Taming Big Remote Sensing Data to Aid First Responders in Damage Assessment in a Post-Flood Scenario

Joshua Ortner
Department of Computer Science
Middle Tennessee State University
Murfreesboro, Tennessee USA
jo3f@mtmail.mtsu.edu

Kuldeep Kurte
Computational Sciences and Engineering Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee USA
kurtekr@ornl.gov

***Abstract*—The availability of high quality imaging technology, an increased presence of orbital satellites, and the innovations associated with unmanned aerial vehicles have led to a spike in the quantity and quality of remote sensing (RS) imagery. RS data is useful in a variety of applications, from scientific research to financial forecasting. One such important use of RS data is humanitarian assistance and disaster response (HADR). Many of the tasks involved in HADR such as resource allocation, aid routing, and rescue and recovery require an accurate and rapid assessment of the damage of an affected region. On-site damage assessment is accurate, but it can be dangerous and very time consuming. Access to high quality aerial images of a region recently affected by a disaster has become fairly common, however due to the size of these datasets, it can be a major challenge to discover images that are useful to the first responders. Straightforward processes like storage and search can become too time consuming to be useful. New techniques for storage, image information mining, image querying, and image classification are needed. The goal of this project is to explore the use of state-of-the-art computer vision and artificial intelligence algorithms in the task of rapid ingestion and analysis of post-disaster aerial imagery. We present a multi-tier example based image querying system. The system utilizes a pre-trained convolutional neural network to extract embedded features from images. The feature space is partitioned using a *K*-dimensional tree to allow for a quick search of the space to find images with features similar to the example provided by the user. We discuss the full details of the application's implementation as well as its shortcomings, present preliminary results when applied to a data set of post-flood images, and discuss potential future directions to improve its performance.**

***Index Terms*—Image Retrieval, Convolutional Neural Networks, Transfer Learning, Remote Sensing, Humanitarian and Disaster Response.**

I. Introduction

Big data has become a colloquial term in nearly every domain that participates in automated digital data acquisition. Thanks to modern innovations and the availability of computers, sensors, and networks, acquiring data is the easy part. Much of the research focus has shifted to developing techniques to wrangle the data and extract high quality information from it in a timely manner [1].

Remote Sensing (RS) data is one domain that has exploded in size over the last decade thanks to advances in satellite, imaging, and unmanned aerial vehicle (UAV) technology. Extracting useful information from the raw pixels of images has proven to be a serious challenge [2]. However, advances in Computer Vision algorithms and the field of Artificial Intelligence (AI) has given researchers powerful tools to begin tackling this challenge. Since RS data can be so valuable and potentially save lives in a post disaster scenario, there is a serious amount of interest in the research and development of novel techniques that can be used to extract information from this data [3].

This research project was focused on exploring and extending the current state-of-the-art computer vision and AI algorithms for use in humanitarian assistance and disaster response. We explore the utility of transfer learning and deep convolutional neural networks in the task of extracting high level features from post-flood aerial imagery. We demonstrate the potential applicability of these algorithms in a post-disaster scenario by designing and implementing an application prototype which allows first responders to use these algorithms to process and analyze their data.

This paper is organized in the following way: Section II provides the implementation details of the AI for Humanitarian Disaster Response Image Retrieval (AI4HDR-IR) application. Next, Section III describes the data used during the project. Further, Section IV present the results of the main feature of the application, i.e. an example based image query system. Finally, Section V concludes the paper and discusses potential future directions for this project.

II. AI4HDR-IR Framework

Figure 1 shows the framework of the AI4HDR-IR system. It is a multi-tier application consisting of three main layers: a web based user interface, an image processing and analysis pipeline, and data storage. Its designed to give first responders the tools required to sift through large quantities of aerial imagery, finding what is most important for their work. Here, we demonstrate the workflow of a hypothetical user to showcase a

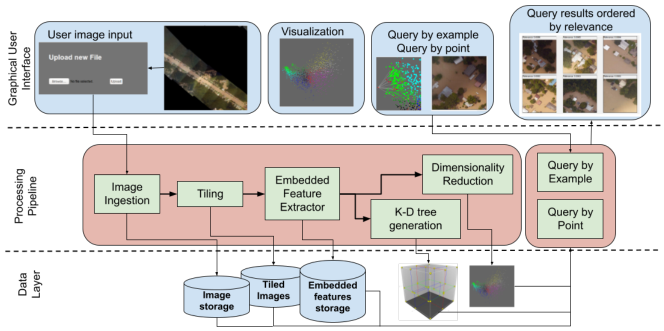

Fig. 1. Diagram of the AI4HDR-IR system.

potential use case for the system, and then describe, in detail, the implementation of the system itself.

A. A Potential User

An important first step in responding to a natural disaster is data collection. Often this is in the form of aerial images collected via satellites or unmanned aerial vehicles. Once data has been collected, a user of the AI4HDR-IR system will launch the application and upload their data through the provided interface. This ensures that images are ingested properly by the system and stored in an appropriate format. Next, the user can navigate to the data exploration page. On the exploration page, the user is presented with an interactive three-dimensional network visualization, where nodes represent an image in the database, edges represent query results and the width of the edges corresponds to the degree of similarity between two connected nodes: wider edges mean the nodes are more similar than thinner edges. The user can zoom in and out and rotate the space to get a better view of the network and images of interest. From here the user can click on a node, triggering the query functionality. The image corresponding to the clicked point will appear on the right side of the interface along with the top ten images that were determined to contain similar content. Nodes will be added to the 3D environment for each image that was returned in the query. As a future feature, the user will be able to mark relevant images (i.e. a relevance feedback mechanism), building a training set to train a classifier which can further aid in extracting images from the database that are important for their needs.

B. Implementation

The first major process in the system is image ingestion. RS images collected by a particular sensor have the same image characteristics, such as spatial and spectral resolutions, but that isn't necessarily the case across the images acquired by other types of sensors. For example, the dimension and resolution of drone images captured in the aftermath of a hurricane in North America may differ drastically from satellite images of a flood in Asia. Accommodating multi-sensor imagery is a challenging task. In this work we focused only on images acquired by UAV. More information on the data set is provided in Section III. After image ingestion, each ingested image is decomposed using an image tiling process that takes an image as input and then produces non-overlapping, uniform slices of that image. Currently, slices with more than three percent of their pixels blank will not be stored or moved onto subsequent stages of the pipeline. This percentage can be altered though, based on the requirements of job. This tiling component of the pipeline is motivated from the fact that the user of the system is interested only in a specific region of the big RS image to be retrieved.

The next step is feature extraction. Feature extraction in this pipeline takes advantage of a method from AI known as transfer learning. Its been shown that the early layers of a convolutional neural network have the ability to learn how to detect the presence of general features such as edges and shapes. This means that the early layers of a model trained in one domain, may be re-purposed for use in another domain, decreasing the amount of time and training samples required to achieve a reasonable accuracy. This is typically implemented by initializing a model with pre-trained weights and replacing the final classification layer with one thats designed for the new problem. Because of the constraints present following a natural disaster, first responders don't typically have access to large training sets or the time required to train deep neural networks.

The feature extraction mechanism of the system is powered by a pre-trained convolutional neural network know as VGG16 [4]. To construct the feature extractor, the final classification layer of the VGG16 network is removed, and the other layers are initialized with weights trained on the ImageNet data set [5]. The features are produced from the output of the new final layer which is a 4096 element feature vector. In this way, the embedded feature vectors are extracted from each of the images and saved in the database for further processing.

A key and challenging component to the visualization of the embedded feature space of the data is dimensionality reduction. Since each of the image feature vector contain 4096 elements, they are impossible to visualize directly. The major downside with any data reduction method is the loss of information. In this case, however, it isn't a concern since the reduced dimension feature vectors are used only for visualization purposes. All other analytical procedures are performed on the full sized feature vectors. We use Principal Component Analysis (PCA) to produce three-dimensional feature vectors. PCA is one of the most popular methods of dimensionality reduction and aims to preserve the information that is most important in defining the given samples [6].

The mechanism powering the example based query process is a K-dimensional tree (KDTree). A KDTree is a binary search tree that's generated from points in a multidimensional space. We use the Euclidean distance between embedded feature vectors to quantify image similarity [7]. A KDTree is designed to represent this kind of relationship, and provides a search complexity of O(log N).

III. Dataset used

The data set used in the development and testing of this system is post-flood imagery collected by the coordinated effort of the Department of Commerce, the National Oceanic and Atmospheric Administration, the National Ocean Service, and the National Geodetic Survey, Remote Sensing Division. The images are of the Louisiana flooding of 2016, collected at an altitude of 5000 feet, with ground sample distance for each pixel of 50 cm. The dimensions of the images are 8015 x 8015. For the purposes of this project, we chose to slice the images into 256 x 256 sized tiles, removing images that contained mostly blank pixels [8].

IV. Results

To evaluate the performance of the system we picked four image content categories: buildings and vegetation (both with and without standing water present). Figure 2 shows an example of a query from each category. This was to ensure that the results demonstrate the applicability of the system in a post disaster scenario. In the case of flood assessment, it may be important to estimate how many buildings are flooded and the degree of that flooding so that regions of the affected area can be prioritized.

With these categories in mind, we selected twenty samples of each category from a database of roughly 18,000 post-flood images. Then we used the samples to query the same database for the top five and ten most relevant images. To quantify the performance of a query we use precision:

$$\frac{TruePositive}{TruePositive + FalsePositive} \text{..........}(1)$$

Fig. 2. A query example of each category showing the top three results: Buildings with no standing water present, buildings with standing water present, vegetation with no standing water present, and vegetation with standing water present (From left to right).

The overall performance of a category is aggregated using the harmonic mean of each query. Finally, the system's total performance is calculated using the harmonic mean of the aggregates. The overall precision of the system is 73 percent and 66 percent using the top five and top ten most relevant images respectively. A detailed breakdown of the performance can be seen in Table I.

TABLE I
Image retrieval performance of four selected queries

Image Description	**Precision Harmonic Mean**	
	P@5	**P@10**
Buildings (No Standing Water)	0.99	0.91
Buildings (Standing Water)	0.55	0.57
Vegetation (No Standing Water)	0.92	0.85
Vegetation (Standing Water)	0.63	0.50
Total System Precision	0.73	0.66

The major conclusion of the results is that the performance of the query system decreases as the queries become more specific and complex. For example, querying the system for images that contain both buildings and standing water, using an image that contains both, under performed compared to a query for images containing only buildings, and no standing water.

It isn't surprising that more complex queries present a greater challenge, especially considering the feature extractor was trained on non-disaster related images. A feature extractor trained on disaster related images may produce higher quality features. The performance could also be related to the distance metric used by the KDTree, which is Euclidean Distance. Other distance metrics such as Manhattan Distance may be more suitable for use in high dimensional spaces. Also, a KDTree may not be the best structure to use for high dimensional data, methods such as hash based indexing may be more appropriate. We discuss these potential paths further in the following section.

V. Conclusion and Future Work

This paper described the application of feature extraction and classification to the task of post-flood damage assessment. Also presented was a system prototype utilizing these techniques that allows first responders to generate a database of images collected from a disaster affected region, and provides them with tools to explore and better understand their data by enabling quick detection of images of interest. The results show a precision of 73% (P@5) and 66% (P@10) when used to query a database of 18,000 post-flood images.

In the future, we plan to test a few techniques to improve the performance of the query system for more complex queries. First we will replace the KDTree as the central querying mechanism with a hash based lookup system [9]. This method may be more sensitive to similarities between the elements of feature vectors and thus may lead to a higher accuracy. Also, a hash based approach would have the positive side effect of faster lookup times which will ensure continued utility even as the size of the database increases. Second, we plan to initialize the weights of a custom built auto encoder with the weights from a pre-trained VGG16 model and then train the auto encoder on the data provided to the system. The motivation behind this is that by training an auto encoder to reproduce the images in the users data set, it will learn features that are more representative of the user's data (such as buildings, roads, and the presence of flood water) instead of the features representing the images from ImageNet. Third, because of the computational requirements of large scale data, we plan to assess the performance of a hash based approach to querying as well as a KDTree with alternative distance metrics. Finally, we would like to add an active learning feature to the system, giving the user the ability to create classification models which can be trained to further hone the query process.

Another potential feature to add is semantic segmentation [10]. Semantic segmentation is the pixel-wise classification of objects in an image. It results in a new image where objects are filled in using unique colors based on their classification. For example, if a model was trained to find cats in an image, the result may be a new image where the form of the cat is colored blue, and all other pixels are black. This method could be very useful in a post-disaster scenario thanks to the high degree of detail, however it tends to be very data hungry and computationally intensive. One potential way of tackling this may be to consider the combination of transfer learning to create a feature extractor and a random forest algorithm to do the pixel-wise classification [11].

ACKNOWLEDGMENT

This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI) program at Oak Ridge National Laboratory, administered by the Oak Ridge Institute for Science and Education.

REFERENCES

[1] Y. Li, J. Ma, and Y. Zhang, "Image retrieval from remote sensing big data: A survey," *Information Fusion*, vol. 67, pp. 94–115, 2021.

[2] J. Zhang, W. Hsu, and M. L. Lee, "Image mining: Issues, frameworks and techniques," in *Proceedings of the 2nd ACM SIGKDD International Workshop on Multimedia Data Mining (MDM/KDD'01)*. University of Alberta, 2001.

[3] M. Rahnemoonfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari, and R. Murphy, "Floodnet: A high resolution aerial imagery dataset for post flood scene understanding," *arXiv preprint arXiv:2012.02951*, 2020.

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[5] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[6] R. Bro and A. K. Smilde, "Principal component analysis," *Analytical methods*, vol. 6, no. 9, pp. 2812–2831, 2014.

[7] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Communications of the ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[8] Department of Commerce (DOC), National Oceanic and Atmospheric Administration (NOAA), National Ocean Service (NOS), National-Geodetic Survey (NGS), Remote Sensing Division.

[9] B. Demir and L. Bruzzone, "Hashing-based scalable remote sensing image search and retrieval in large archives," *IEEE transactions on geoscience and remote sensing*, vol. 54, no. 2, pp. 892–904, 2015.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[11] "158b - transfer learning using cnn (vgg16) as feature extractor and random forest classifier," Sep 2020. [Online]. Available: https://www.youtube.com/watch?v=IuoEiemAuIY