Milkweed (Asclepias syriaca) plant detection using mobile cameras

KORAY OZCAN,1†, ANUJ SHARMA,1 STEVEN P. BRADBURY,2,3 DANA SCHWEITZER,2 TEREŞA BLADER,2 AND SUE BLODGETT2,3

1Institute for Transportation, Iowa State University, Ames, Iowa, USA
2Department of Entomology, Iowa State University, Ames, Iowa, USA
3Natural Resource Ecology and Management, Iowa State University, Ames, Iowa, USA

Abstract. Milkweed (Asclepias spp.) are host plants of monarch butterflies (Danaus plexippus). It is important to detect milkweed plant locations to assess the status and trends of monarch habitat in support of monarch conservation programs. In this paper, we describe autonomous detection of milkweed plants using cameras mounted to vehicles. For detection, we used both aggregated channel features (ACF) for running the detectors on embedded computing platforms with central processing unit and faster region-based convolutional neural network (Faster R-CNN) with a ResNet architecture-based detector that is suitable for graphics processing unit optimized processing. The ACF-based model produced 0.89 mean average precision (mAP) on the training dataset and 0.29 mAP on the test dataset, whereas the ResNet-based Faster R-CNN model provided 0.98 mAP on training and 0.44 mAP on the test dataset. The detections were used to calculate approximate densities of milkweed plants in geo-referenced locations based on global positioning system point correspondences of recorded images. Probability-of-count distributions are compared for the actual milkweed plant locations near roadsides. This is one of the first examples of using automated milkweed plant detection and density mapping using a vehicle-mounted camera.

Key words: aggregated channel features (ACF); Asclepias syriaca; autonomous plant detection; faster region-based convolutional neural network; milkweed.

INTRODUCTION

The eastern North American population of the monarch butterfly (Danaus plexippus) has declined by 80% over the past two decades (Semmens et al. 2016). Increasing reproductive success in the Midwest of the United States during the summer is identified as a high priority for monarch conservation (Flockhart et al. 2015, Oberhauser et al. 2017). Monarch butterflies oviposit only on milkweed species (mainly Asclepias spp.) and primarily on common milkweed (Asclepias syriaca) in the Midwest (Malcolm et al. 1993). To increase monarch populations to a level that would reduce the probability of quasi-extinction by 50% over 20 yr, Pleasants (2017) and Thogmartin et al. (2017) estimated that 1.3–1.6 billion additional milkweed stems need to be added to the landscape in the U.S. Midwest. The current amount of common milkweed stems in the Midwest is estimated to be approximately 1.3 billion, the majority of which is in publicly owned grasslands; land enrolled in conservation programs, such as the United States Department of Agriculture’s Conservation Reserve Program (CRP); and road rights-of-way (ROWs), which...
are the strips of public or private property along both sides of roads (Pleasants and Oberhauser 2013, Pleasants 2017, Thogmartin et al. 2017). Flockhart et al. (2015) estimate that roadsides account for 10% of the remaining milkweed in the central U.S. region. Based on landscape-scale modeling (Thogmartin et al. 2017, Grant et al. 2018), increasing milkweed density in rural road

Fig. 1. Camera setup attached to the vehicle.

Fig. 2. Examples illustration showing the processes of aggregated channel features.
ROWs will play a significant role in reaching monarch conservation goals. Thogmartin et al. (2017) estimate that the density of milkweed in road ROWs in the Upper Midwest needs to increase approximately 1.5–2-fold. Model estimates assessing the impact of habitat increases in 22 land cover classes are most sensitive to assumptions of habitat establishment rates in marginal agricultural land and protected grasslands, followed by secondary road ROWs (Thogmartin et al. 2017).

Current road ROWs are heterogeneous in common milkweed density and are typically detected in patches of various sizes (Hartzler 2010, Blader 2018). Kasten et al. (2016) surveyed 212 five mile long, non-urban, non-forested roadside ROWs once per site in Minnesota, Wisconsin, South Dakota, and Iowa from July to October 2015. These authors detected milkweed in approximately 60% of the sites. For those sites with milkweed, the mean stem density (±standard error) was 35 stems/acre (±4.8), with a range of 3.3–2100 stems/acre. Blader (2018) sampled four 1-mile gravel roadside ROWs in Story County, Iowa, four times from June through August 2017. These sites all contained milkweed. Blader (2018) reported a mean stem density (±standard deviation) of 575 stems/acre (±357). Thogmartin et al. (2017) reported a mean of 57 stems/acre for several roadside surveys completed prior to Blader (2018) and Kasten et al. (2016). These authors also proposed that 100–200 stems/acre in

Fig. 3. (a–d) Example training images used from the milkweed dataset collection.
roadside ROWs was a biologically reasonable upper limit stem density based, in part, on survey results used in their analyses. The variability in reported stem densities across surveys completed to date is likely due, in part, to the large amount of effort required to sample roadsides, which in turn limits the means to assess stem densities adequately across space and time. Our proposed method provides efficient and automated sampling of roadside for milkweed plant density. Large-scale monitoring programs for milkweed are being designed and piloted (MJV 2017); however, the level of effort for sampling sites needs to be reduced to support a statistically rigorous sampling scheme.

Automated sampling of the landscape to detect and quantify milkweed plant density could significantly expand survey coverage by increasing the efficiency of sampling and reducing the level of effort. To increase efficiency in sampling, while maintaining or enhancing accuracy, airborne remote sensing of milkweed plants has been proposed by Burai et al. (2011), consistent with advances in using low-altitude drones to survey for other beneficial plants (Cruzan et al. 2016) and weed species (Barrero et al.

Fig. 4. (a–d) Example detection results of the detector trained with aggregated channel features. The numbers show the detection score (higher is better).
Burai et al. (2011) proposed milkweed plant mapping using airborne hyperspectral imagery. They were able to classify various milkweed plant locations using the spectral angle mapper (SAM) method. This classification method measures pixels close to the measured spectra on field according to previously determined search conditions; that is, the method classifies hyperspectral images without spectral transformation. In a similar vein, Barrero et al. (2016) implemented neural networks to detect weeds in rice fields using aerial images captured with a camera mounted on an unmanned aerial vehicle (UAV).

While drones, or UAVs, could provide a useful platform for roadside ROW milkweed sampling, a vehicle-mounted camera may better complement state- or county-level departments of transportation, which are beginning to implement remote sensing technology with vehicle platforms to assess the condition of 713,000 road miles per year (U.S. Department of Transportation and HM-64 2015). In addition, UAVs require special permissions and licenses to be operated over public and private property, which creates additional regulatory and logistical constraints.

Here, we report a prototype, vehicle-mounted, remote sensing technology that provides fast processing time and is especially suitable for embedded systems running on central processing unit (CPU)- or graphics processing unit (GPU)-based implementation of vision methods and algorithms. We developed an aggregated channel features (ACF; Dollár et al. 2014)-based model and a Faster R-CNN (region-based convolutional neural network; He et al. 2016) model using a ResNet backbone, which is an end-to-end training model that selects convolution neural networks intelligently while omitting handcrafted intermediary algorithms and features. End-to-end learning does not employ handcrafted or intermediary algorithms; learning is directly based on the features or patterns for a given problem in the sampled dataset (i.e., the detection of a milkweed plant from the other vegetation in the roadside). The ACF-based model is especially useful for CPU on embedded system implementations of image processing algorithms. The Faster R-CNN model using ResNet backbone is more suitable to be employed in GPUs while providing higher accuracy for detection. To the best of our knowledge, this study is the first to demonstrate the potential application of automated detection and mapping of milkweed plants in roadside ROWs using vehicle-mounted cameras.
METHODS

Milkweed dataset

Images for the training dataset were primarily collected in Boone and Story counties in central Iowa, USA, during a six-week period from 10 July through 15 August 2017. Images were captured in the morning (at least two hours after sunrise) and afternoon (up to 3 h before sunset). Weather conditions during data collection included sunny, partly cloudy, and cloudy days; no data was collected during precipitation events. A total of 2746 images were collected from areas with milkweed plants. Collected images include variable brightness, texture, and maturity stage of milkweed.

Table 1. Comparison of mAP on the training and test datasets.

<table>
<thead>
<tr>
<th>Model</th>
<th>mAP training</th>
<th>mAP test</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>0.89</td>
<td>0.2861</td>
</tr>
<tr>
<td>Faster R-CNN with ResNet50</td>
<td>0.98</td>
<td>0.444</td>
</tr>
</tbody>
</table>

Notes: ACF, aggregated channel features; Faster R-CNN, faster region-based convolutional neural network; mAP, mean average precision.

Fig. 6. (a–d) Example detection results of the detector trained with model trained with Faster R-CNN. The numbers show the confidence score between 0 and 1 (higher is better).
milkweed plants. The camera setup used for collecting images is illustrated in Fig. 1. The camera model and its lens were specifically selected to eliminate noise from the moving camera (i.e., Blackfly S Color 1.3 MP USB3 Vision camera; FLIR, Wilsonville, Oregon, USA) with a Techspec HP Series 16 mm fixed focal length lens (Edmund Optics, Barrington, New Jersey, USA). The camera was attached to a vehicle traveling with an average speed from 3 to 10 m/s depending on the sampling location and time.

Fig. 7. Milkweed plant locations around Ames, Iowa. Various regions are selected for the experiments, and regions are marked as R11, R22, R6, R7, and R8 for image sequences collected. The red pinpoints show the milkweed plant locations by the side of the road. The arrows indicate the vehicle direction when recording continuous images of the roadside. R6: Lat (min, max): 42.03470, 42.03477, Long (min, max): −93.7954, −93.7765; R7: Lat (min, max): 42.03470, 42.03477, Long (min, max): −93.7953, −93.7765; R8: Lat (min, max): 42.10704, 42.10705, Long (min, max): −93.6593, −93.6592; R11: Lat (min, max): 42.10704, 42.10705, Long (min, max): −93.6593, −93.6592; R22: Lat (min, max): 41.96475, 41.96476, Long (min, max): −93.6584, −93.6592; R8: Lat (min, max): 42.12312, Long (min, max): −93.6593, −93.6592; R11: Lat (min, max): 41.96474, 41.96492, Long (min, max): −93.6584, −93.6592; R22: Lat (min, max): 42.96475, 42.96488, Long (min, max): −93.6583, −93.6392.
Training milkweed plant object detectors

Model training employing aggregated channel features.—Aggregated channel features include feature channels of gradient, histogram of oriented gradients, and LUV color space consisting of luminance and chromaticity coordinates (Dollár et al. 2014). Image gradient is a directional change in color intensity calculated with

Table 2. Environmental properties of roadside.

<table>
<thead>
<tr>
<th>Site</th>
<th>Time</th>
<th>Weather</th>
<th>Temperature (°C)</th>
<th>Wind</th>
</tr>
</thead>
<tbody>
<tr>
<td>R6</td>
<td>15:02–15:09</td>
<td>Fair/windy</td>
<td>32.7</td>
<td>SSE 35 kph</td>
</tr>
<tr>
<td>R7</td>
<td>16:12–16:17</td>
<td>Fair</td>
<td>32.7</td>
<td>SSE 32 kph</td>
</tr>
<tr>
<td>R8</td>
<td>15:57–16:03</td>
<td>Fair</td>
<td>32.7</td>
<td>SSE 32 kph</td>
</tr>
<tr>
<td>R22</td>
<td>13:50–13:57</td>
<td>Cloudy</td>
<td>31.6</td>
<td>SSE 22 kph</td>
</tr>
</tbody>
</table>

Fig. 8. (a–d) Example detections from sequential images captured from roadside R6.
derivatives in horizontal and vertical axes. Histogram of oriented gradients is a histogram vector definition based on oriented gradients that was proposed initially for pedestrian detection in images (Dalal and Triggs 2005). In this model, we used a decision tree of depth 3. The total number of training stages was five, and the final stage had 4096 trees. During training, the total number of negative samples was limited to 40K, while the number of negative samples per image was limited to 500. The maximum number of accumulated negative samples was set to 80K. Since a high number of negative samples were used for better classification, the depth of the decision tree was selected to be three for better detection performance. Decision tree boosting was used over the pixel features in order to train a model that separates objects from background. ACF approximates state-of-the-art detections in pedestrian detection among sliding-window-based approaches (Dollár et al. 2014). Fig. 3 presents example images of the positive training samples of milkweed plants at different stages throughout the sampling period. Some example images, used for training the detectors, are displayed in Fig. 3. The model training took approximately 2.5 h using a computer with an Intel Xeon 3.7 GHz processor with 16 GB of memory (Intel, Santa Clara, California, USA). Example detection results for this model are presented in Fig. 4.

Model training using Faster R-CNN architecture using ResNet backbone.—For the deep learning-based model employing Faster R-CNN architecture, we used the Detectron software environment (Girshick et al. 2018). The ResNet backbone

Fig. 9. Example milkweed detection that is quite hard to detect with human eye.
architecture (network-depth-features) had a depth of 50 layers (He et al. 2016) for feature extraction, as illustrated in Fig. 5. The network was initialized with the weights pre-trained on the ImageNet dataset (Krizhevsky et al. 2012). The features of the original Faster R-CNN (Lin et al. 2017) with ResNet were extracted from the final fourth stage of the convolution layer. Hyper-parameters were set following the existing implementation of Faster R-CNN with ResNet50 (He et al. 2016). The image region was cropped from a proposal region, and it was warped to $224 \times 224$ pixels in image size and fed into the classification network. We used stochastic gradient descent with a mini-batch size of 256. The network was fine-tuned on the training set using a mini-batch size of 256 in the RoI-centric fashion. We trained on a GPU for 60K iterations, with a learning rate of 0.1 that was divided by 10 when the error showed plateau behavior. We used weight decay of 0.0001 and a momentum of 0.9. It took approximately 2.25 h to complete the training for a single NVIDIA GTX 1080 Ti GPU (Nvidia, Santa Clara, California, USA). The details of the implemented Faster R-CNN model are as follows. For testing, the region proposal

![Fig. 10. (a) Missed detections, (b) false positive, (c) multiple plants detected as one, and (d) false positive.](image-url)
network generates the highest scored 1000 proposals for the milkweed class. For best results, we adopted the fully convolutional form and averaged the scores at multiple scales. The images were resized so that the shorter side was either 224, 256, 384, 480, or 640 pixels. Also, the RCNN network was used to update the proposal scores and box positions. The inference for each testing image took about 0.12 s per image or about 8 frames per second (fps) on the GPU. Example detection results for this model are presented in Fig. 6.

Testing stage

To evaluate the milkweed plant detection performance on continuous image sequences, the intersection over union (IoU) criterion used was similar to the Pascal Visual Object Classes (Everingham et al. 2015) challenge. This evaluation criterion is given in Eq. 1, where $B_d$ and $B_{gt}$ represent the bounding boxes of the detected region and the ground truth, for the objects, respectively. Whenever the IoU is >0.5, we counted these detections as true positives (TP). Similarly, when the proposed detection $B_d$ was not overlapping with the ground truth bounding box of $B_{gt}$, or when IoU criterion is smaller than 0.5, it was considered as a false positive (FP). The precision corresponds to TP over the sum of TP and FP as provided in Eq. 2. Furthermore, the mean average precision (mAP) corresponds to mean APs of all the objects available for detection in the object detection literature. Since there is only milkweed class for this project, mAP is equal to AP of the milkweed class.

$$\text{IoU} = \frac{\text{area}(B_d \cap B_{gt})}{\text{area}(B_d \cup B_{gt})}. \quad (1)$$

![Fig. 11. R6 roadside actual milkweed location and predicted density. (a) Probability of count correlation distance: 0.1059. (b) R6 roadside heatmap: Green, yellow, and red color regions show density in an increasing order for the milkweed plant location densities.](image-url)
Average Precision (AP) = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}.

(2)

We compared the milkweed plant detection performance of the two detectors on images captured from vehicle-mounted mobile cameras. While training, mAP reached 0.89 for the ACF-based model and 0.98 for the Faster R-CNN model with ResNet as given in Table 1. On the test dataset, we observed mAP of 0.29 and 0.44, respectively, for ACF-based model and the Faster R-CNN model with ResNet.

RESULTS

Model evaluation: Milkweed detection from roadside images

The ACF-based model was evaluated on images from continuous recordings of ROWs near Ames, Iowa. The results were evaluated based on a set of collected images and sequential images collected from the roadside. In these experiments, we collected a variety of ROW images around Ames, Iowa (sites R11, R22, R6, R7, and R8 in Fig. 7). These roadside images were not used to train the model; that is, the model had no prior information or bias regarding the colors and shapes of milkweed plants recorded in the ROWs. Data to evaluate the model were collected on 15 September 2017; environmental properties of each roadside such as time of the day, weather, temperature, wind direction, and speed are provided in Table 2. The weather conditions on that particular day of the experiment were fair, cloudy, and windy with winds from 22 to 35 kph from the SSE. Temperatures across all five ROWs were approximately 32°C. Example milkweed detections from continuous image recordings are provided in Fig. 8. The model was able to extract milkweed plants when they were distinctly visible compared to other types of plants even in densely populated roadside regions. Fig. 9 provides an example to illustrate the difficulty of detecting milkweed visually (i.e., with the human eye). However, we

![Image](a) Actual Milkweed Locations

![Image](b) Estimated Milkweed Locations

Fig. 12. R7 roadside actual milkweed location and predicted density. (a) Probability of count correlation distance: 0.1224. (b) R7 roadside heatmap: Green, yellow, and red color regions show density in an increasing order for the milkweed plant location densities.
also observed missed detections especially when milkweed leaves were desiccated and small (Fig. 10a). This is most probably due to the reason that our training dataset contained limited examples of dried milkweed plants (Fig. 10b). In addition, multiple milkweed plants were occasionally detected as one due to non-maximum suppression when the detections are too close and overlapping (Fig. 10c). We also observed that in some cases, other more distant plant species create forms that are similar to milkweed causing the detection model to propose object detection (Fig. 10d).

**Model evaluation: Milkweed plant density**

Where milkweed plants were located, we compared the milkweed plant density in terms of probability of counts for actual global positioning system (GPS) locations against estimated plant locations based on object detection outputs. We compared milkweed probability-of-count distributions by using Eq. 3, wherein \( r \) and \( s \) are N-dimensional feature vectors. In our case, the distributions of probability of count for milkweed have a vector size of 20. Correlation distance is a value between 0 and 1, which is used to derive vector similarity. If the two vectors are similar to each other, the correlation distance measure approaches 1.

\[
D = 1 - \frac{\sum_{i=0}^{N-1} (r_i - \bar{r})(s_i - \bar{s})}{\sqrt{\sum_{i=0}^{N-1} (r_i - \bar{r})^2 \sum_{i=0}^{N-1} (s_i - \bar{s})^2}},
\]

\[
\bar{r} = \frac{1}{N} \sum_{i=0}^{N-1} (r_i), \bar{s} = \frac{1}{N} \sum_{i=0}^{N-1} (s_i).
\]

Fig. 13. R8 roadside actual milkweed location and predicted density. (a) Probability of count correlation distance: 0.2567. (b) R8 roadside heatmap: Green, yellow, and red color regions show density in an increasing order for the milkweed plant location densities.
Normalized probability-of-count histogram distributions were calculated for milkweed plant locations noted as R6, R7, R8, R11, and R22 (Fig. 7). Figs. 11–15 provide normalized probability-of-count histograms constructed with detections from the Faster R-CNN model with ResNet due to higher overall accuracy. The histograms are compared with actual normalized milkweed plant densities based on GPS point correspondences. For Figs. 11, 12, 13, and 15, the correlation distances were calculated to be 0.1059, 0.1224, 0.2567, and 0.0364, respectively. For Fig. 14, the correlation distance was calculated to be 0.8470, implying reasonably good representation of the milkweed distribution.

In Figs. 11–15, heatmaps of the milkweed plant densities are also plotted based on estimated milkweed densities given on the right-hand side of the figures for each particular ROW site. When detection-based heatmaps are compared with actual count-based histograms, we can observe the matching peak points of milkweed plant densities. The detection distributions are reasonably concordant with the actual milkweed plant locations; however, mapping accuracy needs to be improved to match exact GPS point locations.

**CONCLUSIONS**

The ACF-based model and the ResNet-based Faster R-CNN model used in this study enable automated detection of milkweed plants in ROWs or in any other image that includes milkweed plants. The ACF-based model produced 0.89 AP on the training dataset and 0.29 AP on the test dataset, whereas the Faster R-CNN-based model provided 0.98 AP on the training dataset and 0.44 AP on the test dataset. Compared to challenging object detection datasets such as MS COCO (Ren et al. 2015), our developed models showed comparable performance on our collected and annotated datasets. On average, the original Faster R-CNN model achieved 0.427 on the MS COCO test dataset (Ren et al. 2015) for various object classes. Although human annotators...
performed well on marking the objects, every annotator is biased to mark bounding boxes in one way or another affecting the overall accuracy of trained and tested models. Faster R-CNN requires a GPU for efficient implementation, while the ACF-based model can run on embedded platforms such as smartphones.

Our models can provide a reasonable estimate of the milkweed plant locations. We observed a high correlation rate, that is, 0.847, between the R11 location milkweed plant density distribution and milkweed plant locations based on images captured with the moving vehicle. We also observed a correlation score of 0.2567 for milkweed plant locations for R8. For other locations such as R7 and R11, the correlation score was quite low for probability-of-count distribution histograms. Since our specialized cameras could only capture image sequences with 3 fps, our continuous image recordings were lower than regular video frame rates, which are typically greater than or equal to 30 fps. Our vehicle was moving with an average speed from 3 to 10 m/s depending on the experiment; hence, image sequence detections could only provide discrete sampling of roadside with an image approximately every 1–3 m traveled. A higher correlation between actual milkweed counts and discrete sampling-based detection counts could likely be achieved with a higher sampling rate to record continuous images. For more accurate mapping of milkweed plant densities, future models should be developed with higher speed cameras to achieve more frequent sampling of milkweed plant distribution. Another reason for lower correlation scores at some ROW sites is that the milkweed plant locations were not exactly within the camera’s field of view. The camera is able to capture plant locations that are closer to the roadway. If the milkweed plant is occluded with other plants/objects or it is not visible due to scattered earth surface within the scene, it becomes harder to detect milkweed locations with a single camera. Using multiple

![Fig. 15. R22 roadside actual milkweed location and predicted density. (a) Probability of count correlation distance: 0.0364. (b) R22 roadside heatmap: Green, yellow, and red color regions show density in an increasing order for the milkweed plant location densities.](image-url)
cameras with aerial view on moving vehicles might provide better density estimations.

Large-scale monitoring programs for milkweed are being designed and piloted (MJV 2017) to assess current habitat condition and to assess trends in habitat expansion with the implementation of monarch conservation programs. The level of effort for national, state, and local sampling programs needs to be reduced through the use of autonomous systems to support statistically rigorous sampling schemes. Increasing milkweed density in rural road ROWs by approximately twofold is needed to reach monarch conservation goals (Thogmartin et al. 2017). Variability in current estimates of ROW stem densities based on traditional surveys is likely due, in part, to the level of effort required to sample roadways, which in turn limits the means to assess stem densities adequately across space and time. Our research establishes a proof of concept for an automated sampling of publicly owned ROWs for milkweed plant density by highway department vehicles.

Small UAV-based techniques, as explained in Cruzan et al. (2016), could also be feasible for monitoring milkweed if flowers, foliage, or other plant structures are grouped and distinctly visible from other plant species in an area. Habitat mapping by UAVs is possible when associated flora and land forms have significant spectral or altitude differences, which has good potential for quantifying milkweed densities in privately owned agricultural land and grasslands. Milkweed plants in ROWs are usually scattered along the length of a road and are not easily distinguished from surrounding plants when densities are low. Based on our prototype models, vehicle-mounted technology can address these sampling challenges and they could be used to efficiently estimate ROW milkweed densities to help provide landscape-scale milkweed density estimates within and across the states in the Upper Midwest.

ACKNOWLEDGMENTS

We would like to thank Iowa State University and Ames High School students Cole Albritton, Joaquin Chavez, Logan Creez, Tyler Delaney, Cory Haggard, Darian Hans, Blake Kassmeyer, Aashwin Lamsal, Riley Nylin, and Logan Peters for collecting and annotating the milkweed image dataset.

LITERATURE CITED


Pleasants, J. 2017. Milkweed restoration in the Midwest for monarch butterfly recovery: estimates of milkweeds lost, milkweeds remaining and milkweeds that must be added to increase the monarch population. Insect Conservation and Diversity 10:42–53.


U.S. Department of Transportation, H. S. H., Federal Highway Administration, and HM-64. 2015. Road condition. https://www.bts.gov/content/road-condition