

CSCE 585: Machine Learning Systems: Project Proposal: Computer Vision Assisted Disaster Rescue

Cole Lewis @colelewis, David Duggan @davidduggan54, Daniella Mallari @dmallari

September 27, 2022

Contents

1	Project Repository	2
2	Problem	2
3	Contextual Work	2
4	Data	3
5	Methodology	3
6	Evaluation	3
7	References	3

1 Project Repository

<https://github.com/csce585-mlsystems/CVDisasterRescue>

2 Problem

The problem we are investigating is that of conducting search-and-rescue operations in the wake of disaster in such a way that the most victims possible can be rescued while the risk to the rescue team is minimized. The authors are seeking to develop a scene recognition model that will take data from real-time computer vision and microphone input to help locate survivors in disaster zones. These data streams would be used in real-time detection of both human anatomy within an environment, as well as distinguishing speech from other noise that may be present. Our model can then be deployed on an autonomous or remotely operated device. This would lessen the strain on human rescuers by improving the efficiency of locating possible survivors in disaster areas where visibility is low and other risk factors are prevalent. Our project takes the approach of drawing from data streams such as camera and microphone input in order to simulate human sensory faculties in making inferences whether or not there are any victims in the presence of the robot.

3 Contextual Work

The application of machine learning to disaster management, mitigation, and prediction is well documented and prolific. This includes predicting disasters before they occur, rapidly managing disasters as they occur [2], and handling disasters after the fact [1]. With the modern advent of this technology, the capability to prepare for and process calamitous events has been greatly aided, and will continue to prove useful now and into the future [3]. Autonomous robots are already widely used in disaster areas with machine learning models. One paper in particular outlines how one such autonomous deployment utilizes behavior trees to determine action from a range of data sources. [5] Our system could easily be deployed on an existing similar robot or drone. In particular, Our system would act as the perception and inference mechanism for the robot to detect humans in need of help after the disaster occurs, and when being operated remotely, the system gives feedback to the operator to assist with inference in ways an autonomous system may fall short absent human supervision. While voice detection has been developed and implemented for use in disaster zones, existing approaches [6] opt to determine what is being said rather than whether or not anybody is speaking, which our system does.

4 Data

The project data is being taken from Google's Open Images Version 6 dataset. We plan to use a subset of the entire dataset because the original size is 1.7 million images. A dataset like this would make model training very time consuming. This means that running experiments to optimize model hyperparameters and test multiple models would be infeasible. The Open Images dataset is accessible using the FiftyOne tool, which allows a user to choose the data that they download. Data can be downloaded by: class, types of annotations, amount of data, and which split to download.

5 Methodology

Our idea uses a human-in-the-loop perceptual system where a rescuer will be watching the camera feed and our system will assist in detection of humans in need of rescue. Our model will employ a convolutional neural network architected for object detection that is fed images of people and their specific parts (arms, legs, hands, etc.) in order to train it to identify any visible part amidst rubble and other obstacles. In addition to finding human parts, the system will use movement detection as a secondary sense to help it locate humans. In addition to visual detection, our system will utilize voice detection via filtering microphone input into a trained voice detection model.

6 Evaluation

We plan to use multiple different evaluation metrics to evaluate the performance of our model. The first is Intersection over Union (IoU), which evaluates the overlap between the ground truth and the predicted bounding box. A perfect overlap is represented by an IOU score of 1, and no overlap gives an IoU score of 0. Using IoU, we can create a threshold for deciding whether the model made a correct prediction or not [4]. This will allow us to use additional metrics such as precision and recall to determine the types of errors our model is making. In our problem, Type II error is far more dangerous than Type I error. A Type II error would cause our model to predict that there is no one in the video frame, despite there actually being someone. This could cause that person to not be saved by the machine. Precision and Recall could be extended to mean average precision (mAP), which would allow us to map the average precision across the different classes using the area under the curve for Precision and Recall.

7 References

[1] GFDRR. 2018. Machine Learning for Disaster Risk Management. Washington, DC: GFDRR. https://www.gfdr.org/sites/default/files/publication/181222_WorldBank_DisasterRiskManagement.pdf

- [2] Ochoa, K. S., & Comes, T. (2021). A machine learning approach for rapid disaster response based ... - arxiv. Retrieved September 5, 2022, from <https://arxiv.org/pdf/2108.00887.pdf>
- [3] Guikema, S. (2019, October 09). Why machine learning is critical for disaster response. Retrieved September 05, 2022, from <https://blogs.scientificamerican.com/observations/why-machine-learning-is-critical-for-disaster-response/>
- [4] Koech, K. E. (2022, March 28). On object detection metrics with worked example. Retrieved September 05, 2022, from <https://towardsdatascience.com/on-object-detection-metrics-with-worked-example-216f173ed31e>
- [5] Rosas, Francisco, et al. "Autonomous Robot Exploration and Measuring in Disaster Scenarios Using Behavior Trees." 2020 IEEE 10th International Conference on Intelligent Systems (IS), 2020, <https://doi.org/10.1109/is48319.2020.9199973>.
- [6] F. Alifani, T. W. Purboyo and C. Setianingsih, "Implementation of Voice Recognition in Disaster Victim Detection Using Hidden Markov Model (HMM) Method," 2019 International Seminar on Intelligent Technology and Its Applications (ISITIA), 2019, pp. 445-450, doi: 10.1109/ISITIA.2019.8937290.