Development of Framework for Highway Lawn Condition Monitoring using UAV Images

Y. Kim[a], S. Kim[a], Y. Yajima[b], J. Irizarry and[c], and Y.K. Cho[a]

[a]School of Civil and Environmental Engineering, Georgia Institute of Technology, USA. [b]Institute for Robotics and Intelligent Machines, Georgia Institute of Technology, USA.

[c]School of Building Construction, Georgia Institute of Technology, USA.
E-mail: ykim858@gatech.edu, skim3310@gatech.edu, yyajima@gatech.edu, javier.irizarry@gatech.edu, yong.cho@ce.gatech.edu,

Abstract –

Planning, monitoring, and maintenance of highway assets is an essential, long-term operation for successful civil infrastructure management. These monitoring and maintenance activities are usually carried out manually, suffering from time-consuming, costly, potentially dangerous tasks. The advancements in Unmanned Aerial Vehicles (UAVs) and computer vision technologies have demonstrated the potential to enable automation of the monitoring workflows. Existing UAS-based approaches are used for various management; however, there was no study to examine the feasibility of aerial image-based computer vision algorithms for the purpose of lawn condition monitoring. This study aims to provide periodic and easy-to-use UAV technology for civil infrastructure maintenance. We developed the comprehensive framework from UAS data collection to build a deep learning model suitable to distinguish areas of interest with vague boundaries robustly, process the outputs into geo-database, and visualize them through a Geographic Information Systems (GIS) platform. The outcome of the proposed framework displays the overall mowing quality in the highway environment in an intuitive way to support decision-making in the management.

Keywords – Civil infrastructure management; Computer-vision; UAVs; Deep Learning; Lawn mowing condition assessment

1 Introduction

Roadways are one of the most vital infrastructures necessary for public transportation and industrial operations. To secure the robustness of roadway infrastructure, the roadway infrastructure condition (e.g., road condition monitoring, lawn condition monitoring) is monitored and assessed for maintenance needs. These inspection, maintenance, and post-maintenance inspection are often carried out manually with visual inspection by trained inspectors. Thus, it is often time-consuming, laborious, and expensive. One of such tasks is lawn mowing control on the roadway infrastructure. Typically, the goal of lawn mowing maintenance personnel is to verify if the mowing performance carried out by a third-party contractor is acceptable, identify any poorly mowed regions that require additional mowing, and bared regions that require maintenance. Thus, such lawn mowing control requires significant time and resources due to visual observation and manual measurement to verify the mowing performance. Failure to do so may cause traffic safety issues because tall grass can obscure drivers' vision and may not see traffic control devices or pedestrians [1].

The recent advancement of computer vision algorithms and the prevalent use of UAVs offers a great potential for automation in highway landscape maintenance practices. In particular, UAVs equipped with RGB cameras provide an easy-to-use and inexpensive platform to obtain lawn condition information of the field readily. UAVs can capture images of areas where maintenance work was performed, including areas of difficulty (e.g., steep slope). Maintenance personnel can monitor the lawn mowing conditions more frequently and easily by analyzing these aerial images.

To fully utilize the computer vision-based monitoring method, the following limitations should be addressed: 1) to deal with the variability of lawn condition images due to the range of seasonal lawn condition and the different resolutions of aerial images captured at different heights and angles, and 2) the analyzed output needs to be directly mapped onto the actual physical 3D coordinates. To overcome these limitations, we propose a comprehensive, data-driven automated mowing condition monitoring framework integrated with a user interface that visualizes the outcome in a user-friendly manner. We developed a computer vision-based lawn condition monitoring system that adopts fully convolutional neural networks to perform pixel-wise classification of the aerial images. We examined the feasibility of utilizing computer-vision technologies in the mowing condition maintenance work and evaluated the proposed system with actual test cases. The outcomes of this study are particularly beneficial for practical applications and support the decision-making process.

2 Related Work

2.1 Civil Infrastructure Monitoring

Monitoring and maintenance of civil infrastructure are paramount and have been a significant interest in society. It is crucial to promptly detect and identify the location of possible areas requiring necessary maintenance because the failure to do so can lead to traffic congestion, driver discomfort, and potential safety and operational problems. Traditionally trained personnel carried out onsite visual observation [2]. These practices are non-trivial tasks requiring significant resources, and they can be subjective and error-prone as the outcome is highly dependent on the expertise of the onsite personnel. Acknowledging the limited amount of resources in contrast to the vast volume of areas to be monitored, the recent advancements in visual data collection hardware such as UAVs as well as computer vision techniques to analyze the collected visual data have demonstrated its potential opportunities to automate civil infrastructure monitoring processes [3]. Several examples of UAV-based inspection applications are agricultural crop and weed monitoring [4–6], road inspection applications [7], and construction and infrastructure monitoring [8].

Despite the increasing applications of UAV and deep learning in the civil engineering domain, UAV application has not yet been widely researched in the field of highway mowing and maintenance services. Lawn mowing control on the roadway infrastructure requires significant time and resources for prompt maintenance. However, lawn mowing control poses several challenges. Firstly, it requires the person to distinguish areas of interest with vague boundaries granularly. Also, the processed information needs to be transferred to spatial context information to capture the physical location of interest areas.

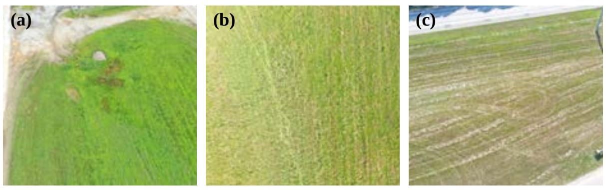
2.2 UAV Image Segmentation and Classification

With the prevalent use of UAVs, UAVs gained growing attention for various monitoring applications because they can cover large areas with high-resolution imagery in an inexpensive way. To this end, image segmentation using aerial images captured by UAVs is also a widely studied topic. For example, in the agricultural areas, UAV images were used to identify weeds from the grass areas [4]. In civil engineering domain, aerial and satellite images were used to assess infrastructure damage in natural disaster areas [9]. In addition, aerial images coupled with machine learning techniques were used to map the rural environment, classifying them into roads, vineyards, asphalt, and roofs [10].

Although there have been studies that demonstrated the capacity of the UAV image segmentation method, there are several challenges to be addressed to deploy in the lawn mowing condition monitoring application. First of all, there is a lack of annotated datasets to develop a segmentation model. Secondly, there is a high intra-variability within the grass conditions based on the seasonal and weather factors (e.g., lighting conditions) and camera factors (e.g., image resolution). To address these limitations, this study aims to develop a framework that can collect and annotate datasets and develop a robust segmentation algorithm for the variance in the collected dataset.

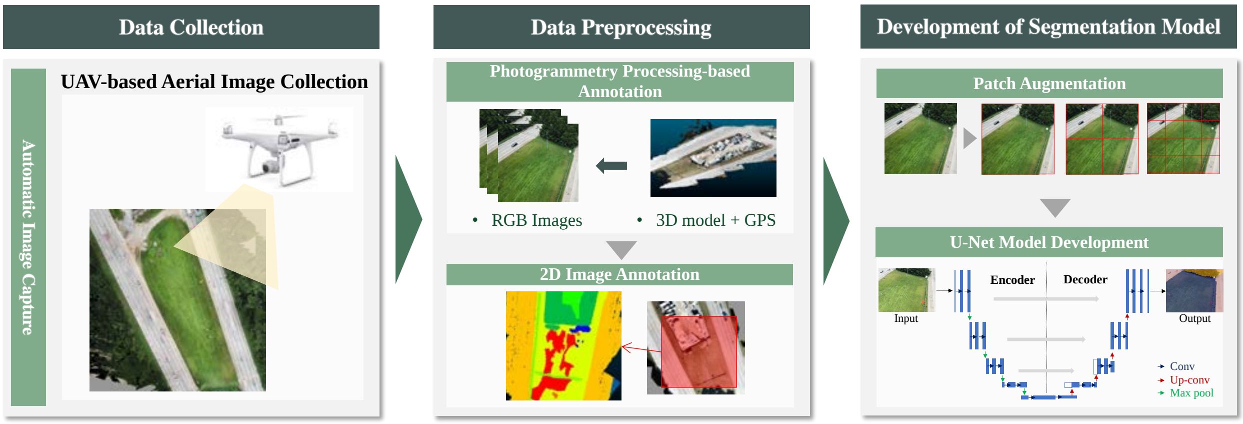
3 Methodology

This study proposes an overall lawn condition classification framework leveraging UAVs equipped with camera sensors. Our proposed framework aims to collect a large volume of dataset for lawn mowing condition monitoring and develop a segmentation algorithm to classify the vegetation conditions into three classes of maintenance personnel's interests: mowed, un-mowed, and bared spot, as illustrated in Figure 1. We defined a *mowed* condition as regions where grass height is less than 6 inches, an *un-mowed* condition as regions where grass height is over 6 inches, and a *bare spot* as regions where no grass is detected.



[Fig 1] Examples of lawn conditions: (a) bare spot, (b) un-mowed, and (c) mowed lawn condition

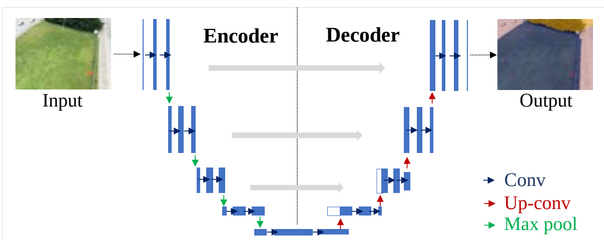
We developed a pixel-wise classification that assigns one of the lawn condition labels to each input image pixel. To automate the annotation process, we labeled the dataset using the 3D semantic segmentation method that converts 3D points into 2D images [11]. The overall framework is depicted in Figure 2.



[Fig 2] Overview of the proposed framework

3.1 Network Architecture

This section describes how our overall framework utilized the image processing algorithms to distinguish the various grass conditions. We adopt U-Net model [12], which is one of the most widely-used network architecture for semantic image segmentation. U-Net has an encoder-decoder structure with fully convolutional network, as illustrated in Figure 3. The first component of the architecture, the encoder, consists of alternating convolution and pooling operations, allowing the model to progressively increase the number of feature maps by downsampling feature maps. The decoder part of the model then upsamples the feature map and semantically projects the features learned by the encoder onto the pixel space to obtain classification results. Overall, this structure of contracting and expansive operations enables to capture of the localized segmentation of the input image and propagates the feature information to successive layers with higher resolutions. Due to this robustness to precise segmentation, this study utilized this architecture to capture fine details of the images.



[Fig 3] U-Net architecture

3.2 Data Collection and Preprocessing

3.2.1 Aerial Image Collection

For data collection, a DJI Marvic Pro and DJI Marvic 2 Pro with 4K cameras and a GPS sensor were deployed. Using a public open mobile application, DJI Go 4 and PIX4Dcapture, a licensed drone pilot set up the flying path and selected the boundaries of the vegetation areas to be collected. The drones automatically captured the images at different elevations and angles based on the flight parameter input, including flight speed, altitude, and camera angle.

3.2.2 Image Annotation Process

As the image collected from UAV inherently contains a significant area of overlap, a 2D-3D co-labeling method [11] that can label the same area only once was adopted to generate image annotations efficiently. First, the collected aerial images were processed through the Structure-from-Motion (SfM) and converted into 3D point cloud data. With respect to the manual 3D annotation, semantic labels were assigned to the original drone images by utilizing the camera projection equation. Since each image's camera intrinsic and extrinsic parameters were known after the SfM step, each 3D point can be associated with a 2D pixel. This approach offers an efficient annotation method because it can produce labels of all aerial images with a one-time manual annotation process.

3.2.3 Data Augmentation

We applied a patch augmentation strategy to make the classification model robust to the variability of vegetation conditions and Ground Sampling Distance (GSD). It processed an n×n grid over the original image and its corresponding ground truth. The image is then divided into small pieces, which are treated individually, as depicted in Figure 4. Since the input size of deep learning architecture is much smaller than that of the raw drone images, it can be expected to increase of GSD range in the training dataset.

[Fig 4] Patch-based data augmentation

4 Experiment

4.1 Data Collection

We applied our method to several highway locations in Atlanta, Georgia. Figure 5 shows the drone data collected at actual highway environments: the I-675 Highway that is maintained by the Georgia Department of Transportation (GDOT). The total area of the location is approximately 1.742 (acre), and the length along the primary axis is 150 (m). The data collections were regularly performed to monitor a wide range of grass conditions at different mowing conditions and seasons, as illustrated in Table 1. The test environment includes roads, vegetation, trees, and other objects.

The flight parameters of the UAV consist of flight elevation between 30 to 50 (m) from the ground, GSD of 0.99 – 1.64 (cm/pixels), and camera angles of 45, 60, and 90 degrees. An individual combination of the flight parameters was tested to find the optimistic performance of assessing mowing quality. A trained pilot with an observer controlled the drone's flight and monitored the process to take control of the automated flight if needed.

To acquire ground truth of lawn mowing condition, visual inspection and tape measurements were manually conducted at several locations. Based on the 15cm threshold of grass height, the ground truth of indicating mowed and un-mowed areas was generated for each data collection. These ground truths are used to validate the proposed annotation process.

[Fig 5] Test location under different field conditions collected on (1) Jul 22 and (b) Nov 11, 2021.

4.2 Development of grass condition assessment algorithm using image data

A customized dataset of over 700 aerial images was used to train the proposed U-Net, and the corresponding labeled segmented images. We employed a total of 4 classes, namely mowed, un-mowed, bare spot, and miscellaneous classes, which include non-grass regions such as roads, guardrails, signages, and trees. To improve the model's performance, we experimented with different the GSD of drone images and determined the final input image size. That is, after downsampling to the deep learning input size, the GSD of 20~25cm/pixels, and can be prepared by dividing each original image into four split images in our flight parameters.

Finally, the dataset was divided for training, validation, and testing. The validation dataset was used to tune the hyperparameters of the model that maximize the model accuracy and minimize the sparse categorical cross-entropy loss function. The validation loss is monitored to ensure that there is no overfitting or underfitting during the training process.

4.3 Experimental Evaluation

For evaluation, we produced pixel-wise classification results for input images. To quantitatively validate the performance of our proposed classification framework, we utilize a metric of accuracy, precision, recall, and F1 score, which is the dominant evaluation criteria for the image classification task.

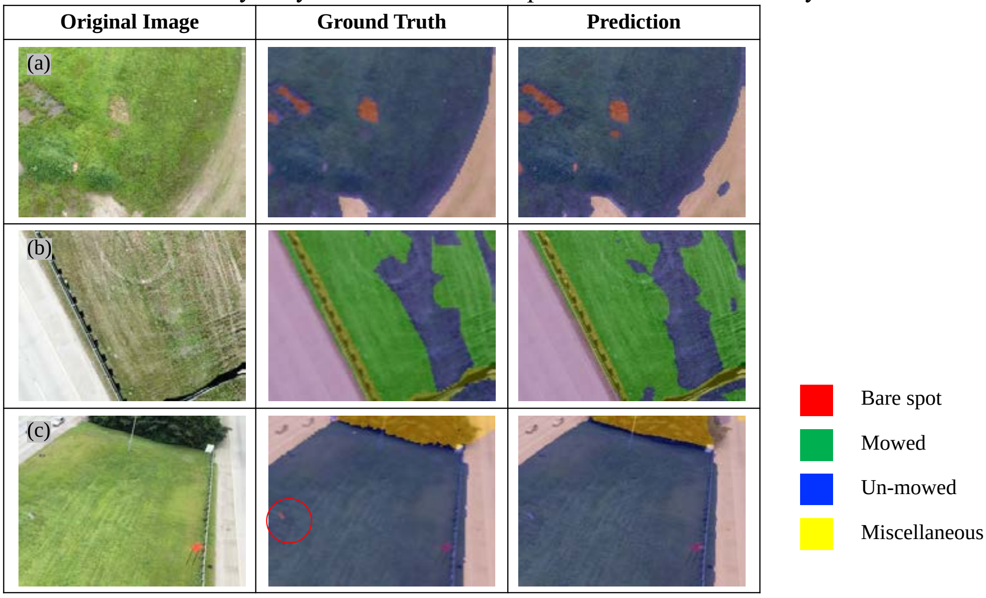
5 Results

5.1 Experiment Results

We analyzed the performance of our proposed framework. Figure 6 shows the examples of classification results. The green, blue, and red color legends indicate the lawn conditions of mowed, un-mowed, and bare spots, respectively.

We tested our model in different datasets of grass conditions (Table 1) to check the robustness of our model in terms of handling intra-variability. The overall accuracy was 87.5%, indicating that our model can produce acceptable outcomes across different grass conditions, except for data collected on Nov 11, which had precision and recall values lower than 80%. We believe that this low performance is attributed to the distinctive characteristics of this dataset due to the new construction of building in the test location as illustrated in Figure 5. This resulted in drastic changes in the features of the images.

In addition, we evaluated our model in terms of different classes. Table 2 shows the accuracy results for each class. Compared to other classes, bare spot scores the lowest accuracy by a large margin. This is due to class imbalance, the lack of a bare spot in the whole dataset. Figure 6 (c) illustrates the relatively very small area of bare spot compared to other grass condition classes. Therefore, the failure of spotting a small bare spot impacted the overall accuracy a lot.



[Fig 6] Image segmentation results

Table 1. Overall Performance of the Proposed Framework

Collection Date	May 14, 2021	July 22, 2021	July 30, 2021	November 11, 2021
Accuracy	0.872	0.826	0.891	0.912
Precision	0.876	0.768	0.815	0.411
Recall	0.876	0.730	0.874	0.401
F-1	0.811	0.729	0.823	0.405

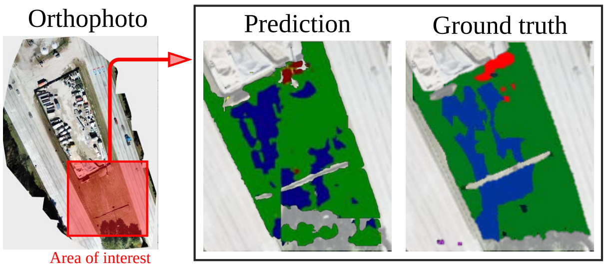
Table 2. Performance of Lawn Mowing Condition Class

	Mowed	Un-mowed	Bare	Misc
Accuracy	0.989	0.994	0.995	.985
Precision	0.945	0.888	.473	.766
Recall	0.959	0.908	.497	.988
F-1	0.946	0.895	.468	.863

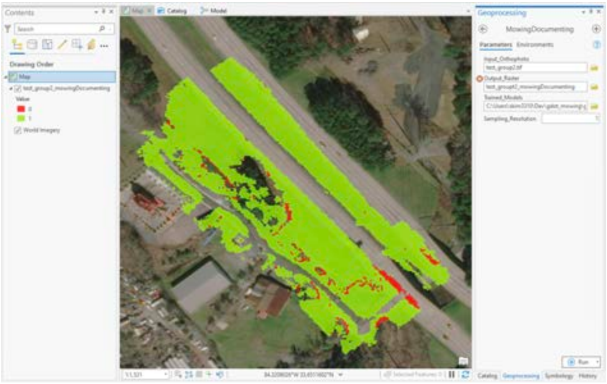
5.2 Geo-Referencing and Digitization

In addition, we mapped the image-level classification results onto the geo-referenced highway site to provide a more user-friendly, informative results. Each orthophoto was produced by stitching and smoothening whole drone images into the whole site was deployed.

To do so, an orthophoto of the highway site was firstly split into multiple patches of the interpretable size in the trained model. Then, each patch was inferenced and reverted into the original position while preserving coordinate reference systems. Fig 7 shows an output of being geo-referenced and digitized classification map in a raster data format. This enables lawn monitoring managers to visualize mowing quality and make decision upon Geographic Information Systems platform. Fig 8 shows the final output in the GIS platform.



[Fig 7] Result with Geo-referencing and Digitization



[Fig 8] Implementation on the GIS platform

6 Discussion

The results demonstrated the potential benefits of our proposed system, there are several limitations to overcome for practical issues for real-world application. First, for practical purposes, orthophoto provide more useful information than a single drone image because it provides the overview with the location information. However, the accuracy of orthophoto classification is lower compared to that of the original aerial image. This is because generating orthophoto causes some texture and color information loss while multiple images are smoothed out and switched together. As a result, orthophoto suffers from low-resolution, incompleteness, lost texture information.

Secondly, the capacity of our proposed model is limited in terms of the highway assets coverage. Several vertical highway assets such as poles are not well detected.

In the future works, these above-mentioned issues can be mitigated by the integration of 3D model. Aerial images from tilted views will be utilized to generate 3D highway scene, and the 3D model will be overlapped with 2D images to further train the classifier.

7 Conclusion

In this study, we showed the feasibility of aerial image-based computer vision algorithms for the purpose of lawn condition monitoring for highway assets. Our proposed framework expanded the application of computer-vision and drone technology. With the user-friendly GIS-based visualization, it will support the lawn monitoring personnel to easily verify the mowing performance of the contractors and better manage the highway assets condition.

Acknowledgements

The work reported herein was supported by the National Science Foundation (Award #: OIA-2040735) and the Georgia Department of Transportation (GDOT) project RP 20-09/T.O. 2014-99. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF or GDOT.

References

[1] R.W. Eck, H.W. McGee, United States. Federal Highway Administration. Office of Safety, Vegetation Control for Safety: A Guide for Local Highway and Street Maintenance Personnel: Revised August 2008, 2008. https://rosap.ntl.bts.gov/view/dot/42594 (accessed February 21, 2022).

[2] K.L. Rens, T.J. Wipf, F.W. Klaiber, Review of Nondestructive Evaluation Techniques of Civil Infrastructure, J. Perform. Constr. Facil. 11 (1997) 152–160. https://doi.org/10.1061/(ASCE)0887-3828(1997)11:4(152).

[3] B.F. Spencer, V. Hoskere, Y. Narazaki, Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring, Engineering. 5 (2019) 199–222. https://doi.org/10.1016/j.eng.2018.11.030.

[4] F. López-Granados, J. Torres-Sánchez, A.-I. De Castro, A. Serrano-Pérez, F.-J. Mesas-Carrascosa, J.-M. Peña, Object-based early monitoring of a grass weed in a grass crop using high resolution UAV imagery, Agron. Sustain. Dev. 36 (2016) 67. https://doi.org/10.1007/s13593-016-0405-7.

[5] A. Brilhador, M. Gutoski, L.T. Hattori, A. de Souza Inácio, A.E. Lazzaretti, H.S. Lopes, Classification of Weeds and Crops at the Pixel-Level Using Convolutional Neural Networks and Data Augmentation, in: 2019 IEEE Lat. Am. Conf. Comput. Intell. -CCI, 2019: pp. 1–6. https://doi.org/10.1109/LA-CCI47412.2019.9037044.

[6] A.I. De Castro, J. Torres-Sánchez, J.M. Peña, F.M. Jiménez-Brenes, O. Csillik, F. López-Granados, An Automatic Random Forest-OBIA Algorithm for Early Weed Mapping between and within Crop Rows

Using UAV Imagery, Remote Sens. 10 (2018) 285. https://doi.org/10.3390/rs10020285.
[7] T. Sun, Z. Di, P. Che, C. Liu, Y. Wang, Leveraging Crowdsourced GPS Data for Road Extraction from Aerial Imagery, Proc. IEEE CVF Conf. Comput. Vis. Pattern Recognit. (2019).
[8] S. Kim, J. Irizarry, Knowledge-Based Considerations for Developing UAS Operational Procedures on Infrastructure and Construction Task Environments, (2020) 268–277. https://doi.org/10.1061/9780784482865.029.
[9] Y. Pi, N.D. Nath, A.H. Behzadan, Detection and Semantic Segmentation of Disaster Damage in UAV Footage, J. Comput. Civ. Eng. 35 (2021) 04020063. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000947.
[10] S. Bhatnagar, L. Gill, B. Ghosh, Drone Image Segmentation Using Machine and Deep Learning for Mapping Raised Bog Vegetation Communities, Remote Sens. 12 (2020) 2602. https://doi.org/10.3390/rs12162602.
[11] Y. Yajima, M. Kahoush, S. Kim, J. Chen, J. Park, S. Kangisser, J. Irizarry, Y. Cho, AI-driven 3D Point Cloud-Based Highway Infrastructure Monitoring System using UAV, in: 2021.
[12] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, Int. Conf. Med. Image Comput. Comput. Assist. Interv. (2015).