed image analysis technique is shown in Algorithm 1.

C. Machine Learning

In this step we take the DCT coefficients obtained from the image analyses and train a machine learning model that can classify a previously unseen image into one of three classes: images that describe flooding conditions, images that describe damages, presumably in the aftermath of the floods, and images that are related to flash floods but are not direct

Input: a corpus with p images;
Output: list of top k DCT coefficients for each image;
for $i \leftarrow 1$ **to** p **do**
    Convert image i to grayscale;
    Preform DCT on image i;
    Select the k largest coefficients such that 90% of energy is preserved;
end

Algorithm 1: Analyzing Images

images of either flooding or damages. These three classes are referred to as *water*, *nowater*, and *others* respectively in our subsequent analysis. We created two different machine learning models using the decision tree and Naive Bayesian algorithms respectively. The classification results are discussed in Section V.

IV. DATA COLLECTION

Social sensing can be based either on the physical sensors that one uses to transmit their data such as GPS, accelerometer etc., from their phone or the meta-information via text messages, social media posting and micro-blogging etc. Such data can be post processed to create intelligent applications. Regardless of the types of the sensor (physical or meta), the process flow for development of an information abstraction model can be be generalized by the process flow diagram as shown in Figure 1. We abide by same process flow for our data collection.

First step in the data collection process was to sense them which included creating a filter criteria for Twitter. The search phrase "Ellicott City Flash Flood" was used as a driver to our "Social Sensing Module" to find tweets related to the event. We used advanced search criteria of Tweeter to limit the Tweets time range from July 30 to August 1 2016. We wrote a custom software application that extracted the data from Twitter and stored in MySQL database. Since the Tweets are highly unstructured data, we parsed the Tweets and provided some structures to the Tweets data. The data was then stored in our local database for further analysis. Our research aimed at creating a multimodal data fused predictive model from the dataset, thus we had to integrate information extracted from images with its textual counter part. To that effect, the tweets with images and without the images were separated. Table II below depicts the dataset we had for the event window.

TABLE II: Total Tweets

Tweets with Images	152
Tweets without Images	477
Total	648

As a part of pre-processing, we separated the tweets that had image associated with them from the ones that did not. There were many tweets without images and also retweets. For our analysis, we lead the image as a control point. In other words, Tweets without images were excluded as shown in Table II. Our approach aims to establish correlation of the text content

Fig. 6: Images and Density Functions
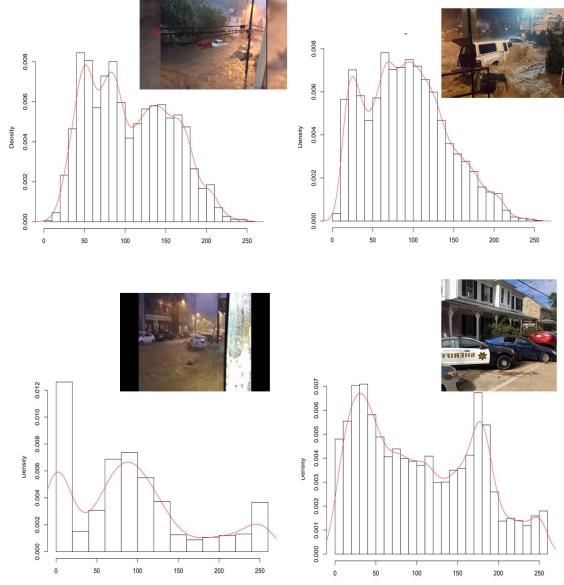
Fig. 7: Images and Density Functions

with the image counterpart in a Tweet. Therefore, special care was taken not to loose the connection between Tweets and their images. We use ROWID as an identifier to make this connection such that there would be one to one connection between image and tweet text.

V. Experimental Results

We employed word frequency analysis, which provides a list of words that occur in a text and the number of times they occur. We used the information to validate the contents of tweets and how specific are they to the event in general. We plotted the top-12 words, where first two words are about the Ellicott City. As seen from the Top-12 list word, one easily can make a meaningful contextual sense from it. First step in analyzing the text content of the Twitter text was to find the most frequent words amongst collected datasets (see Figs. 3, 4, 5 for the word clouds for three distinct clusters of tweets). However, the image analysis step shows completely different patterns than the tweets. There are three distinct patterns that are apparent in the set of images, e.g., those during the floods, those of the damages caused by the floods, and others that are not directly descriptive of but related to the flooding and damages.

Figure 6 and Figure 7 show the histograms and probability density functions (PDFs) of three sample images that are representative of the three themes mentioned above: *water*, *nowater*, and *others*. While the PDFs show distinct patterns, there are significant overlaps between PDFs of images in the categories of *water* and *nowater* while images in the category *others* are somewhat distinct. We validated this conjecture using mixture models based on the Gaussian kernel. Hence we
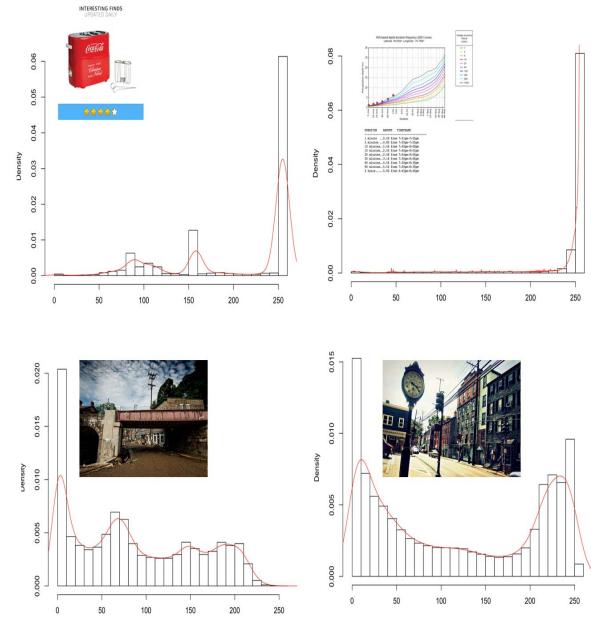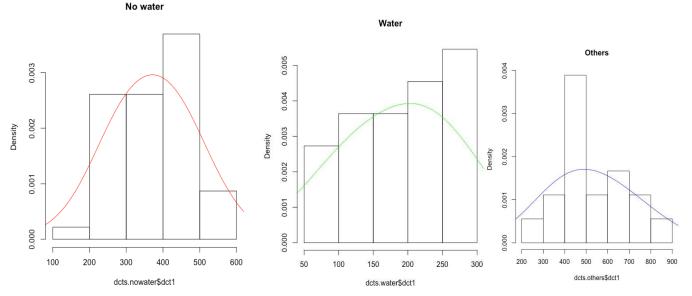

Fig. 8: Density Functions for 1D-DCTs
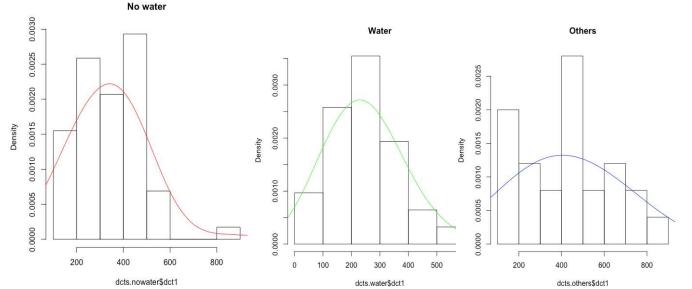
Fig. 9: Density Functions for 2D-DCTs

developed a training model based on DCTs for automatically classifying the images into the three classes mentioned above. We developed two different DCT-based models: 1D-DCT of the gray-scaled images after vectorization and 2D-DCT of the gray-scaled images. While for most of the images the top 5 DCTs captured more than 90% of the energy, Figures 8 and 9 show the PDFs for the first DCT coefficient, which had by far the largest absolute values. Furthermore Figures

TABLE III: Classification Results using DCTs with and without Outliers

| Classifiers	With Outliers		Without Outliers	
	1D DCT	2D DCT	1D DCT	2D DCT
Decision Trees	27%	43%	65%	63%
Naive Bayes	52%	65%	67%	66%
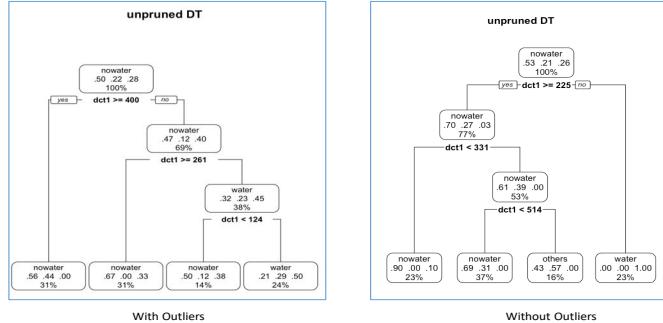
Fig. 10: Decision Tree Pruning with and without Outliers

8 and 9 show that the means of the density functions of the first DCT coefficient are well-separated with images of the class *water* having the smallest mean values and those of class *others* having the largest mean values, with those of the class *nowater* in the middle. However, there is a wide variance in the values for all three classes. This suggests that there are images in each class whose DCT values overlap with those in the other two classes. Hence we expect that a classifier trained on the DCT values of the images will have poor classification accuracy. One of the main reasons for this is that images in the class *water* include objects such as dry land, cars, humans, and other objects that are commonly found in the images of class *nowater*. However, images in the class *others* are relatively easier to separate from those in the other two classes. This intuition was tested and confirmed by training classifiers using decision trees and naive Bayesian algorithms. After removing some of the images in each of the three classes we found that the classification accuracies improved significantly. Both classification models were trained using 50% samples of the data and tested against the other 50% holdout samples. The classification results are shown in Table III. Note that Naive Bayes approach achieves 67% and 66% image classification accuracy respectively using 1D and 2D DCT vectors on twitter images. The decision trees for the classifiers using the 1D DCT are shown in Figure 10. Both trees are unpruned but are relatively small trees. Also, as can be seen, the trees are predominantly based only on the first DCT coefficient.

VI. Conclusion

In this paper we investigated how the social media feeds such as twitter can be analyzed using single as well as multiple data modalities. We showed that while the independent analysis of texts and images from the consistent twitter feeds has high correlation the presence of out-of-context information makes the separation and inference process non-trivial. Our proposed approach achieves $\approx 67\%$ accuracy for classifying

the images and $\approx 94\%$ accuracy for classifying the texts based on the real twitter traces from a recent flash flood event. We plan to investigate a bipartite graph based approach in future to develop a more robust model for establishing the mutual correlations and meaningful mappings between our text and image analysis processes. We believe our approach in this work and the preliminary results would help build predictive data analytic tols for different mission critical human-in-loop cyber-physical systems applications in smart cities.

References

[1] F. Abel, C. Hauff, G.-J. Houben, R. Stronkman, and K. Tao. Twitcident: fighting fire with information from social web streams. In *Proceedings of the 21st International Conference on World Wide Web*, pages 305–308. ACM, 2012.

[2] M. T. A. Amin, C. Aggarwal, S. Yao, T. Abdelzaher, and L. Kaplan. Unveiling polarization in social networks: A matrix factorization approach. Technical report, IEEE, 2017.

[3] Z. Ashktorab, C. Brown, M. Nandi, and A. Culotta. Tweedr: Mining twitter to inform disaster response.

[4] P. Giridhar, T. Abdelzaher, and L. Kaplan. Social fusion: Integrating twitter and instagram for event monitoring. 2017.

[5] A. Guille and C. Favre. Event detection, tracking, and visualization in twitter: a mention-anomaly-based approach. *Social Network Analysis and Mining*, 5(1):1–18, 2015.

[6] K. Jayarajah and A. Misra. Can instagram posts help characterize urban micro-events? In *Information Fusion (FUSION), 2016 19th International Conference on*, pages 130–137. IEEE, 2016.

[7] S. Kumar, G. Barbier, M. A. Abbasi, and H. Liu. Tweettracker: An analysis tool for humanitarian and disaster relief. 2011.

[8] T. M. Lehmann, M. O. Güld, T. Deselaers, D. Keysers, H. Schubert, K. Spitzer, H. Ney, and B. B. Wein. Automatic categorization of medical images for content-based retrieval and data mining. *Computerized Medical Imaging and Graphics*, 29(2):143–155, 2005.

[9] J. Lingad, S. Karimi, and J. Yin. Location extraction from disaster-related microblogs. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 1017–1020. ACM, 2013.

[10] Y. Lu, X. Hu, F. Wang, S. Kumar, H. Liu, and R. Maciejewski. Visualizing social media sentiment in disaster scenarios. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1211–1215. ACM, 2015.

[11] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller. Processing and visualizing the data in tweets. *ACM SIGMOD Record*, 40(4):21–27, 2012.

[12] F. Morstatter, S. Kumar, H. Liu, and R. Maciejewski. Understanding twitter data with tweetxplorer. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1482–1485. ACM, 2013.

[13] T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web*, pages 851–860. ACM, 2010.

[14] M. Walther and M. Kaisser. Geo-spatial event detection in the twitter stream. In *European Conference on Information Retrieval*, pages 356–367. Springer, 2013.

[15] F. Wanner, A. Stoffel, D. Jäckle, B. C. Kwon, A. Weiler, D. A. Keim, K. E. Isaacs, A. Giménez, I. Jusufi, T. Gamblin, et al. State-of-the-art report of visual analysis for event detection in text data streams. In *Computer Graphics Forum*, volume 33. Citeseer, 2014.