U.S. Department of Transportation
Federal Railroad Administration

RESEARCH RESULTS
FRA OFFICE OF RESEARCH, DEVELOPMENT & TECHNOLOGY

RR 22-04 | February 2022

Railroad Trespass Detection Using Deep Learning-Based Computer Vision

SUMMARY

The U.S. Department of Transportation (US DOT) John A. Volpe National Transportation Systems Center (Volpe), under the direction of the Federal Railroad Administration (FRA) Office of Research, Development, and Technology (RD&T), developed an Artificial Intelligence (AI) software application for automating the detection of grade crossing violations and trespass activities from static camera video feeds. Volpe researchers conducted the work from 2020 to 2021. The Grade Crossing Trespass Detection (GTCD) software application outputs predicted grade crossing violations and right-of-way (ROW) trespassing as tabular data in MS Excel format along with annotated video files of trespass events. An example of video processing output showing pedestrian and vehicles traversing a grade crossing in Ramsey, NJ, during an activation is shown in Figure 1.

Accurately detecting when a trespass event occurs using standard video input reduces the time needed to collect safety data. Currently, railroads and many state DOTs have a wealth of video data on their systems, but that data is generally only analyzed if there is a documented incident. Automated identification and processing of trespass events from the existing video data may yield significant safety data currently not being analyzed. This software application is available for download at https://public.huddle.com/b/jPDLGE/index.html.

BACKGROUND

Fatalities resulting from grade crossing collisions or ROW trespassing account for approximately 95 percent of all rail-related fatalities over the past 10 years. In 2020, 1,901 grade crossing incidents resulted in 197 fatalities and 688 injuries; and there were also 525 trespass fatalities and 557 injuries at non-crossing locations [1].

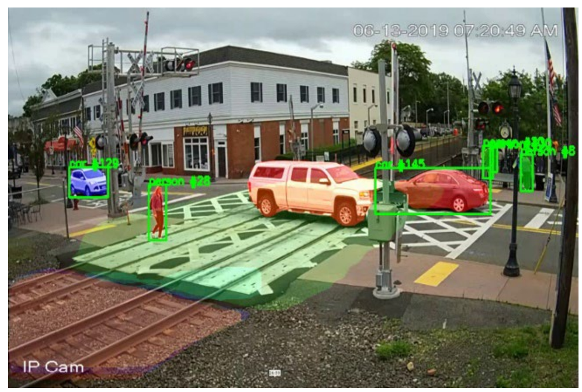

Figure 1. Example of Video Processing Output Showing Pedestrian and Vehicles Traversing a Grade Crossing in Ramsey, NJ

Analysis of past trespass incidents is one way railroads and Federal, State, and local agencies select trespass mitigation locations. However, this approach omits most railroad ROW locations from consideration for trespass mitigation since they have not yet experienced an accident. A 2018 FRA Report to Congress, “National Strategy to Prevent Trespassing on Railroad Property,” identified data gathering and analysis as one of four strategies to prevent trespasser incidents [2]. This, in itself, is not enough. The key element of this strategy is identification of *new data sources*. Railroads and many state DOTs have vast networks of closed-circuit televisions that monitor their infrastructure. The video data from these systems contain a wealth of information regarding trespass activities but are generally not analyzed due to resources needed to manually review the video data.

In response to this recommendation, FRA RD&T supported the research and development of an AI tool to process video data for identifying trespassers at grade crossings and along railroad ROWs.

OBJECTIVES

The main objective of this project was to develop a deep learning-based computer vision tool for automating detection of grade crossing violations and railroad ROW trespass activities from static camera video feeds.

METHODS

Figure 2 shows the data processing overview of the GTCD software application to identify rail trespass events.

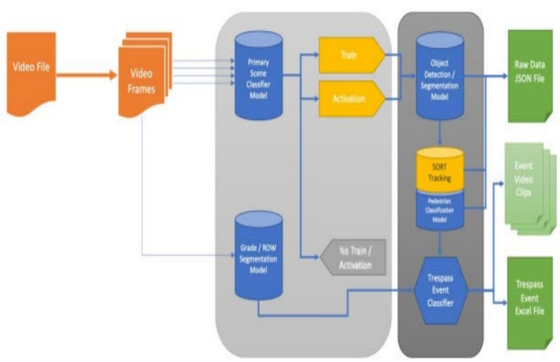

Figure 2. GCTD Software Application Data Processing Overview

The following steps provide an overview of the algorithm:

<u>Step 1 (Parsing the video data)</u>: The event detection process begins with extracting each frame of video along with video timestamp data and then passing it through the computer vision processing pipeline.

<u>Step 2 (Grade crossing and ROW segmentation mask)</u>: This step defines the region of interest (ROI) for grade crossing and railroad ROW. Previous approaches required a predefined ROI, limiting the utility of the solution to locations where those regions had been manually defined. The solution presented here instead defines ROIs using a computer vision model trained on a sufficient corpus of data, annotated for both rail grade crossing and ROW regions. The Mask R-CNN [3] architecture, combined with a ResNet50 feature extraction backbone, was selected to be the initiation point for the active learning process to create the grade/ROW segmentation model.

<u>Step 3 (Scene classification)</u>: The next step is the identification of a grade crossing activation within the video stream that is required for a violation to occur. SqueezeNet image classification was selected as the network architecture for the scene classification component of the GCTD application architecture.

<u>Step 4 (Object detection)</u>: The Mask R-CNN architecture, combined with a ResNet50 feature extraction backbone, was trained using the Common Objects in Context (COCO) dataset to search and provide segmentation masks for the seven object types identified as the most relevant for this project: person, car, truck, bus, train, motorcycle, and bicycle.

<u>Step 5 (Trespass prediction)</u>: Once the predictions have been returned by the computer vision pipeline, the following post-processing steps generate final trespass event predictions: 1) object tracking data is processed using the Simple Online and Realtime Tracking (SORT) [4] algorithm; and 2) a calculation is then performed to test for segmentation mask overlap between the grade crossing/ROW segmentation masks and the segmentation mask for both pedestrians and vehicles.

This approach has the following benefits:

- It is fully generalized, meaning that it can be applied to nearly any rail grade crossing/ROW where video can be collected without requiring

additional training of the computer vision models.
- Data collected using this approach can also perform other rail grade safety activities, such as counting vehicles and pedestrians.

One drawback to this approach is the significantly higher computational requirements to process videos and/or real-time video feeds. In addition, the application is quantitatively less accurate when using significantly skewed camera angles or in poor visibility.

RESULTS

The software application accepts multiple video input file formats, and video codexes and can be executed via a command line or using the included graphical user interface (GUI). Figure 3 shows the screenshot of the GUI.

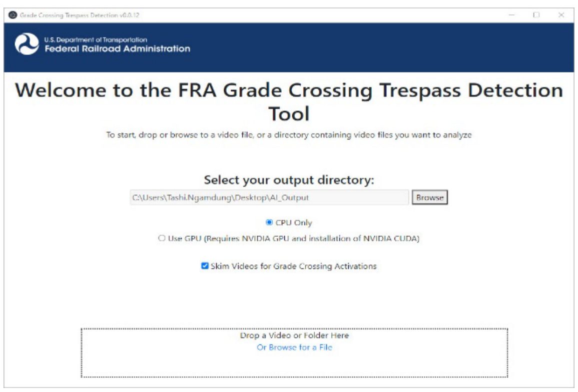

Figure 3. Screenshot of the GCTD Software Application Graphical User Interface

The accuracy of the computer vision models were measured using average label accuracy (for the grade activation model) and intersection-over-union for both precision and recall (for the segmentation model). Accuracy for the activation detection model was 92 percent on the validation set, and the segmentation model scored an average precision of 0.978 and an average recall of 0.967.

Application inference (prediction) performance was maintained across multiple videos with various framerates and resolutions by normalizing the input during image pre-processing to a standard resolution and framerate. This results in the ability for application to process full HD video at approximately 5 fps using a single graphics processing unit. Lower resolutions process only moderately faster, with a standard HD video yielding a 6 fps inference rate.

Trespass detection accuracy was measured across a subset of eight video clips for a known number of human-validated events, and where the event type distribution was considered representative of the larger target distribution. Quantitative accuracy for this subset was effectively 100 percent; all known and expected events were detected. However, a significant false positive rate of ~30 percent was incurred on some videos due to poor segmentation overlap detection between the ROI (grade or ROW) and the object (person or vehicle).

CONCLUSIONS

Currently, railroads and many state DOTs have a wealth of video data on their systems, but that data is generally only analyzed if there is a documented incident due to the human capital required to analyze it. Automated identification and processing of trespassing and crossing violations may yield significant safety data currently not being analyzed. The GTCD software application can automatically detect trespassers at grade crossings and along railroad ROWs from static camera video feeds. It outputs predicted trespassers as tabular data along with annotated video files of trespass events. The application is available at https://public.huddle.com/b/jPDLGE/index.html.

FUTURE ACTION

This project successfully demonstrated that a computer vision model can be developed and generalized across multiple locations. The following are potential next steps for this study:

- Add additional locations and training data to improve the segmentation model and trespass detection accuracy.

- Add visualization to 1) show trespasser source and destination, 2) detect and visualize environmental factors which play a contributing role in trespass events, and 3) show both pedestrian and vehicle pathing.
- The current GCTD system collects all the raw data required to generate activity metrics. However, it does not output those counts in tabular format. Update the output to include pedestrian and vehicle counts.
- The current algorithm is trained to detect trespassers from static video data. Update and train this model to detect trespassers from a dynamic video data collected from locomotive cameras.

REFERENCES

[1] FRA Office of Safety Analysis. 2.08-Highway-Rail Crossings. 4.08-Casualty Summary Table. Accessed October 2021.

[2] FRA. (October 2018). National Strategy to Prevent Trespassing on Railroad Property.

[3] He, K., Gkioxari, G., Dollár, P., and Girshick, R. (Januray 2018). Mask R-CNN.

[4] Bewlwy, A., Ge, Z., Ott, L., Ramos, F., and Upcroft, B. (July 2017). Simple Online and Realtime Tracking.

ACKNOWLEDGMENTS

This work was performed under interagency agreements between the RD&T Train Control and Communication Division and the Volpe Center's Systems Safety and Engineering Division.

CONTACT

Francesco Bedini Jacobini
Program Manager
Federal Railroad Administration
Office of Research, Development, and Technology
1200 New Jersey Avenue, SE
Washington, DC 20590
(202) 493-0800
Francesco.Bedini@dot.gov

Tashi Ngamdung
Operations Research Analyst
Volpe National Transportation System Center
Systems Safety and Engineering Division
55 Broadway, Cambridge, MA 02142
(617) 494-2937
Tashi.Ngamdung@dot.gov

KEYWORDS

Automated railroad trespass detection, artificial intelligence, AI, machine learning, railroad trespassing, grade crossing violations, railroad safety

CONTRACT NUMBER

DTFR5317X00026