Crop Field Segmentation

CSCE 5222 Feature Engineering

Group 7:

1. Pradhyumna Poralla
2. Akruti Khurana
3. Shiva Karpe

Abstract

The main goal of the project is to segment all the crop fields from the satellite images that are collected from the Satellite Aerial Images from the U.S. Department of Agriculture (USDA); about nine different photos have been set to undergo several segmentation techniques, and we discuss all the methods and their significant drawbacks in our particular use cases where we also discuss the best methods gave us better results, a. Ween discuss the accuracy and precision of technique and detailed description on the formation of our evaluation sets to compare the accuracy of our predicted ground truths and our expected ground truths followed the methods and tools used to form required ground truths.

Keywords

The main keywords mentioned throughout the report are Crop fields, non-crop fields, trees, and ground truth, which are the most repeated words.

Introduction

Segmentation is the most popular word used in the field of computer vision. This is the proper term that has significant meaning, for it that is to differentiate the background from the ground truth, where ground truth means to be the region of interest in the image. Our problem statement is to segment the crop fields from the given dataset of nine satellite images that are collected from the U.S. Department of Agriculture (USDA); Where the images are photos of open areas from some unknown location in the USA where the fields that are captured from satellites are images of open fields with cultivating crop fields and bare land such as non-cultivating lands. The photos also have different areas with roads and partitions between the fields and a large area of trees. Small regions of residential and water bodies are also in images with clear boundaries in the pictures.

Problem Statement

Our problem statement is that segmenting crop fields means focusing on two cases of ground truth in
Case 1: We focus on crop fields along with the trees in the images where crop fields and trees/**forest.**
Case 2: We focus on crop fields **without** considering trees/**forests**; together, we considered to be cultivating crop fields where they are regarded as our region of interest, which is the ground truth of our project, where the rest is considered to be a non-cultivating field that is included with residential

areas in the image and all the dry fields in the pictures along with the water bodies that are present in few of photos. Our goal is to prepare the mask of the ground truth, which we are focusing on cultivating crop fields as described above. As we mentioned, we also focus on the trees as our ground truths. In the image, the other parts are set with zero intensity and were to be put in black in color.

Cite: Field Segmentation

Tasks

1. Extract Ground Truths
2. Prepare Ground Truth for Evaluation
3. Calculate Accuracy and Precession

Region of Interest in Images:

The above image is the minimal representation of all the areas or regions present in the image, where each element is represented in a different format. The roads in the image are illustrated in red, whereas a few similar pathways or roads in the above picture aren't represented for understanding purposes. However, those boundaries and roads aren't considered ground truths, and there are dry lands described in black color where orange stands for water bodies.

The rest of the green region in the image represents all the actual ground truth.

In the above image, the area with a white box of "1" stands for a part of ground truth cropland, where it has boundaries and other crop fields in its neighbor. We can see that it is one such crop field in the image where we have many.

Case 1: Description of Ground Truth with Considering Forest:

In the above image, the area with a white box of "F" stands for a part of the ground truth representing the bushy trees/Forest, which are also included in the ground truth.

Case 2: Description of Ground Truth without Considering Forest:

In the above image, the area with a white box of "F" stands for forest or trees, which are not a part of the ground truth representing the bushy **trees/Forest**, which **are not included** in the ground truth.

Methodology For Case 1 (Considering Forest in Ground Truth)

Per our problem statement, we have gone through many different methods and techniques and had different results. Where each of them gives us different results as part of experimentation. Finally, K means solving the major problem with promising results per requirement.

At the first stage, our segmentation function requires an image file path as input where the function reads the image as beginning as the first stage of the process.

At stage one, the image is read where it is always represented in RGB format. In the next stage, where the image is converted into CIELAB color space where the image is converted with an inbuilt conversion function of MATLAB, it expresses color as three values: *L** for perceptual lightness and *a** and *b** for the four of human vision: red, green, blue, and yellow.

At stage three, a dimension of the L*a*b image matrix, a MATLAB double-type matrix, is converted into a single matrix using im2sigle. Where the single matrix undergoes the segmentation process, that is, K means segmentation.

Next step, the single matrix is set to for K clusters by using '**imsegkmeans**,' the inbuilt function in MATLAB. We chose to form two clusters where K is 2 (Elbow Method: When it comes to determining the optimal number of clusters, the elbow method comes into the picture. Here for different values of k, we calculate the Within-Cluster-Sum of Square Errors (WSS) and then choose the k for that WSS value for which it first starts to diminish. We get a plot of WSS v/s k; it is visible as an elbow.).

Where a pixel-labeled matrix is formed as the stage output with vales filled with cluster numbers representing 1s and 2s; we process this for preparing masks in the next stage.

This is the almost final stage of forming the mask to segment the crop field in the image; as discussed before this stage, the labeled matrix is formed into a logical matrix. All the 2's are converted to logic one and others to logic 0 if the labels in the labeled matrix are in reverse where 1's represent our ground truth. In such a case, we decide on the max label method, where we count the max number of labels in the label matrix where our ground truth occupies max pixels. We consider that the maximum pixels present in the label matrix are converted into logic 1 and others to logic 0. A logical matrix is formed by applying these few logical operations on a labeled matrix, and we create an output mask.

Where the mask is the final output, it represents the ground truth with the white pixels, and all the non-considered are with black pixels. As mentioned before, it is the logical matrix where the logical matrix is converted to "uint8" type where general image type in MATLAB and make pixel to pixel multiplication to show the ground truths in the image clearly.

Citation for methodology: K Means

Methodology For Case 2 (Not Considering Forest in Ground Truth)

As per the problem statement, we consider all the crop fields in the image, and we do not consider forest in the ground truth in this method.

In this method, we also use the K-means technique, as mentioned in the method of case 1, but technically we do some preprocessing in the image while that is set into k means function.

As the First stage and part of preprocessing of an image, we use a few morphological operations on the image “I” as part of initial processing; we try to remove all the foreground elements in the images. That is, we enhanced all the non-green features in the image. For that, we used the morphological operation “**imopen**” in MATLAB, where we used a disk-shaped structuring element with a radius of 5 and formed an image “I2”.

In the Second stage, we subtracted the output of the first stage with the original image “I” with “I2.” We formed an output “I3” with all the edges in the image. To enhance the image contrast of “I3” by saturating 1% of the image, we have used “**imadjust**” on image “I3” and formed the “I4” image.

In the Third stage, we used the “**imbinarize**” inbuilt function of MATLAB on image “I4”. We formed a logical matrix, “I5” where this “I5” image is an inverted mask of our required ground truths. By inverting the “I5” image, we form an operated mask that is “**mask_op**.”

In the Fourth stage, we have to use the same functionality or the method that one which we used in the method of case 1, that is, K-means we read the original image in this method that is the same as “I” the image “I” is converted into CIELAB color space where L*a*b image is formed and, a MATLAB double-type matrix, is converted into a single matrix using im2sigle. Where the single matrix undergoes the segmentation process, that is, K means segmentation.

The image is transformed into the inbuilt function “**imsegkmeans**” of MATLAB. We chose the K value to be 2 (Elbow Method: When it comes to determining the optimal number of clusters, the elbow method comes into the picture. Here for different values of k, we calculate the Within-Cluster-Sum of Square Errors (WSS) and then choose the k for that WSS value for which it first starts to diminish. We get a plot of WSS v/s k; it is visible as an elbow.).

Where a pixel-labeled matrix is formed as the stage output with vales filled with cluster numbers representing 1s and 2s, we process this for preparing masks as we have detailed explained in the method of **case 1.**

The mask that is formed using K means be “**mask_kmean,**” where now, at the final stage, we use “**mask_op**” and do some logical operations in this stage where we do logical and operation between “**mask_kmean**” and “**mask_op**” to form the final “**output_mask**” that represents all the segments as per the ground truths which we require the accuracy is calculated with using a confusion matrix.

The Additional discussions of the method are discussed below.

Output Discussion of Case 1(considering forest in ground truth):

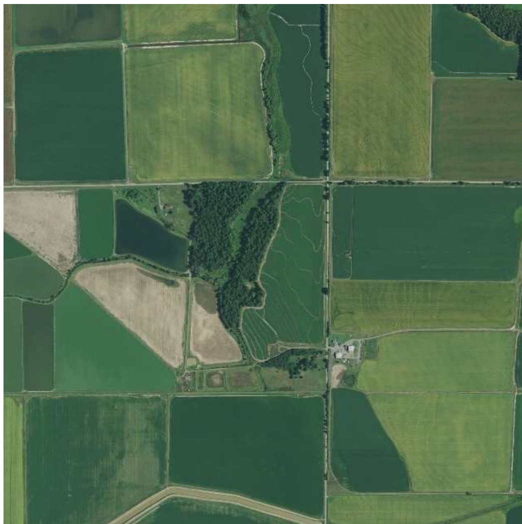


Fig 1: Normal Input Image

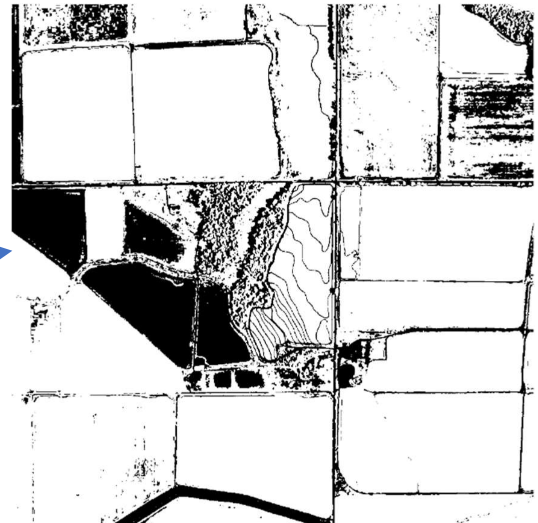


Fig 2: Output Mask

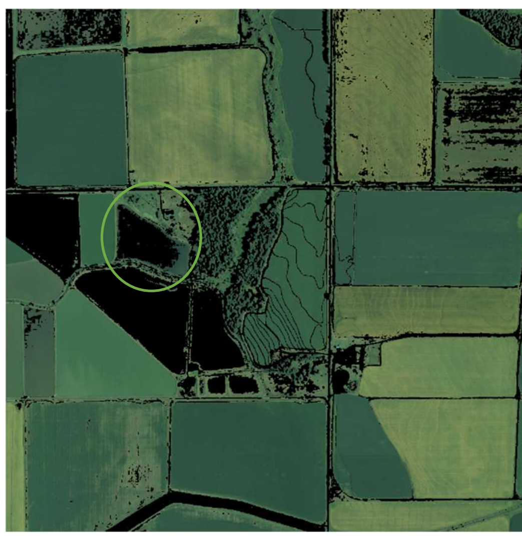


Fig 3: Output Image Applied Mask

From the above images, consider Fig 3. This image represents the ground truth we mentioned above in the ROI section, where we could have done better, but we misclassified a few areas, as mentioned. As an example, in Fig 3, observe the circle. The water pound is partially blacked out, as it is not part of our ground truth, where such misclassifications might give us low accuracy.

Similar misclassifications are also made at a few points where the model ideally needs to segment the required ground truths.

In the ideal case, the evaluation matrix for images is made where the mask is prepared by using various techniques we discuss in upcoming topics in detail where we compare the ideal evaluation mask and the output mask individual image to check the accuracy of the prepared crop segmentation method that is mentioned above.

Output Discussion for Case 2 (not considering forest in ground truth):

Fig a: Original Image

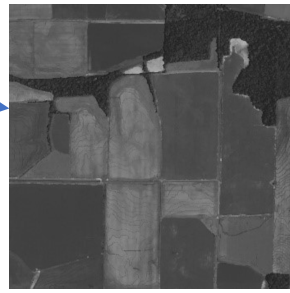


Fig b: RGB to Gray

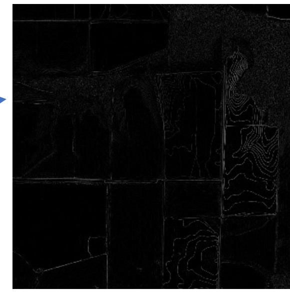


Fig c:After Morphological Transform

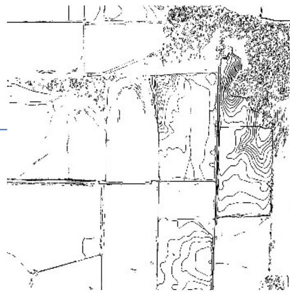


Fig f: Inverted (e)

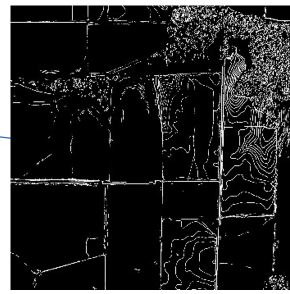


Fig e: Binarized Output

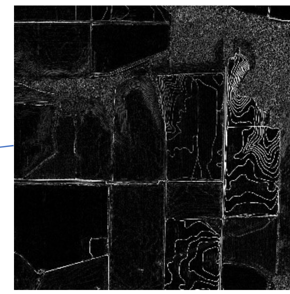


Fig b: Subtract (a-c)

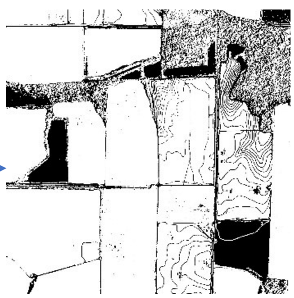


Fig g: Output K-means

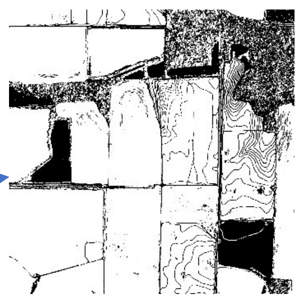


Fig h: Final Output with (e logical (&) g)

The above Images are the flow of the method of case 2, which we have proposed, where the output is formed as per the problem statement of case 2, where the forest and non-crop fields and roads are excluded from the image.

In the ideal case, the evaluation matrix for images is made where the mask is prepared by using various techniques we discuss in upcoming topics in detail where we compare the ideal evaluation mask and the output mask individual image to check the accuracy of the prepared crop segmentation method that is mentioned above.

Evaluation Discussion for Case 1 (considering forest in ground truth):

The ideal ground truths per problem statements are made using different techniques, which are prepared to compare with the proposed output.

Initially, each image is cropped into 16 equal images.

Fig 4: Cropped Images

Each piece image is used to prepare ideal ground truths masks using inbuilt image segmentation tools of MATLAB. “Draw ROI (Polygon)” and “Flood Fill” are the tools to draw the ground truths. The “Draw ROI” tool is often used, but “Flood Fill” is rarely used for complex elements to draw.

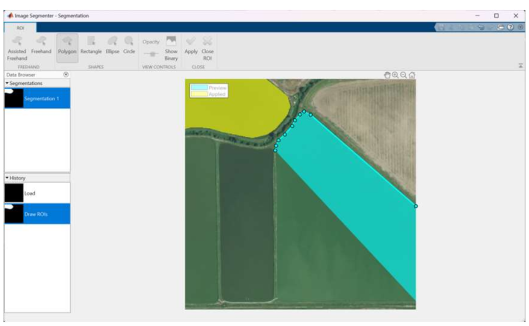


Fig 5: Screen Shot of “Draw ROI” tool

Every image part ground truth mask is compared with the same part of the output mask.

The mask is compared with the confusion matrix, and accuracy and precision are calculated for each part of the image. They are averaged to the accuracy of the complete image.

Fig 6.1: Part of Image

Fig 6.2(a): Output Mask

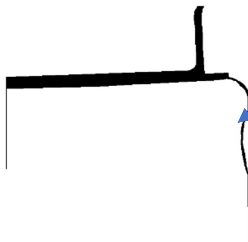


Fig 6.3(b): Evaluation Mask

Fig 6.3(a): Image with Output Mask

Fig 6.3(b): Image with Evaluation Mask

Consider Image L97a for Case 1 (considering forest in ground truth):

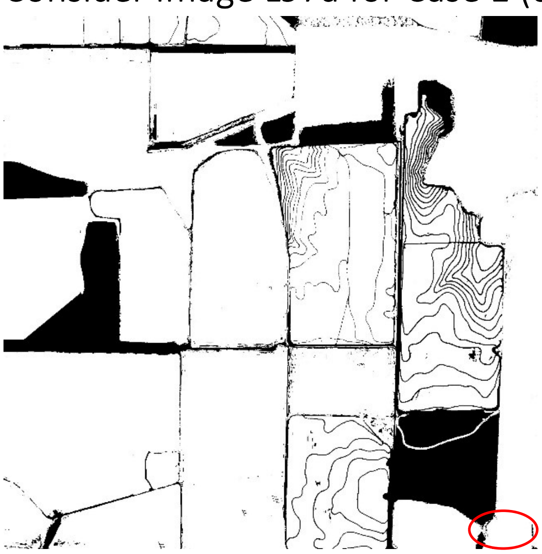


Fig 6.4(a): Ground Truth

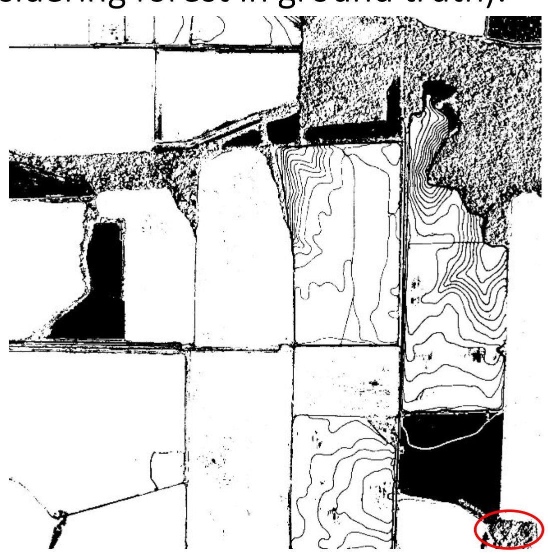


Fig 6.4(a): Output

The above images are ground truth, and the output image is as per ground truth image here is considered with the forest and all the crop fields where the bushy trees in the output image are still yet not classified as per the ground truth where the model fails in such regions where the forest are not classified adequately in output where this is the drawback of our proposed method.

Confusion Matrix for the above images:

	Positive	Negative
Positive	11241	3899
Negative	4400	242604

The calculated accuracy for the above confusion matrix is 96.83%.

The calculated precession for the above confusion matrix is 98.22%

As mentioned in the above paragraph, below the images, the ground truth and output as per the confusion matrix shown in the above example are 3899 true negatives and nearly 4400 false positives.

Based on the above images, we represent a red circle at the images' edge, clearly stating a part of the misclassification and representing the true negatives in the confusion matrix. The confusion matrix forms such misclassifications as true positives and negatives.

Evaluation Discussion for Case 2 (not considering forest in ground truth):

To create the mask for the ideal case that is ground truths, we have not cropped the images as per the previous evaluation discussion of the case2. In this, we have created a mask of ground truths directly on the image where each image is taken and by using inbuilt image segmentation app in MATLAB we perform or used two techniques as we have mentioned for case1.

"Draw ROI (Polygon)" and "Flood Fill" are the tools to draw the ground truths. The "Draw ROI" tool is often used, but "Flood Fill" is rarely used for complex elements to draw.

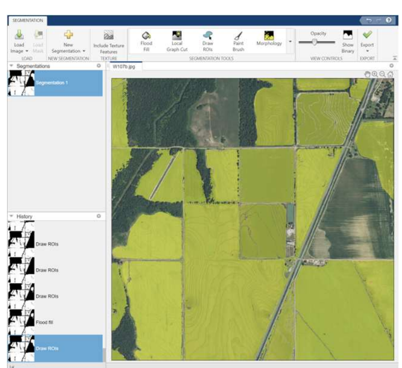


Fig 7: Screen Shot of "Draw ROI" tool

Every image ground truth mask is compared with the same image output mask.

The mask is compared with the confusion matrix, and accuracy and precision are calculated for each part of the image. They are averaged to the accuracy of the complete image.

Consider Image L97a for Case 2 (not considering forest in ground truth):

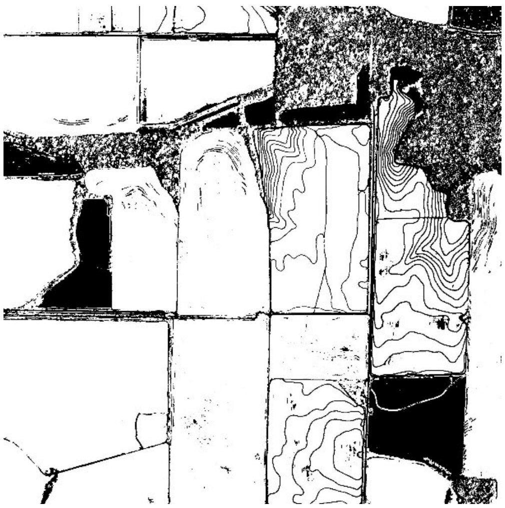


Fig 7.1 (a): Output of Proposed Technique

Fig 7.2 (b): Prepared Ground Truth

The above images are ground truth, and the output image is as per ground truth image here; we do not consider the forest as part of the ground truths where we compare both the masks using a confusion matrix.

Confusion Matrix for the above images:

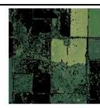
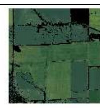
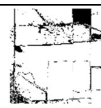
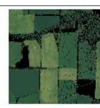
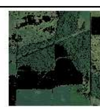
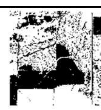
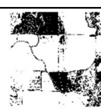
	Positive	Negative
Positive	873343	289407
Negative	81150	2950404

The calculated accuracy for the above confusion matrix is 91.17%.

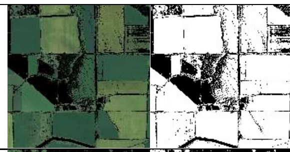
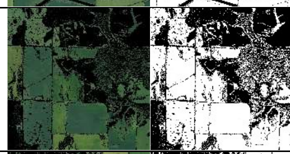
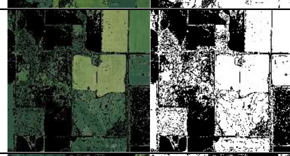
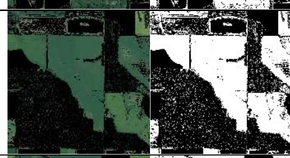
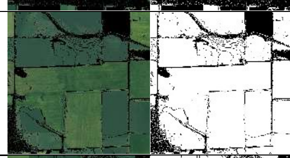
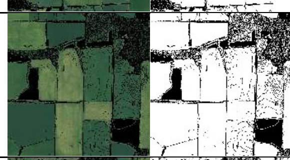
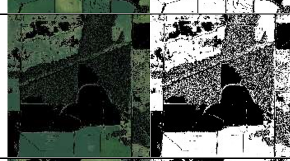
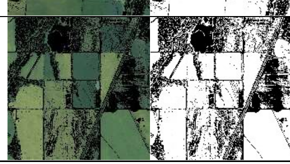
The calculated precession for the above confusion matrix is 97.32%

In the above images, a few parts are misclassified as per the ground truth.

Results of Case 1 (considering forest in ground truth)

Image	Ground Truth Obtained	Accuracy	Precession
	 	0.8869	0.9254
	 	0.7964	0.7710
	 	0.8709	0.8462
	 	0.8766	0.8538
	 	0.9259	0.9608
	 	0.9318	0.9529
	 	0.7854	0.9494
	 	0.8318	0.8425
	 	0.8679	0.8634

Results of Case 2 (not considering forest in ground truth)

Image	Ground Truth Obtained	Accuracy	Precession
		0.9046	0.9347
		0.8314	0.8789
		0.8122	0.8717
		0.9007	0.8969
		0.9556	0.9692
		0.9117	0.9732
		0.7847	0.8148
		0.8970	0.9202
		0.8169	0.8711

Conclusion

As the concluding point, we have tried to achieve the maximum best results. But as per the results shown at the top, the results of case 2, ground truths without considering forest, are better than the case 1 results. Still, a few of our ground truths couldn't turn out as we expected through the function or the model we developed to segment crop fields from the pictures. We end this up with pretty good and high accuracy.

References

Andriyanov, N., & Dement'ev, V. (2018). Application of mixed models of random fields for the segmentation of satellite images. Collected Selected Papers from the IV International Conference on Information Technology and Nanotechnology.

https://doi.org/10.18287/1613-0073-2018-2210-219-226

G. Chamundeswari, Prof. G. Pardasaradhi Varma, Prof. Ch. Satyanarayana (2012). An Experimental Analysis of K-means Using Matlab. International Journal of Engineering Research & Technology (IJERT) Volume 01.

https://www.ijert.org/research/an-experimental-analysis-of-k-means-using-matlab-IJERTV1IS5347.pdf