February 2024

Report No. 23-049

Maura Healey
Governor

Kim Driscoll
Lieutenant Governor

Monica Tibbits-Nutt
MassDOT Secretary & CEO

Artificial Intelligence Framework for Crosswalk Detection Across Massachusetts

Principal Investigator (s)
Dr. Jimi Oke
Dr. Chengbo Ai
Dr. Yuanchang Xie

University of Massachusetts Amherst
University of Massachusetts Lowell


Research and Technology Transfer Section
MassDOT Office of Transportation Planning

U.S. Department of Transportation
Federal Highway Administration

Technical Report Document Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
20-024	n/a	n/a

4. Title and Subtitle	5. Report Date
Artificial Intelligence Framework for Crosswalk Detection across Massachusetts	
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Atanas Apostolov, Yuanchang Xie, Chengbo Ai, Francis Tainter, Jimi Oke	20-049
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
University of Massachusetts Amherst 130 Natural Resources Road, Amherst, MA 01003 University of Massachusetts Lowell 1 University Avenue, Lowell, MA 01854	n/a 11. Contract or Grant No.
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Massachusetts Department of Transportation Office of Transportation Planning Ten Park Plaza, Suite 4150, Boston, MA 02116	Final Report - February 2024 [May 2023 - February 2024]
	14. Sponsoring Agency Code n/a

15. Supplementary Notes
Project Manager - Bonnie Polin, MassDOT

16. Abstract
The goal of this project was to develop an artificial intelligence (AI) framework to detect crosswalk locations across the state of Massachusetts, as well as their type classification (continental, parallel lines, or solid) and location category (intersection, midblock, or driveway). Aerial images downloaded from MassGIS were annotated, and these were used to train the AI model, which was then used to detect crosswalks across the entire state using images from both 2019 and 2021. About 88,000 crosswalks were detected in 2021. In terms of location category, 89% of these were intersection crosswalks, 8% were midblock, and 3% were at driveways. In terms of type, just under 65% were continental (zebra-style) crosswalks, about 35% were parallel lines, and less than 1% were solid/painted crosswalks. A python script was developed to perform the post-processing. The model and the post-processing script will be available to MassDOT to perform further analyses, and results will be expected to inform maintenance and safety initiatives.

17. Key Word	18. Distribution Statement
crosswalk, artificial intelligence, safety	unrestricted

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
unclassified	unclassified	41	n/a

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized.

This page left blank intentionally.

Artificial Intelligence Framework for Crosswalk Detection across Massachusetts

Final Report

Prepared By:

Atanas Apostolov
Graduate Research Assistant

Francis Tainter
Co-Principal Investigator

Chengbo Ai
Co-Principal Investigator

Yuanchang Xie
Co-Principal Investigator

Jimi Oke
Principal Investigator

University of Massachusetts Lowell
1 University Avenue, Lowell, MA 01854

University of Massachusetts Amherst
130 Natural Resources Road, Amherst, MA 01003

Prepared For:
Massachusetts Department of Transportation
Office of Transportation Planning
Ten Park Plaza, Suite 4150
Boston, MA 02116

February 2024

This page left blank intentionally.

Acknowledgments

Prepared in cooperation with the Massachusetts Department of Transportation, Office of Transportation Planning, and the United States Department of Transportation, Federal Highway Administration.

The Project Team would like to acknowledge the efforts of all state transportation agencies and metropolitan planning organizations who participated in the interviews and shared information with the research team.

Disclaimer

This page left blank intentionally.

Executive Summary

This study of Artificial Intelligence Framework for Crosswalk Detection across Massachusetts was undertaken as part of the Massachusetts Department of Transportation (MassDOT) Research Program. This program is funded with Federal Highway Administration (FHWA) State Planning and Research (SPR) funds. Through this program, applied research is conducted on topics of importance to the Commonwealth of Massachusetts transportation agencies.

Methodology

Using annotated aerial imagery from 2019 and 2021 across Massachusetts, an artificial intelligence model was trained to detect crosswalks. The results were post-processed to further categorize the crosswalk location (intersection, midblock, driveway) and remove false positives. Crosswalks were then compiled as a GIS layer for delivery to MassDOT.

Results

Based on our analysis, there were 88,440 crosswalks detected in Massachusetts in 2021 and 83,380 crosswalks detected in 2019. Of these, 89% were intersection crosswalks, whereas about 8% were midblocks. The remaining 3% were driveways. These proportions were similar in both years. In terms of type, continental ("zebra") crosswalks accounted for 62–64%, standard/parallel line crosswalks comprised 36–38%, and solid/painted crosswalks accounted for the remainder (<1%).

This page left blank intentionally.

Table of Contents

This page left blank intentionally.

List of Tables

List of Figures

This page left blank intentionally.

List of Acronyms

Acronym	Expansion
AI	Artificial Intelligence
DSC	Dice Similarity Coefficient
ESRI	Environmental Systems Research Institute
GIS	Geographic Information System
MassGIS	Massachusetts Bureau of Geographic Information
MassDOT	Massachusetts Department of Transportation
ODC	Object Detection and Classification
SSL	Subsegment Length

This page left blank intentionally.

1.0 Introduction

Crosswalks are critical parts of our roadway infrastructure. Knowing the locations of crosswalks is important for prioritizing our systemic countermeasure program to enhance pedestrian safety. MassDOT has already developed the risk models for pedestrian safety but does not have a complete inventory of statewide crosswalk locations. Artificial Intelligence (AI) has the potential to automatically detect crosswalks from satellite images and quickly generate the crosswalk inventory data needed by MassDOT. This automatic detection can be far more efficient than manual identification. Also, such an AI tool allows MassDOT to update the crosswalk inventory frequently when new satellite images are available. Given new satellite images, this tool can also be used to conduct change detection and identify crosswalks that need to be repainted.

This project has the following objectives:

- Develop an AI model to detect crosswalks across Massachusetts.
- Use Python scripts to clean up the AI model detection outputs and prepare a clean copy of the detected crosswalks in a polygon feature class within a geodatabase.

This page left blank intentionally

2.0 Data Summary

We downloaded aerial imagery from the Massachusetts Bureau of Geographic Information (MassGIS) website. These images were from both 2019 and 2021. Crosswalks in 44 tiles were annotated for (1) zebra/continental crosswalks, (2) crosswalks marked by two parallel solid lines; and (3) raised crosswalks without pavement markings.

2.1 Annotated Tiles

The 2021 annotated tiles are shown in Table 1. In total, 44 tiles were annotated.

Table 1. Annotated tiles

Tile	Number of Crosswalks Annotated	Tile	Number of Crosswalks Annotated
18TXN960190	39	19TCH165010	88
18TYN005175	39	19TCG240935	382
19TBG655830	55	19TCG195920	236
19TCG255920	471	19TCG105860	97
19TCG285890	441	18TYM005905	5
19TCG165920	87	19TCG180890	99
19TBG880920	30	19TCG090860	46
19TBG865905	60	19TCG210950	178
19TCG270920	531	19TCG150950	28
19TCG300965	392	19TCG255950	350
19TCG330785	116	19TBG685830	300
19TCG120860	68	19TCH270175	1
19TBG895680	28	18TYM035935	121
18TYM140830	28	19TCG285725	41
18TYN200175	64	18TXM990635	262
19TBG565950	25	19TCG165830	35
19TBG640920	44	19TCG300905	472
18TYN275190	40	18TXM930815	40
19TBH670160	9	18TXM960740	324
19TCH345010	59	19TBG640800	45
19TCH360025	88	19TCH285370	179
19TCG195980	112	19TCH285385	37

Figure 1 shows two of the tiles used in annotation.

Figure 1. Example tiles

2.2 Annotated Crosswalks

A summary of the crosswalks annotated is in Table 2. In total, 6,192 crosswalks were annotated: 5,157 crosswalks were used to train the original model, and 666 of the remaining 1,038 crosswalks were used for validation. Examples of annotated crosswalks are shown in Figure 2.

Table 2. Summary of annotated crosswalks used for training

Label	Description	Count	Percentage (%)
1	Zebra/continental/ladder	4,411	85.53
2	Parallel lines (solid or dashed)	615	11.93
3	Fully painted/solid	131	2.54

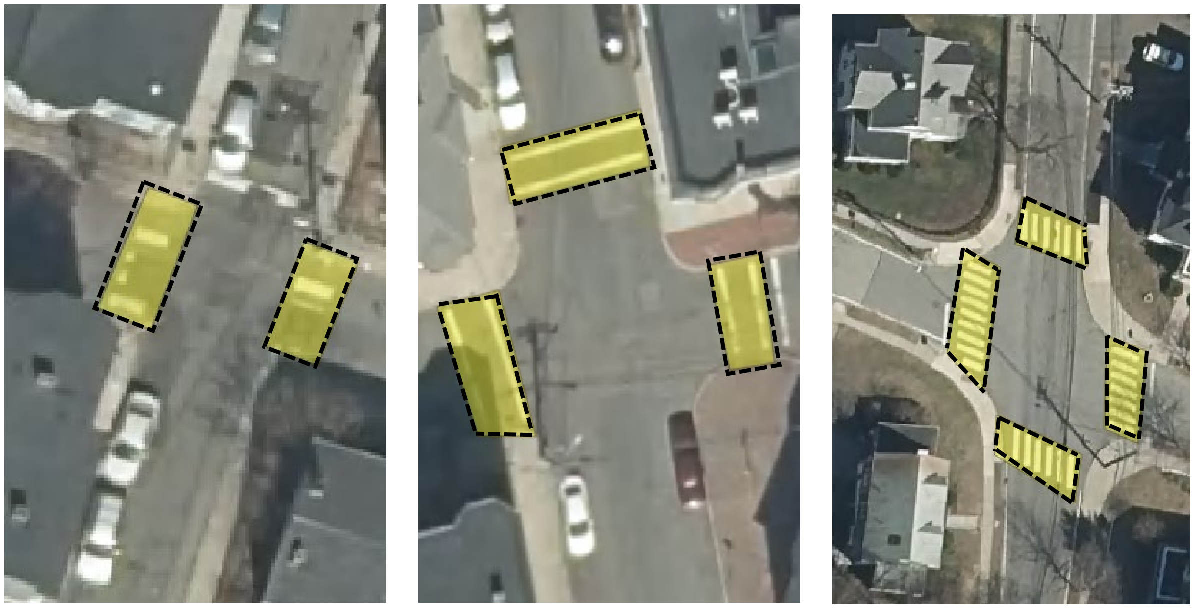


Figure 2. Annotated crosswalks

This page left blank intentionally.

3.0 Research Methodology

3.1 Research Framework

The research was conducted according to the following tasks:

1. Data collection and annotation
2. Pre-processing
3. AI model training (image segmentation)
4. Post-processing
5. Quality checks

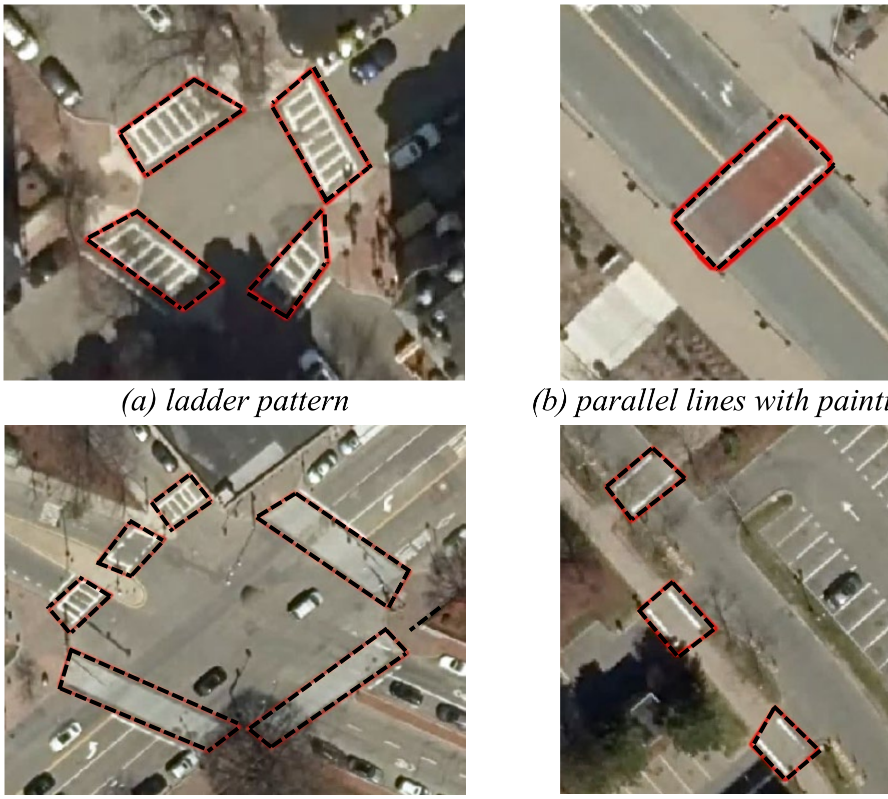
3.2 AI Model Development and Training

The AI model was developed using the Environmental Systems Research Institute (ESRI) ArcGIS Pro software. ArcGIS Pro provides a suite of tools that can facilitate the development of AI models to perform tasks such as object detection and classification, image segmentation, and change detection. Object detection and classification (ODC) models generate horizontal bounding boxes around crosswalks, while image segmentation models classify each pixel as crosswalk or non-crosswalk-related.

Given our large training data set, training an AI model using ArcGIS Pro typically takes between 15 and 20 hours. Therefore, it is impractical to conduct a comprehensive comparison of all the models available in ArcGIS Pro. We evaluated several popular models, including MaskRCNN, FasterRCNN, U-Net, and DeepLabv3Plus. The first two models are for ODC, whereas U-Net and DeepLabv3Plus are for image segmentation. We found that DeepLabv3Plus performs better than the other three models and offers the following advantages:

- Existing ODC models in ArcGIS Pro only generate horizontal bounding boxes. This creates many overlapping and large bounding boxes. These bounding boxes include areas/pixels that do not belong to crosswalks.
- We use the road network to classify the detected crosswalks into different categories, such as intersection-related and driveway-related. Large and overlapping horizontal bounding boxes make this classification task very challenging.
- Although some latest AI models can produce rotated bounding boxes that better match detected crosswalks than horizontal ones, these models are not included in ArcGIS Pro. Integrating them into ArcGIS Pro will take a considerable amount of time and effort.
- Both U-Net and DeepLabv3Plus generate polygons that well match detected crosswalks. By visually comparing their results, DeepLabv3Plus clearly performs

better. In another project funded by MassDOT *(1)*, one of the Co-PIs of this project evaluated U-Net and DeepLabv3. Given our prior experience with these two models and the preliminary results in this research, we decided to choose DeepLabv3Plus for all subsequent modeling work in this study.



(a) ladder pattern *(b) parallel lines with painting*

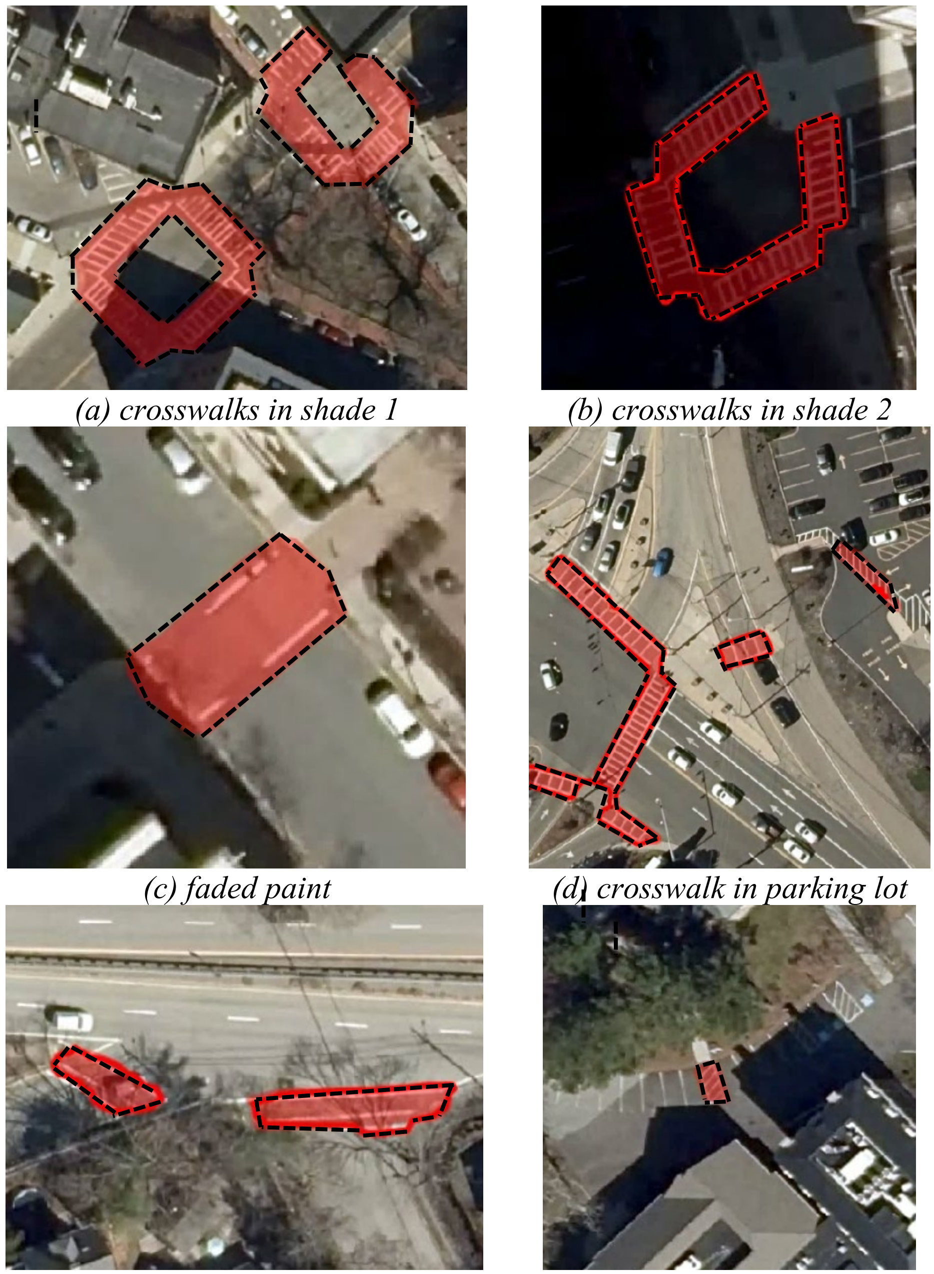
(c) different pavement materials and ladder pattern *(d) parallel lines without painting*

Figure 3. Sample annotations done manually

The AI model development and training involves three key steps:

- Crosswalk annotation: This step is to manually draw bounding boxes around crosswalks using the "Training Samples Manager" tool in ArcGIS Pro. The entire state of Massachusetts is covered by about 10,000 aerials photos. We selected a few photos that well represent urban, suburban, and rural areas. The remaining aerial photos will be processed automatically by the trained AI model. Some examples of the annotations are shown in Figure 3. Note that the crosswalks are in different forms, including (a) ladder pattern, (b) parallel lines with painting, (c) different pavement material, and (d) parallel lines without painting. When annotating these crosswalks, (b) and (d) are combined into one category. In total, we have three categories.
- Exporting training data: This step is to export the annotated images into a standard format that can be fed into the AI model for training. The "Export Training Data For Deep Learning" tool in ArcGIS Pro is used for this task.

- Train deep learning model: This step is based on the ArcGIS Pro "Train Deep Learning Model" tool, and it typically takes around 16 hours for training one model using an Nvidia GeForce RTX 4090 GPU computer.



(a) crosswalks in shade 1   *(b) crosswalks in shade 2*

(c) faded paint   *(d) crosswalk in parking lot*

(e) crosswalks blocked by trees   *(f) false positive in a parking lot*

Figure 4. Sample AI model detection results

The trained AI model is saved in the ".dlpk" format and inputted into the "Classify Pixels Using Deep Learning" tool in ArcGIS Pro for automatic processing of the remaining unannotated aerial photos. The direct outputs of this tool are raster layers consisting of individual pixels classified into four categories: background and three categories of crosswalks. To facilitate the subsequent data post-processing (see Section 3.3), these raster layers are further converted into polygon layers. While the above annotation step was performed manually, Python scripts were prepared to automate the second and third steps.

Overall, the DeepLabv3Plus model performs very well. It accurately detects crosswalks in shade, those blocked by tree limbs, or with faded paint in most cases. Figure 4 presents some sample outputs. Despite its accuracy, there are instances of false positives. For example, some crosswalks in parking lots are detected. Most of these issues can be filtered out by utilizing the road network layer as discussed in the post-processing step.

3.3 Post-Processing

The post-processing entailed identifying and discarding false positive detections, as well as categorizing crosswalk locations as Intersection, Midblock, and Driveway. The framework is shown in Figure 5.

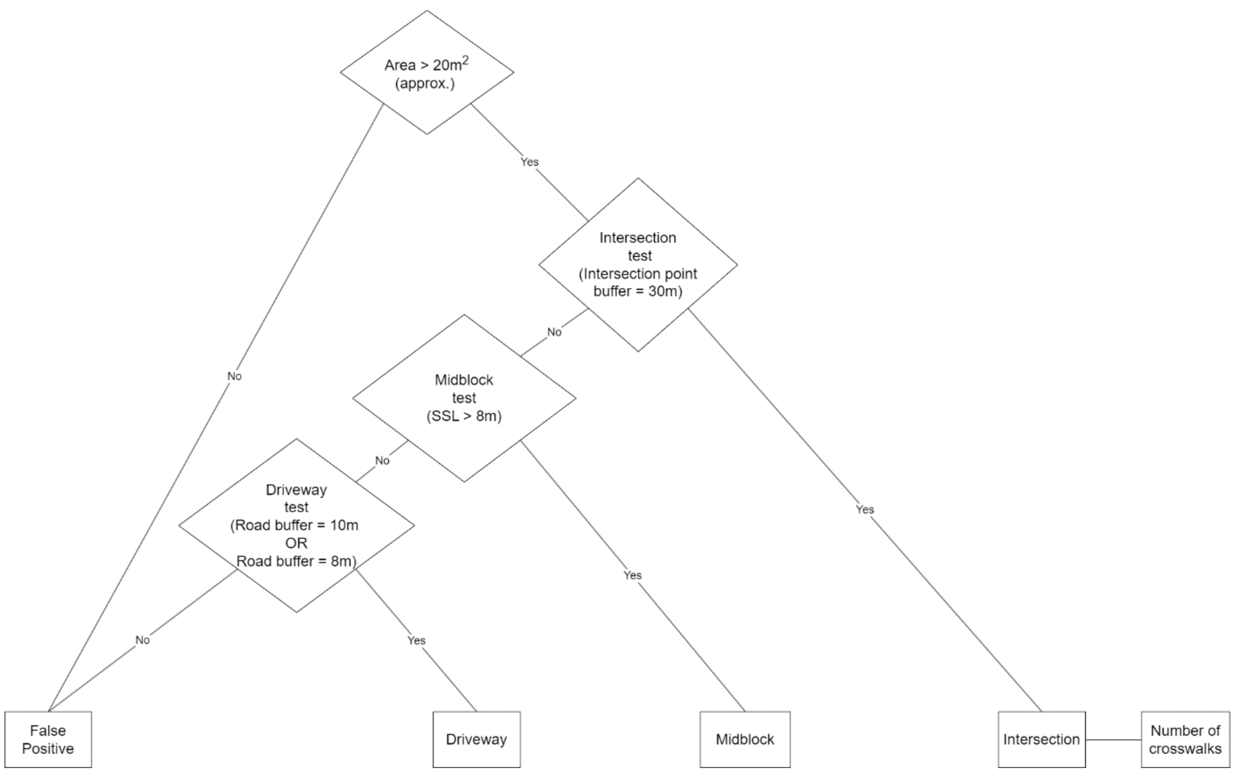


Figure 5. Post-processing framework

We also excluded segments whose jurisdiction indicated that they were either part of a private road (code "H") or that they were unaccepted by city or town (code "0"). This ensured that only crosswalks on public roads were included in the final inventory. We note that the dissolved split road network does not contain jurisdictional information. Thus, there might be some midblock crosswalks which remain on private roads.

3.3.1 Area Thresholding

First, we applied area thresholding to filter out the false positive detections. The area threshold was set to 20 m^2 based on empirical experiments. This was done to ensure uniformity among polygons from different years and parts of the state.

3.3.2 Intersection Test

An intersection points layer was provided by MassDOT. In order to identify which of the crosswalks were located at intersections, we constructed a circular buffer (radius 30 m) around each intersection point. Any crosswalk polygons detected within or intersecting the buffer were categorized as "Intersection." Given that several intersection crosswalks were jointly detected as single polygons (L-shaped, U-shaped, and square annular), we counted the total numbers using the number of road segments intersecting the polygons. If no road segment intersects a polygon, which falls within the 30 m intersection point buffer, the number of crosswalks associated with that polygon was marked as 1 (Figure 6). In Figure 6, the crosswalk polygon buffer is the rectangle.

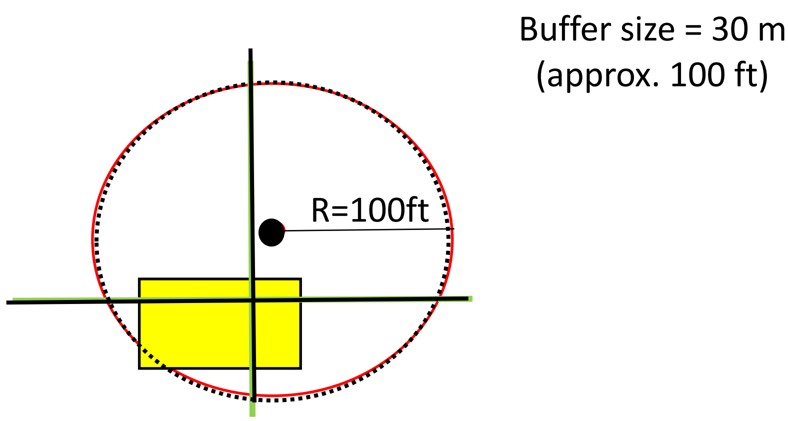


Figure 6. Intersection test

Following their categorization, actual crosswalk counts in joint intersection polygons were obtained by counting the number of segments intersecting each polygon.

3.3.3 Midblock Test

Any polygon that was not categorized as an intersection crosswalk was passed to the midblock test. Here, if a polygon had an intersecting segment, we computed the subsegment length (SSL) on either side of the polygon. If both SSLs were at least 8 m long, then the corresponding polygon was categorized as a midblock. This is illustrated in Figure 7, in which the crosswalk polygon buffer is shown as a rectangle.

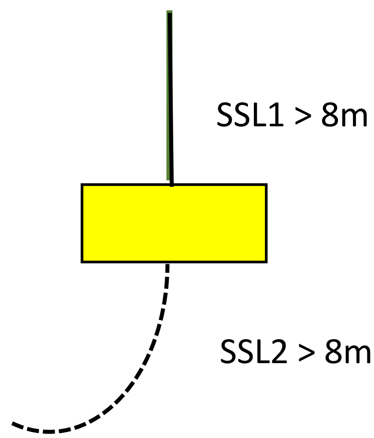


Figure 7. Midblock test

The original road network layer ("Roads Arc") provided by MassDOT was initially used to conduct all the categorization tests. However, this network included streets that were partitioned into segments at town boundaries. We observed that this led to some midblock crosswalks being miscategorized as driveway crosswalks due to short subsegment lengths. Thus, we obtained the "Dissolved Split" road layer whose road segments were merged across town boundaries. We then used this layer to conduct the midblock test instead.

3.3.4 Driveway Test

The driveway test was finally applied to polygons that failed the midblock test. First, we constructed a large polygon buffer around the polygon. If that buffer intersected with a roadway, we checked the road type and then created a corresponding road segment buffer around it. The road types and corresponding buffer sizes are shown in Table 3. For simplicity, only two buffer sizes (10 m and 8 m were considered). If the polygon then intersected with the road segment buffer, then it was categorized as a driveway.

Table 3. Summary of buffer size selection for road type

Road type	**Buffer Size**
1. Limited Access Highway 2. Multilane Highway, not limited access 3. Other numbered route 4. Major road - arterials and collectors	10 m
5. Minor street or road (with Road Inventory information, not class 1–4) 6. Minor street or road (with minimal Road Inventory information and no street name) 7. Ramp 8. Tunnel	8 m

The driveway test is illustrated in Figure 8, in which the crosswalk polygon buffer is shown as a rectangle.

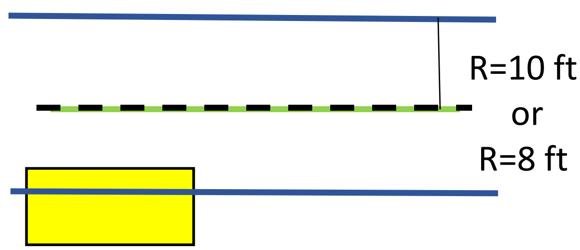


Figure 8. Driveway test

This page left blank intentionally.

4.0 Results

4.1 AI Model Results

We report the performance of the trained AI model in Table 4. As shown, the model has high accuracy and precision. The recall metric is also very high. An epoch signifies the use of all training data in estimating model parameters. Typically, several passes have to be made in order to converge on the optimal state of the model. The dice similarity coefficient (DSC) measures the spatial overlap/similarity between the detected crosswalk areas and the annotated ones. The formula is given by:

$$DSC(A, B) = 2\frac{A \cap B}{A + B}$$

where A is the detected area, and B is the annotated or ground truth area.

Table 4. Model training performance

Epoch	Training loss	Validation loss	Accuracy (%)	DSC (%)
0	0.55	0.89	99.69	79.94
1	0.54	0.87	99.69	80.05
2	0.55	0.89	99.69	79.71
3	0.54	0.87	99.70	80.19
4	0.53	0.87	99.69	80.16
5	0.53	0.85	99.70	80.41
6	0.53	0.85	99.70	80.51

The validation performance metrics of the model are given in Table 5. The accuracy (*Acc*) captures the overall correctness of the model in detecting crosswalk space and non-crosswalk space. The precision (*Pr*) is defined by the proportion of correct crosswalk pixels compared to all the detected crosswalk pixels. Finally, the recall (*Re*) captures the percentage of annotated crosswalk space that was correctly detected by the model. The equations for these metrics are as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Pr = \frac{TP}{TP + FP}$$

$$Re = \frac{TP}{TP + FN}$$

$$F_1 = 2\frac{Pr \times Re}{Pr + Re}$$

Class 1 refers to continental/zebra crosswalks, Class 2 to parallel lines/standard, and Class 3 to fully painted. The "No Data" column refers to non-crosswalk space in the images.

Table 5. Model validation performance

Metric	No Data (%)	Class 1(%)	Class 2 (%)	Class 3 (%)
Precision	99.81	92.42	87.64	69.88
Recall	99.89	87.33	80.38	61.20
F_1	99.85	89.80	83.85	65.25
Accuracy	99.70%			

4.2 Post-Processing Results

The post-processing involved identifying false positives and categorizing the location of the crosswalks. In Table 6, we compare the number of true positives to the false positives filtered out by the model. We see that roughly the same number of false positives as true positives are obtained. However, the false positives comprise mostly smaller-area polygons (area under 20 m^2).

Table 6. Total crosswalk detections (true positives and false positives).

Metric	2019 Count	2021 Count
True positives	83,380	88,440
False positives	80,567	91,098
Total polygon count	163,947	179,538

Examples of categorized intersection crosswalks are shown in Figure 9. In these images, the dots indicate the location of the intersection point, while the dotted lines indicate the road segments. The circles indicate the 30 m buffer around the intersection point. We note that each image is specific to the corresponding crosswalk detected. Thus, as in the third image on the top row, there are detected crosswalks not highlighted. (But these are highlighted in the images specific to them.) The "Polygon Area" in each image title indicates the area of the detected crosswalk polygon in the corresponding image.

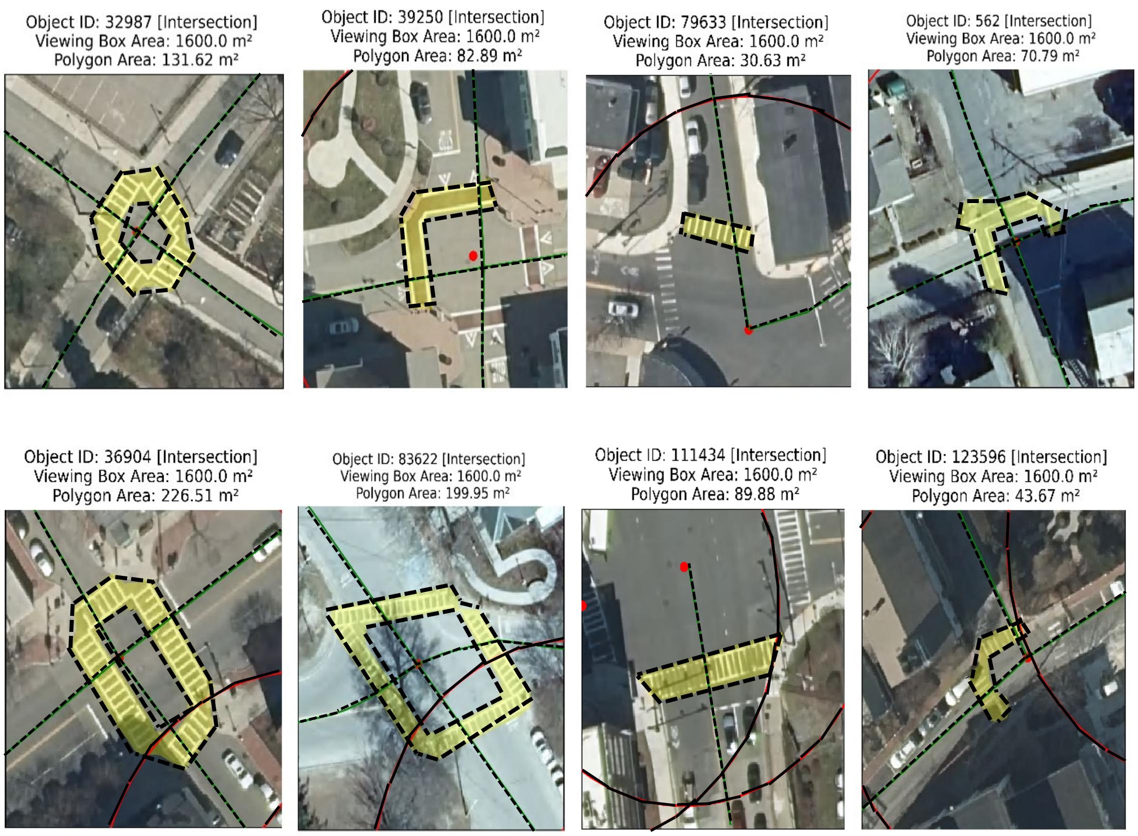


Figure 9. Intersection crosswalks

Not all intersection crosswalks intersect with the road network, as discussed earlier. However, they lie within the vicinity of an intersection, based on the 30 m buffer test. Figure 10 shows two intersection crosswalks for which the polygon does not intersect with the road network.

Figure 10. Intersection crosswalks

Detected crosswalks that are categorized as midblock are shown in Figure 9. As discussed earlier, these categorizations were made using the "dissolved split" layer. Each image indicates the two subsegments (SS1 and SS2) on either side of the crosswalk polygon (dashed rectangle) and their corresponding lengths.

Figure 11. Midblock crosswalks

Examples of categorized driveway crosswalks are shown in Figure 12. The road segments are shown in dotted line, while the surrounding solid lines indicate the buffer around the respective segment. Crosswalks are highlighted by dashed lines.

Figure 12. Driveway crosswalks

4.2.1 Dissolved Road Network Layer

A summary of the midblock crosswalk counts using the “Dissolved Split” GIS layer compared to the “All Roads” layer is shown in Table 7. In both years, using the dissolved layer increased midblock crosswalk detections by 2,000 or more. Thus, we use the dissolved road network to report final results.

Table 7. Crosswalk count comparison

Crosswalk Category	All Roads Layer		Dissolved Split Layer	
	2019	2021	2019	2021
Midblock	4,426	4,577	6,722	6,927
Driveway	2,188	2,362	2,994	3,205

4.2.2 Crosswalk Counts

Table 7 shows the breakdown by category of the true positives obtained from the post-processing framework. From the results, we see that intersections account for 88% of the detections in both years, whereas the midblock crosswalks account for about 8% of the detections, also in both years. The driveway crosswalks make up roughly 4% in both years.

Table 8. Crosswalk counts by category

Crosswalk Category	2019 Count	2019 Proportion (%)	2021 Count	2021 Proportion (%)	YoY Pct. difference (%)
Intersection	73,664	88	78,308	88	6.3
Midblock	6,722	8	6,927	8	3
Driveway	2,994	4	3,205	4	7
Total	83,380	100	88,440	100	6.1

In Table 8, we summarize the crosswalk types (continental, standard, solid) predicted by the model in both years. We observe that continental (zebra) crosswalks are the majority, accounting for 62–64% of the total counts. Standard/parallel line crosswalks are the second most ubiquitous, comprising 36–38% of total counts. Solid/painted crosswalks account for the remainder (<1%).

Table 9. Crosswalk counts by type

Crosswalk Type	2019 Count	2019 Proportion (%)	2021 Count	2021 Proportion (%)	YoY Pct. difference (%)
Continental	51,558	61.83	56,381	63.75	9.35
Standard/parallel	31,606	37.91	31,816	35.97	0.66
Solid/painted	216	0.26	243	0.27	12.50
Total	83,380	100	88,440	100	6.07

4.2.3 Validation

To quantify the performance of the post-processing steps, we sampled 3% of the crosswalks from each category. This proportion was chosen in consideration of the number of crosswalks that had to be manually checked, given the time constraints of the project. In each of these samples, we checked the crosswalk image to ensure that the crosswalk was correctly detected (true positive) and that its category was correctly assigned. Any incorrect detections or categorizations were considered as errors. Based on these, we computed accuracy scores (as shown in Table 10).

Table 10. Validation results using the dissolved split layer

Category	Intersection	Midblock	Driveway
Sample size	1,724	197	95
Accuracy	99%	93%	84%

4.3 Test Performance Metrics

Five of the annotated tiles were not used ("unseen") in model training. We then applied the AI model to detect crosswalks in these tiles and computed area/pixel-based test performance metrics. Across these five tiles, 666 crosswalks had been annotated. The test results are summarized in Table 11. We note that if the model does not detect the entirety of a crosswalk area (due to, for instance, faded paint), then the pixel-based metrics may seem to underestimate performance. Thus, even though the recall score is much lower than the precision, a count-based metric may produce higher performance numbers.

Table 11. Model test performance results

Number of Polygons	Area	Accuracy	Precision	Recall
666	27,531.63	99.95%	95.76%	83.35%

4.4 Comparison of 2019 and 2021 Results

We performed a comparison of the 2019 and 2021 results by applying an area-based overlay. We calculated the geometric difference of the detected polygons from both years and discovered a mismatch of 1.034 million m^2. This represents the polygon area that can be found in 2021 but not in 2019. The total area of the detected polygons for each year is given in Table 12. The dice similarity coefficient between the detected crosswalk area in both years is 79%.

Table 12. Comparison of detection results between 2019 and 2021

Year and Metric	2019	2021	Difference	Overlap	Dice Coefficient
Total detected crosswalk area (million m^2)	4.36	4.54	1.03	3.50	79%

There are a few explanations for the nearly 20% discrepancy. First, given that the recall rate of the model is about 83%, some of the polygons undetected in one year likely are detected in another year and vice versa, and this difference could be up to 17% or even more, in terms of area. Yet, we note that the difference in the number of polygons detected in 2019 and 2021 is only 7%, which might indicate that in terms of actual crosswalks counted, the discrepancy is not as severe. The reason that the area-based difference may magnify differences may be because not all parts of a crosswalk may be detected in either year. There may also be some shifts or wear and tear, as shown in the examples in Figure 13, which shows 2019 (dashed shape with hash marks) and 2021 (solid line) detection results. Areas of overlap are shown by dashed shapes. Thus, a more robust analysis may yield a better comparison between both years in terms of unique crosswalk detections. This might include (1) accounting for translational shifts and partial detections in comparing results for both years; and (2) incorporating the detection of wear and tear as part of the AI modeling framework in order to obtain a comprehensive understanding of crosswalk inventory and condition in each year. Ultimately, the union of detections in both years could be combined for a final result in order to ensure that the complete number of crosswalks is detected using data from both years.

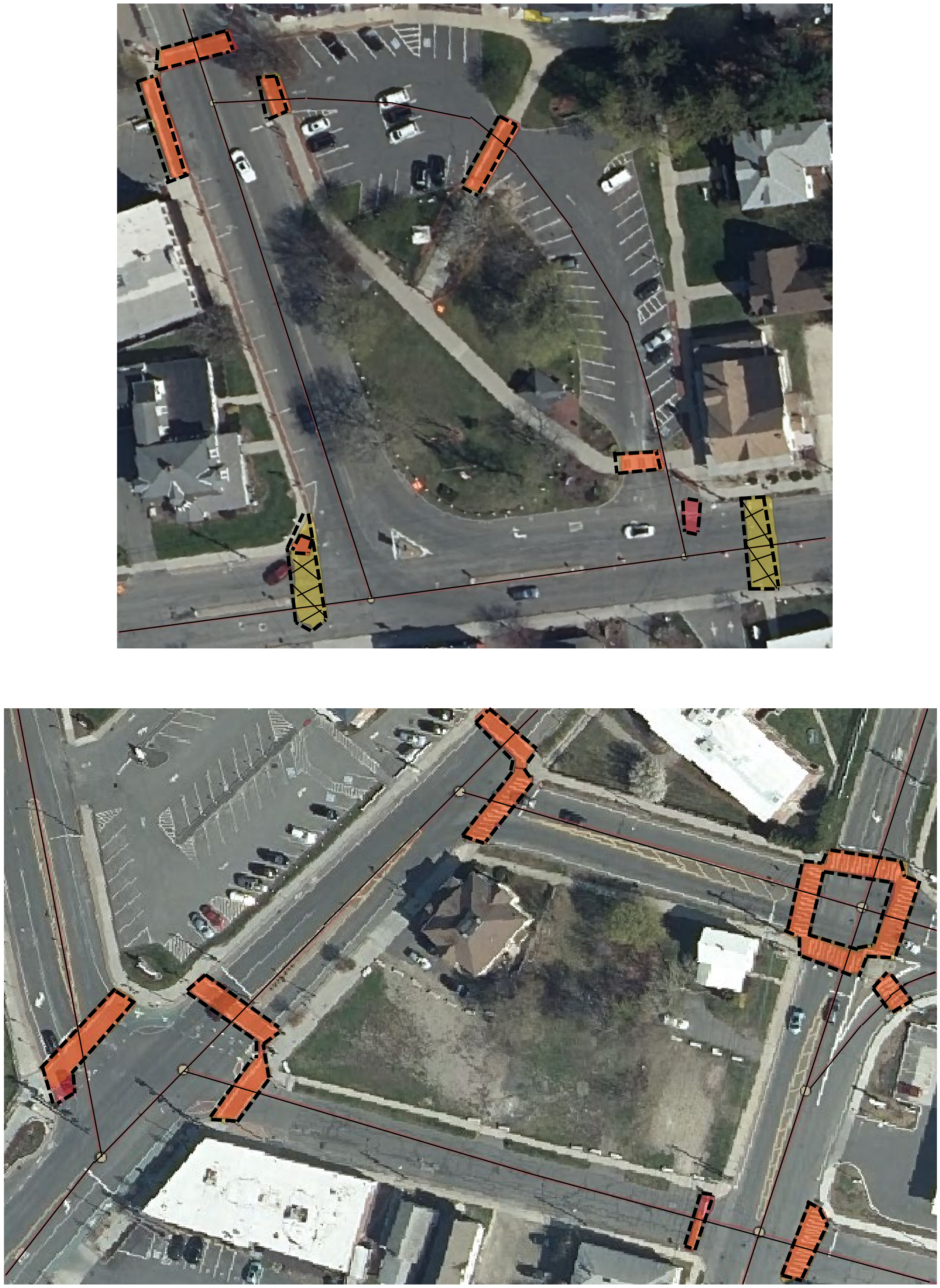


Figure 13. Overlays of 2019 and 2021 detection results

This page left blank intentionally.

5.0 Conclusion

In summary, we have trained an image segmentation AI model for detecting crosswalks in the state of Massachusetts. We applied a post-processing procedure for categorizing intersection, midblock, and driveway crosswalks. The model achieved high validation accuracy for all three categories. We analyzed visualizations of each crosswalk type to assess the viability of our procedure.

Our results and approach demonstrate that even with AI, human expert knowledge is still required in order to best realize the benefits of the AI model. With a bit of effort in annotation of just a fraction of the aerial images prior to training, and further effort in post-processing and cleaning up false detections, we were able to generate crosswalk detections across 10,000 aerial images in a given year for the entire state. If the crosswalks were to have been manually identified from aerial images for the entire state without AI assistance, this may have taken almost two years of person-hours! Thus, we see that artificial intelligence, when properly used, can save significant amounts of time (labor) and costs, even with the errors notwithstanding.

Ultimately, the robust framework we have developed can be further fine-tuned and can assist policymakers in pedestrian safety and traffic planning initiatives as it aids manual inspection of the state's road infrastructure assets.

This page left blank intentionally.

6.0 References

1. Sun, X., Y. Xie, L. Jiang, Y. Cao, and B. Liu. DMA-Net: DeepLab with Multi-Scale Attention for Pavement Crack Segmentation. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 10, 2022, pp. 18392–18403.