Acta Scientiarum

http://periodicos.uem.br/ojs
ISSN on-line: 1807-8664
Doi: 10.4025/actascitechnol.v44i1.55708

COMPUTER SCIENCE

Tensor based statistical segmentation of green vegetation canopy images

Umit Cigdem Turhal[1*] and Can Dagdelen[2]

[1]Electric and Electronics Engineering Department, Engineering Faculty, Bilecik Seyh Edebali University, Pelitözü Mah, Fatih Sultan Mehmet Bulvarı, 27, 11230, Bilecik Merkez, Bilecik, Turkey. [2]Postgraduate Education Institute, Electric and Electronics Engineering Department, Bilecik Seyh Edebali University, Bilecik, Turkey. *Author for correspondence. E-mail: ucigdem.turhal@bilecik.edu.tr

ABSTRACT. The increase of developments of electronic and computer has been resulted with the commonly use of these technologies in everyday life in many applications. One of these applications is emerged in agricultural applications as precision agriculture studies. The use of technology in agriculture has lots of benefits such as energy saving, yield increase, time saving and etc. In this study, a novel learning based, pixel-wise segmentation method which uses Common Vector Approach (CVA) for the first time in the literature for segmentation is proposed. In the proposed method, first of all color regions belong to both vegetation and soil are manually cropped and then three different color space representations such as HSV, Lab and Luv of RGB images for each color region are encoded as 3rd color tensors. Then by unfolding the color tensor in the mode-3 direction, 2-D color matrix is obtained. The columns of this 2-D color matrix are the $6x1$ dimensional vectors that include (H, S a, b, u, v) components of HSV, Lab and Luv color spaces of an image pixel, and each column vector is accepted as an object. By applying CVA in object space consisting of column vectors of 2-D color matrix, a common color vector which represents common properties of that color region is obtained and used for segmentation purposes. In the experimental studies two different datasets proposed for open computer vision tasks in precision agriculture before in the literature are used. Three different experimental studies are performed for different dataset combinations in terms of training set and the test set. The performance of the proposed method has been compared with the performance of a deep learning method, Convolutional Neural Networks (CNN) based semantic segmentation method. In all of the three experimental studies proposed method achieves extremely high performance according to CNN, especially in the second and in the third experimental studies where dataset combinations include the two of the datasets.

Keywords: Image processing; 3rd order tensors; statistical pattern recognition; precision agriculture; machine learning; vegetation segmentation.

Received on September 8, 2020.
Accepted on May 17, 2021.

Introduction

Advances in electronic technology have led to the active use of image processing and computer vision in many areas of our daily lives. Thus automatic systems are developed that will be the solution to many problems involving visuality such as precision agricultural applications. Through to agricultural applications performed with image processing and computer vision studies, automatic systems can be developed that can achieve reliable results while saving time and input and service costs. For precision agricultural applications, images covering large areas within short time are acquired by remote sensing systems such as satellite and airplanes imagery (Zhang & Kovacs, 2012). However these technologies are limited by the low temporal and spatial resolution and require special and costly hardware that is not appropriate for the farmers' practical usage. Therefore in recent studies, images with higher resolutions captured by using unmanned aerial vehicles (UAVs) imagery and even mobile phone imagery have been introduced (Nex & Remondino, 2014). The hardware requirement used to obtain multiple spectral images is more costly and it is not so easy to apply according to the hardware of RGB imagery. However an RGB image that includes the visible light region is obtained very easily using consumer-level cameras, even mobile phone imagery.

Segmentation which is a basic step in many image analysis applications (Bao, Javanbakhti, Zinger, Wijnhoven, & De With, 2013; Parra, Boutin, & Delp, 2017) is also a required and an important image processing step in computer vision based precision farming applications that generally focus on crop health monitoring and weed management (Slaughter, Giles, & Downey, 2008;). In such these studies, segmentation is for detection of vegetation in the images of agricultural areas. There are basically two types of approaches used for segmentation. While the first type of these approaches consists of color index-based and threshold-based methods, the second type consists of learning-based classification approaches (Hamuda, Glavin, & Jones, 2016; Wang, Zhang, & Wei, 2019).

In color index-based methods, the vegetation extraction easily can be performed in multispectral imagery by using the vegetation indexes such as NDVI, Difference Vegetation Index (DVI), Normalized Red (NR), Normalized Green (NG), and etc. (Mulla, 2013). In RGB imagery, by using the Red (R), Green (G) and Blue (B) color components some formulations are constructed as color indexes. These formulations are defined as the Normalized Difference Index (NDI), the Excess Green Index (ExG), the Excess Red Index (ExR) and the Color Index of Vegetation Extraction (CIVE). In threshold index based methods the intensity value of each pixel in gray scale images is classified by comparing with the one or more predetermined threshold value (Osunmadewa, Wessollek, & Karrasch, 2015; Hassanein, Lari, & El-Sheimy, 2018; Kaur & Kaur, 2014).

In the learning-based approaches, both unsupervised and supervised machine learning algorithms were investigated for vegetation segmentation. In the literature as unsupervised methods that do not require any training phase, K-means clustering (Kumar & Prema, 2016); Prema & Murugan, 2016) and particle swarm optimization (PSO) based K-means clustering (PSO) (Bai et al., 2014) can be found. As supervised methods that previously require a training phase, decision trees (Guo, Rage, & Ninomiya, 2013; Riegler-Nurscher, Prankl, Bauer, Strauss, & Prankl, 2018) artificial neural networks (Potena, Nardi, & Pretto, 2017) and statistical based (Ruiz-Ruiz, Gómez-Gil, & Navas-Gracia, 2009; Zheng, Shi, & Zhang, 2010) methods can be found. In supervised segmentation methods, the problem of segmentation of green vegetation parts of an image from background can be supposed to be a two-class classification problem. Segmentation includes a feature extraction step to define object classes, and the classification is performed based on these properties. In fact supervised learning algorithms are pattern classification algorithms that require a training step. In this step the classification model is constructed which will be used in the testing step. Therefore in supervised classification datasets are splitted into two at the beginning of the process as to be one of them is training set and the other one is the test set. Compared to threshold-based methods, these methods can provide promising classification results despite of their more computational cost with an appropriate number of training examples.

In the proposed study, a novel segmentation method based on CVA which is a supervised learning algorithm in which color regions of image which will be segmented are encoded as 3^{rd} order color tensors, is proposed. By unfolding the color tensor in mode-3 direction, the 2-D color matrix is obtained. Each column of this matrix are composed from the vectors which includes the color information in the three different color spaces of HSV, Lab and Luv of a pixel of the original RGB image. Then CVA is applied to this color matrix and a common pixel vector which represents the common properties of that color region is obtained and it is used for segmentation.

In the experimental studies it is used two datasets which have been previously proposed in the literature for education purposes in precision agriculture applications. The detailed explanation about these datasets is given in the following section. In the study different experimental studies are performed and their results are evaluated, for varying dataset combinations in terms of the training set and the test set. In each of the experimental studies in the training step, using the training images the segmentation model is constructed. Then in the testing step the constructed model is evaluated using the test images. In this segmentation method training images are obtained by manually cropping different color regions from the green vegetation canopy images. Thus in the training step for both of vegetation and soil, a common vector which represents the common properties of pixel group that belongs to each color region are obtained. These common vectors are defined in the indifference feature subspace of each color region. In testing step, for each of its pixels of a RGB test image, the $6x1$ dimensional color vectors are obtained such as it is in the training step. Than in order to detect the class of each pixel in the test image, it is projected onto the indifference feature subspace belong to both classes' vegetation and soil. The projected vector is called remaining vector of that test pixel vector on that feature subspace. On each feature subspace distances between the common vector of that subspace and the remaining vector of the test pixel are calculated and the test pixel is assigned to the class that the distance is minimum thus segmentation is performed.

The contributions of this paper can be summarized as follows:

A novel segmentation algorithm which uses CVA for segmentation for the first time in the literature is proposed.

While semantic segmentation accepts each image as a single object the proposed algorithm accepts each $6x1$ dimensional image pixel vector that including color components of HSV, Lab and Luv as an object. This difference provides an important superiority to the proposed novel segmentation method in this paper over semantic segmentation methods: While semantic segmentation methods require a huge amount of training data in the model construction step, the proposed method can achieve an extremely high performance even there is only one training image in the dataset. This is the remarkable advantage of the proposed segmentation method over semantic segmentation methods.

While semantic segmentation methods require ground truth images in order to perform segmentation the novel proposed method doesn't require a ground truth image, as it extracts a feature vector learning from the image and uses it for segmentation.

It is less sensitive to the environmental factors as it uses HSV, Lab and Luv color spaces in defining the image data. Besides, it has the property that it can be generalized into different segmentation applications based on using different features instead of using the color information.

While semantic segmentation requires two different steps for segmentation and classification the novel proposed method can perform both the segmentation and the classification at the same time with a single process.

Hardware requirements are less than as in a deep learning semantic segmentation.

Remaining of the paper is organized as follows. In the second section the proposed segmentation method is explained in detail. In the third section the experimental studies are given and then the results are evaluated in the conclusion.

Proposed segmentation method

In this paper a novel learning based segmentation method is proposed based on CVA in which images are encoded as 3^{rd} order tensors. Image tensors include the color information of the three different color spaces of HSV, Lab and Luv instead of RGB images in their frontal slices. In the following basic definitions of tensors are given and CVA is explained briefly.

A. Basic tensor definitions

The tensors, which can be defined as a multidimensional array, consist of a combination of different vector spaces, each with its own coordinate system. Three basic tensor concepts are tensor order, tensor fibers and tensor slices. The order of the tensor is related to the axes number of the tensor. For example, while a scalar is a zero-dimensional tensor, a vector consisting of number sequences is a 1-dimensional tensor , a matrix consisting of vector sequences is a two-dimensional tensor, and a 3-dimensional structure consisting of matrix arrays is a three dimensional tensor and so on (Figure 1 (Kolda & Bader, 2009)).

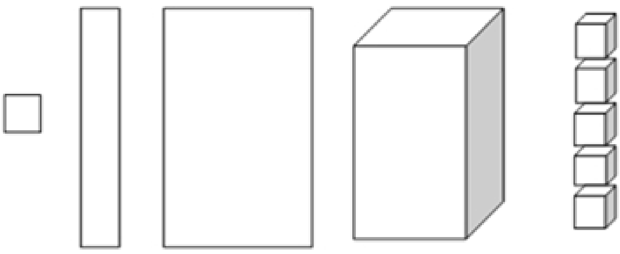

Figure 1. Left to right order of tensors (zero to fourth). Scalar, vector, matrix, 3rd order tensor and 4th order tensor.

Fibers are the higher order extension of matrix rows and columns, and are defined by fixing all indices but one. Three kinds of fibers are defined for a tensor; these are column fibers, row fibers and tube fibers as given in Figure 2. A tensor slice is defined by fixing all indices but two, and is a two dimensional fragment of a tensor. Three type of tensor slices are defined as given in Figure 3 that are horizontal slices, lateral slices and frontal slices.

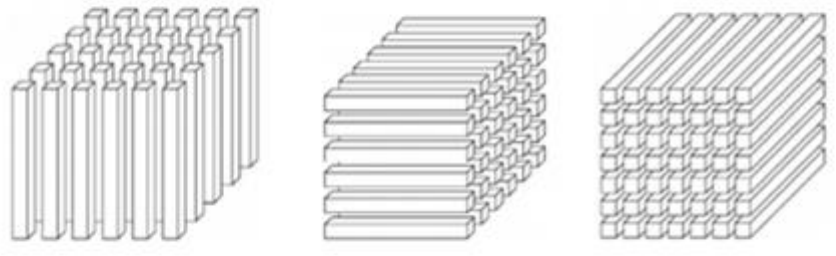

Figure 2. Tensor fibers. Column, row and tube left to right (Kolda & Bader, 2009).

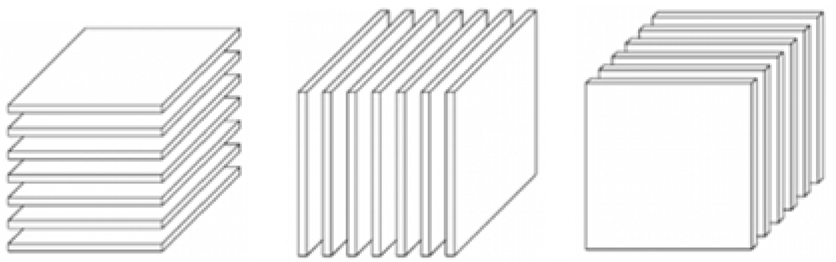

Figure 3. Tensor slices. Horizontal, lateral and frontal slices left to right (Kolda & Bader, 2009).

By unfolding a tensor into a matrix form variants of low dimensional feature subspaces can be obtained. This is also known as matricization or flattening of the tensor. Unfolding of a tensor into a matrix is achieved by constructing a matrix with its all columns include one type of the fibers of the tensor. In this process if the column fibers are used it is called mode-1 unfolding, if row fibers are used it is called mode-2 unfolding and finally if tube fibers are used it is called mode-3 unfolding (Figure 4).

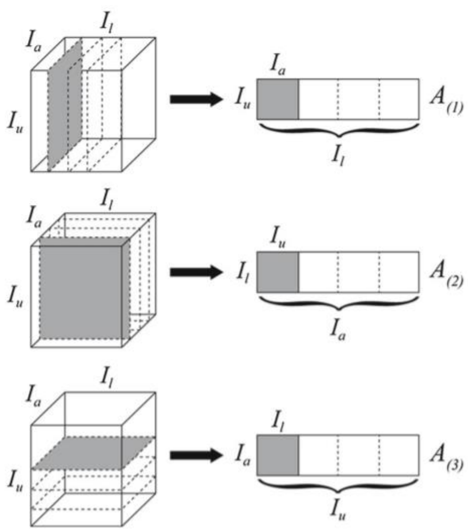

Figure 4. Unfolding a tensor (Sattari, 2015).

B. Common vector approach (CVA)

CVA is a supervised statistical learning algorithm based on the principle of finding a single and constant common vector which represents the common properties of a pattern class. It is first proposed for isolated word recognition when the sample dimension for each class was equal to or greater than the number of samples in the training set (Gulmezoglu, Dzhafarov, & Barkana, 2001). Let $x_1, x_2, \dots, x_m$ be the sample vectors that exist in the training set for each class and are supposed to be linear independent. In this case by using CVA, each $nx1$ dimensional sample vector, x_j can be resolved as

$$x_j = x_{j,diff} + x_{com} + \varepsilon_j \quad j = 1,2, \dots, m \qquad (1.1)$$

where $x_{j,diff}$ represents the differences caused by various factors between the samples in a class, x_{com} which is called common vector and is single and constant for each class represents the common aspects of different samples in that pattern class and ε_j is the error vector. In CVA in order to find $x_{j,diff}$ and the common vector x_{com} for a class, within-class scatter matrix (S_W^c) of all the samples belonging to that class is computed as given in Equation 1.2 and then Eigen decomposition is applied to this matrix. By this decomposition the feature space which is spanned by all the eigenvectors of the within-class scatter matrix is obtained.

$$S_W^c = \frac{1}{m}\sum_{j=1}^{m}(x_j^c - \bar{x}^c)(x_j^c - \bar{x}^c)^T \quad (1.2)$$

where, $\bar{x}^c = \frac{1}{m}\sum_{j=1}^{m} x_j^c$ is the mean vector of class c and m is the number of samples in a class.

The feature space can be decomposed into two complementary subspaces such as the $(n-1)$-dimensional difference and the $(m-n+1)$-dimensional orthogonal indifference subspaces. The difference subspace is spanned by the eigenvectors corresponding to the non-zero eigenvalues of the within-class scatter matrix while the indifference subspace is spanned by the eigenvectors corresponding to the zero eigenvalues. And finally the common vector x_{com} of that class is obtained by projecting of any sample from that class onto the indifference subspace.

$$x_{com}^c = x_i^c V^c, \quad V^c = \begin{Bmatrix} \uparrow & \vdots & \uparrow \\ v_1^c & \vdots & v_k^c \\ \downarrow & \vdots & \downarrow \end{Bmatrix} \quad (1.3)$$

where $k = (m-n+1)$, denote the indices of the eigenvectors which span the indifference feature subspace corresponding to zero eigenvalues of the c-th class within- class scatter matrix.

CVA is first proposed for the pattern classification problems for insufficient data case, in which the number of objects (m) in the training set of each class are equal or less than the size (n) of the object's $(n \leq m)$ (Gulmezoglu et al., 2001).

The above explanation is the application of CVA in insufficient data case. However if the case is sufficient data, where the number of objects (n) is larger than the size of the objects (m), the zero eigenvalues are not emerged from the Eigen analysis of the within-class scatter matrix. This means that the indifference feature subspace has disappeared and CVA cannot be applied. However, Gulmezoglu et al. (2001) showed that, in the case of sufficient data, indifference feature subspace does not disappear indeed and because some eigenvalues of the within-class scatter matrix are too small according to the others, so the class common vector still can be obtained (Gulmezoglu, Dzhafarov, Edizkan, & Barkana, 2007). Thus in case of sufficient data, in order to obtain the common vector of a class the average vector of the class is projected onto the indifference feature subspace that is spanned by the eigenvectors whose value are sufficiently small than the others. The common vector in sufficient data case can be found as,

$$x_{com}^c = \bar{x}^c V^c, \quad V^c = \begin{Bmatrix} \uparrow & \vdots & \uparrow \\ v_1^c & \vdots & v_k^c \\ \downarrow & \vdots & \downarrow \end{Bmatrix} \quad (1.4)$$

where $k = (m-n+1)$, denote the indices of the eigenvectors which span the indifference feature subspace corresponding just near to zero eigenvalues of the c-th class scatter matrix.

In the testing step of CVA the test data is projected onto the indifference subspace of each class and this projected vector is called the remaining vector (Rem_{test}^c) in that class, of the test image . The remaining vector can be found as,

$$Rem_{test}^c = x_{test} V^c \quad (1.5)$$

In order to assign the test data into a class, distances between the classes' common vectors and the remaining vector of the test image in that classes' are computed. Then test image is assigned to the class in which the distance is min.

Experimental studies and discussions

C. Dataset

In the experimental studies two different datasets which are proposed for educational purposes previously in the literature are used. The first dataset consists of the RGB images belong to young carrot seedlings which are taken under natural variable light conditions in the area around Negotino, Republic of Macedonia during the period of February. There are 39 images with the size of 3264 × 2448 in the dataset. The technical information and the sample images of this dataset are given in Table 1 and Figure 5 respectively.

The second dataset that is used in this study comprises field images, vegetation segmentation masks and crop/weed plant type annotations. The dataset which was acquired with the autonomous field robot Bonirob is proposed by (Haug & Ostermann, 2014) and it contains 60 images totally. The technical information and the sample images of this dataset are given in Table 2 and Figure 6 respectively.

Table 1. Technical data of the images in the first dataset.

Image Type	RGB
Camera	10 MP
Flight height	1 m
Number of images	39

Figure 5. RGB image samples of the first dataset.

Table 2. Technical data of the images in the second dataset.

Image Type	RGB/ Near Infrared
Mean distance to ground (d)	450 mm
Ground resolution	~ 8.95 pixels mm^{-1}
Field of view x (at distance d)	~ 145 mm
Field of view y (at distance d)	~ 108 mm
Camera model	JAI AD-130GE [12]
Image resolution	1296 x 966 pixels
Lens	Fujinon TF15-DA-8
Focal length	15 mm
F-number	4
Number of images	60

Figure 6. RGB image samples of the second dataset.

D. Collecting the data

In the proposed novel segmentation method in this paper, each RGB image is represented as a third order tensor as given in Figure 7. Each tensor has 6 frontal slices that corresponds to H and S components of HSV color space, a and b components of Lab color space and u and v components of Luv color space respectively.

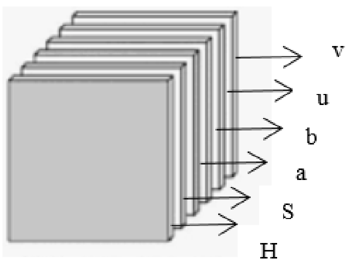

Figure 7. 3rd order color tensor.

The 2-D color matrix which is used as the objects matrix in this study is collected by unfolding the color tensor in mode-3 direction as described in Figure 4. Thus each tube fiber, x_{ij}, (Figure 8) corresponds to a $6x1$ dimensional vector that includes color information belongs to three different color spaces of each image pixel , and its definition is given as,

$$x_{ij} = [H_{ij} \quad S_{ij} \quad a_{ij} \quad b_{ij} \quad u_{ij} \quad v_{ij}]^T \tag{1.6}$$

where $i = 1,..,m,\ j = 1,\dots,n$. Here m,n corresponds to the number of original image rows and columns respectively. According to this definition by unfolding the color tensor in mode-3 direction 2-D color matrix, $X_{(3)}$, which composes the data used in this study can be obtained as,

$$X_{(3)} = \begin{bmatrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ x_{11} & x_{21} & . & . & x_{mn} \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix}_{dxmn} \tag{1.7}$$

Here d equals to the number 6 and, the column vectors of the data matrix $X_{(3)}$ can be resolved as in Equation 1.1 according to CVA.

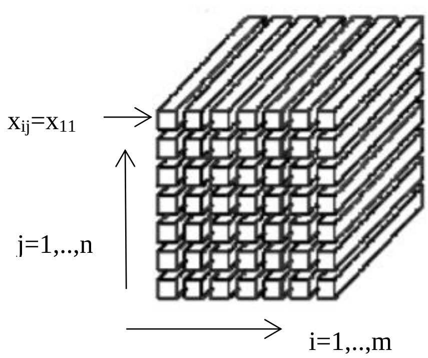

Figure 8. Elements of a tube fiber of a 3D tensor.

D. Feature extraction and segmentation using CVA

CVA uses the statistical information of the data distributions in order to represent the feature space of the data. Therefore to obtain the feature space of the data, the within-class scatter matrix S_W^c of each class is computed as in Equation 1.2. Then Eigen analysis is applied to this within-class scatter matrix and thus feature space which is spanned by all the eigenvectors of the S_W^c is obtained.

As CVA is a learning based classification algorithm, at the beginning the collected dataset is splitted into two as the training set and the test set. While the novel proposed segmentation model is constructing using the training dataset the model performance is evaluated using the test set.

Training Step: In the model construction step, RGB images are first transformed into HSV, Lab and Luv color spaces as the RGB color space is not appropriate for segmentation due to the high correlation among R, G, and B components. The reason of HSV, Lab and Luv color spaces are preferred, are their robustness to illumination and similarity with human color perceptual (Hamuda, Mc Ginley, Glavin, & Jones, 2017). The color components of these three color spaces are gathered together by using 3rd order tensors. Then by unfolding these tensors in mode-3 direction, 2-D matrix representations whose columns corresponds to 6x1 dimensional color vectors (x_{ij}) where the definition is given in Equation 1.6 are obtained. Each of these vectors corresponds to the color components in HSV, Lab and Luv color spaces of one image pixel in the original RGB image. These vectors can be resolved using Equation 1.1 applying CVA as

$$x_{ij}^c = x_{ij,diff}^c + x_{com}^c + \varepsilon_j \quad i = 1,2,..,m \quad , j=1, 2,.., n \tag{1.8}$$

where c denotes the class information namely image color region. $x_{ij,diff}^c$ represents the color vector of each image pixel in difference subspace and x_{com}^c represents the color vector of each image pixel in indifference subspace.

In order to find these subspaces total scatter matrix (S_W^c) of a class of data which corresponds to the color vector in this study, is obtained as in Equation 1.2. Then Eigen analysis is applied to the total scatter matrix of that data class and while the eigenvectors whose eigenvalues are not zero spans the difference subspace the eigenvectors whose eigenvalues are just near to zero spans the indifference subspace. The class common vector for both vegetation and soil can be found as,

$$x_{com}^{c} = \bar{x}^{c} V^{c} \quad (1.9)$$

where $\bar{x}^{c} = \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}^{c}$ is the mean vector of the c-th class and V^{c} is defined as in Equation 1.4.

Testing Step

In the testing step segmentation of the given test image is performed. In order to segment vegetation and soil part in a test image, first off all by unfolding the 3rd color tensor of the test image in mode-3 direction, 2-D color test matrix is obtained (Equation 1.7). Than for the each column vector in 2-D color test matrix by projecting these vectors onto the indifference subspaces of both vegetation and soil the remaining vectors (Rem_{test}^{c}) for both vegetation and soil parts are computed. In order to assign the test data into vegetation or soil, distances between the class common vectors and the remaining vector of the test image in that class are computed. Then test image is assigned to the class in which the distance is min. The flow chart and the pseudo code of the feature extraction process are given in Figure 9 and 10 respectively.

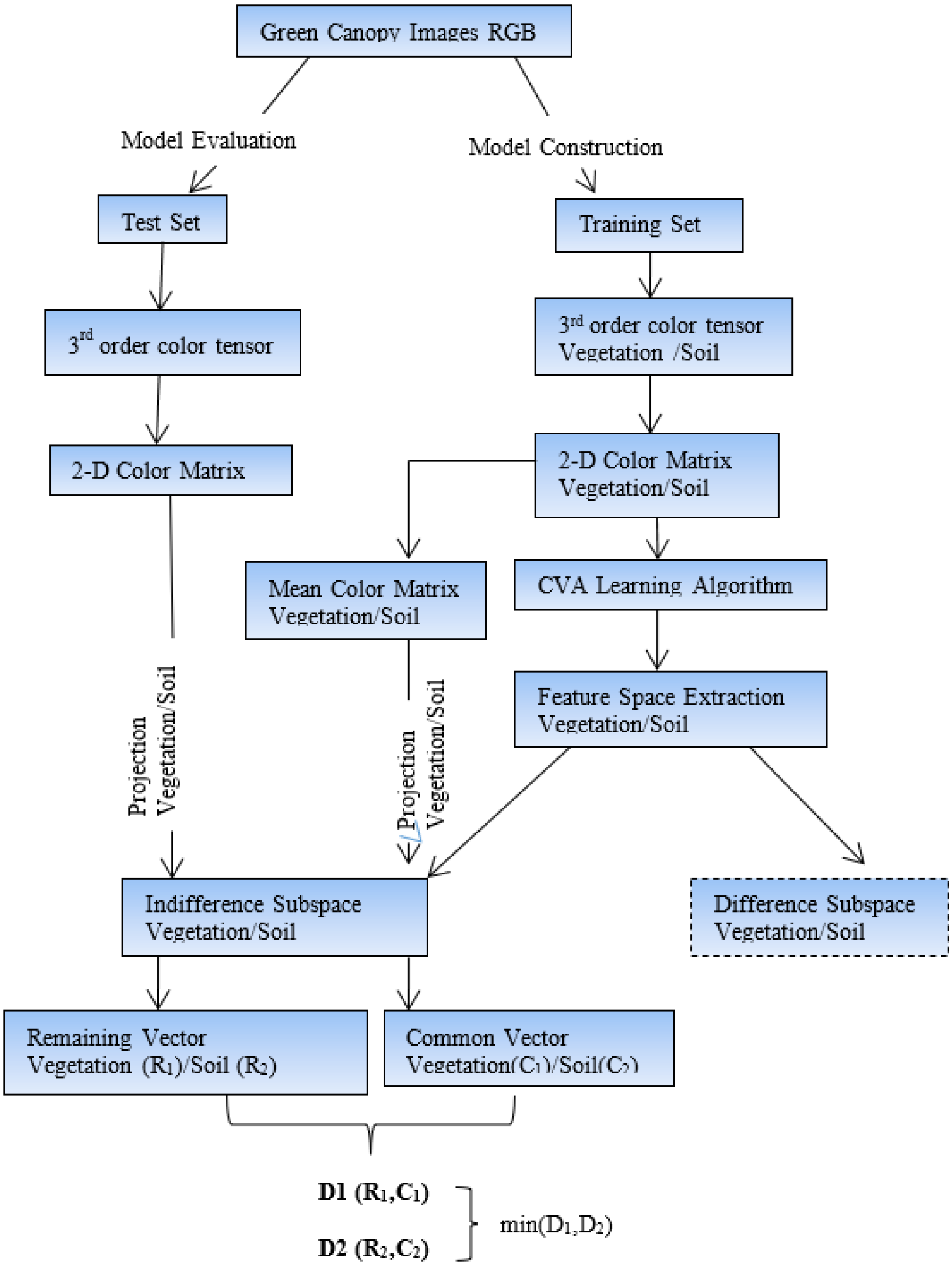

Figure 9. Flow chart of the proposed segmentation method.

Learning based novel sementation algorithm: **Vegetation segmentation method based on common vector approach (CVA) in which images are encoded as 3rd order tensor**.

Data Collection:
-Three color space representation of acquired RGB images are obtained.
-3rd order color tensor representations are obtained
-Unfolding the 3rd order color tensor in mode 3- direction 2D color matrix $X_{(3)}(k),\ k = 1,2,\ldots,mn$ is obtained as given in Equation 1.7.

Feature Extraction/Model Construction:
Taking the columns of 2D color matrices as the objects, for each color region scatter matrices, S_W^c, are obtained and Eigen analysis are performed on this scatter matrices as in Equation 1.2 for each class.

Than the feature space that can be decomposed into two complementary subspaces such as difference and orthogonal indifference subspaces for each color region in an image is defined. And indifference feature subspace is used in order to obtain Common color vector, X_com^c, for a color region.

Model Evaluation:
- 3rd order tensor representation of test image is obtained.
-Then 2D color matrix is obtained.
- Each column of 2D matrix which includes the color information belong to three color spaces of each pixel in an image is projected onto the indifference subspaces of each color region and a remaining vector of each pixel of the test image on each color region is obtained.
- For each color region, distances in that color region between the common color vector and the remaining vector of test image pixel , are computed and the test image pixel is assigned to the color region that the distance is minimum.

Figure 10. Pseudo code for common pixel vector for an image class.

E. Experimental results and evaluations

In this paper using two different datasets, three different experimental studies in which training and test set configurations differ from each other are performed. Thus the performance of the proposed novel segmentation method in this paper is evaluated in case of using different training set and test set combinations. For comparison purposes segmentation is also performed using the most common deep learning model Convolutional Neural Network (CNN) which is used in classification and object segmentation of image related tasks.

All images that are used in both model construction and model evaluation stages are resized standardly as be to 500x500. An image sample with dimension of 500x500 is accepted as a single object in CNN, however as an important superiority of the proposed method over CNN, in the proposed novel segmentation method each image pixel with the dimension of $6x1$ is accepted as an object. These $6x1$ dimensional image pixel vectors are obtained with the unfolding process of the 3rd order tensor representation of images in mode-3 direction as defined above. Thus while it is obtained only one object with the dimension of 500x500 from only one image sample with the dimension 500x500 in CNN segmentation method, it is obtained 250000 objects with the dimension 6x1 from only one image sample with the dimension 500x500 in the novel proposed segmentation method.

In the first experimental study, only one dataset is used as the training and test set and on this dataset 10-fold cross validation is applied and the performance of the proposed method is evaluated. In the second experimental study, the performance of the proposed method is evaluated, when in case of one of the dataset is the training set and the other one is the test set. Finally in the third experimental study, the two datasets are combined. By applying 10- fold cross validation on combined dataset performance of the proposed method is evaluated. Some example images from the two datasets and their corresponding ground-truth images are illustrated in Figure 11.

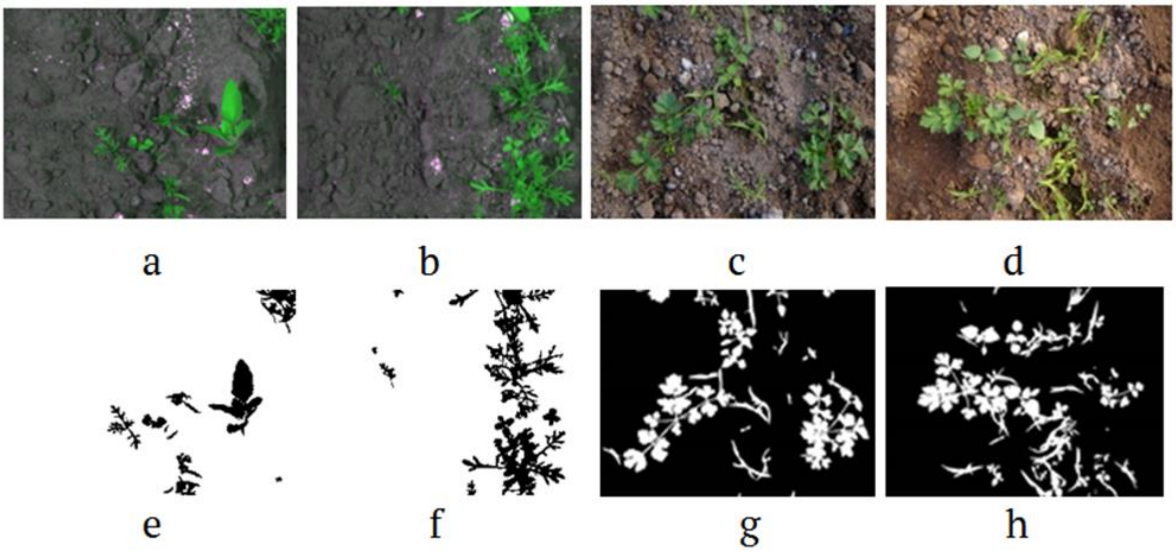

Figure 11. Some example images (upper row) and corresponding ground-truth images (below row). a-e and b-f pairs are from first dataset; c-g and d-h pairs are from second dataset.

In this study two different metrics are used in order to evaluate the results and performance of the novel proposed segmentation model. Both of these two metrics that are accuracy and mean intersection over union (IoU) can be obtained from the confusion matrix of the results (Chen et al., 2017; Rahnemoonfar & Sheppard, 2017). While accuracy is the percentage of correctly classified pixels for each class and it can be computed as the percentage of the summation of the True Positive (TP) and the True Negative (TN) numbers in the confusion matrix of the tested image. Mean IoU which is a well-known metric in semantic segmentation can be defined as the number of pixels that are common between the target and prediction masks divided by the total number of pixels present across both masks (Shen & Zeng, 2019). For this metric, an IoU of greater than 0.5 is considered as an acceptable semantic segmentation.

For a visual evaluation about the proposed pixel-wise segmentation method, two sample segmentation results for two sample images are displayed in Figure 12.

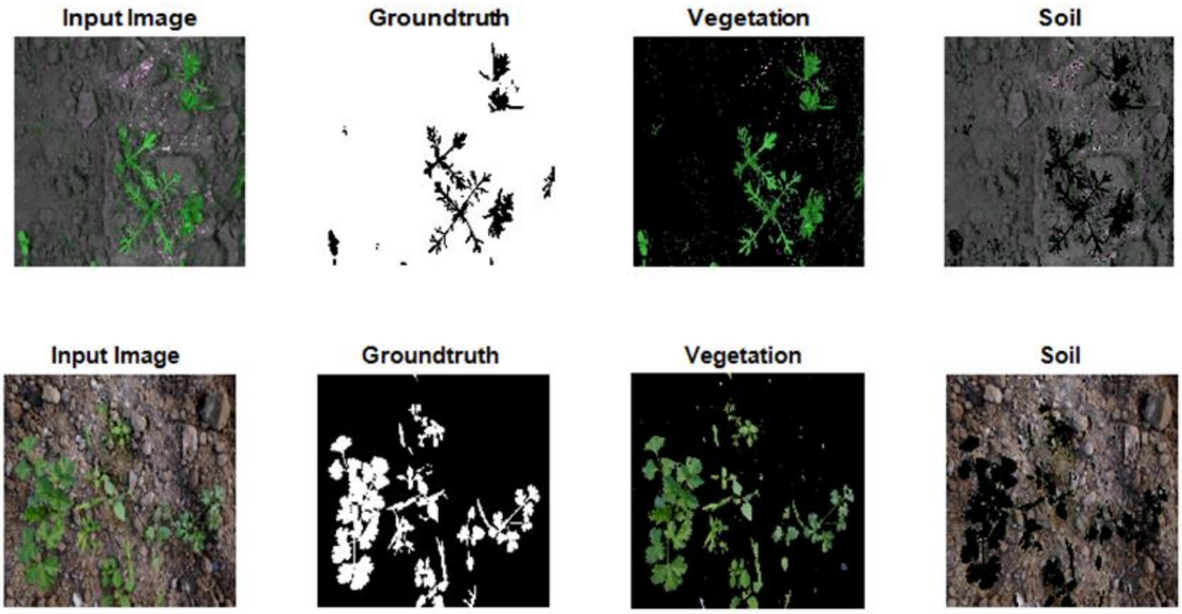

Figure 12. Proposed methods' segmentation results for two different experiments for two sample images

In all of the three experimental studies novel proposed segmentation algorithm achieved extremely good segmentation results. The segmentation result of these three experimental studies in terms of accuracy and mean IoU are presented in Table 3 Table 4 for the proposed method and for CNN respectively. In all of the three experimental studies the segmentation results achieved for the novel proposed algorithm outperforms

CNN-based semantic segmentation method. In the case that studying only on one dataset both in model construction and in model evaluation steps the performance of our proposed method and of CNN are so close. However in the second and in the third experimental studies this is not the case. In the second study in which one dataset is using in model construction step while the other one is using in model evaluation step, between the segmentation performances there is approximately 13.9 % difference that is remarkably high in favor of our proposed method. We think that, this is the result of difference that while CNN accepts the whole image as an object our proposed algorithm accepts an image pixel color vector as an object. This shows that our proposed algorithm is more robust to changes in the datasets as it extracts similar features for specified color regions in different datasets. In the third experimental study the two datasets are combined and, model construction and model evaluation is performed using the joint dataset. A similar result is emerged as it is in the second experiment. In the third experimental study our proposed algorithm outperforms the CNN with about 7.4% high segmentation performance.

Table 3. Results of experimental studies for the novel proposed segmentation method.

Experimental study	mean IoU	Class IoU		Accuracy (%)	Class Accuracy (%)	
		Soil	Vegetation		Soil	Vegetation
1	0.86	0.99	0.78	99.5	99.52	99.56
2	0.58	0.98	0.17	97.9	98.2	89.5
3	0.86	0.99	0.73	99.4	99.45	99.2

Table 4. Results of experimental studies for CNN based semantic segmentation method.

Experimental study	mean IoU	Class IoU		Accuracy (%)	Class Accuracy (%)	
		Soil	Vegetation		Soil	Vegetation
1	0.95	0.99	0.90	99	99	99
2	0.44	0.84	0.04	84	8	90
3	0.80	0.91	0.69	92	91	96

Conclusion

In this study, a novel pixel-wise vegetation segmentation method based on CVA in sufficient data case is proposed. In this method RGB images of green vegetation canopy are encoded as 3rd order tensors and each dimension corresponds to the color component of HSV, Lab and Luv color spaces. CVA obtains a common color pixel vector for each color region in an image which represents the common properties of that color region and it is used for segmentation. In order to compare the performance of the proposed method, Convolutional Neural Networks based semantic segmentation method is implemented for vegetation segmentation of green canopy images.

Using CVA a common color pixel vector for each color region in an image which represents the common properties of that color region is obtained. In CVA the obtained common color vector of a color region represents the common properties of that region and it is used for segmentation in order to determine if a pixel of an input image is belong to a color region or not. One of the main difference between the novel proposed CVA based segmentation method and CNN-based semantic segmentation method is CNN requires a ground truth image but the proposed method does not require. The proposed method is a supervised learning algorithm that makes classification with feature extraction. That is why without need of a ground truth image it can make segmentation. The other important difference between the proposed method and the CNN based method in a segmentation process is while CNN based semantic segmentation requires a huge amount of training data and the corresponding ground truth images, the novel CVA based segmentation method can perform good performances even if there is only one training image sample in the dataset.

The performance of the proposed method has been compared with the performance of CNN based semantic segmentation method and in all of the three experimental studies proposed method has achieved extremely high performance according to CNN. When it is used only one dataset both in training and the test set the performance difference between the proposed method and CNN based semantic segmentation is low but especially, when it is used different dataset combinations as the training set and the test set, the proposed method achieves extremely high segmentation performances according to CNN based semantic segmentation method.

References

Bai, X., Cao, Z., Wang, Y., Yu, Z., Hu, Z., Zhang, X., & Li, C. (2014). Vegetation segmentation robust to illumination variations based on clustering and morphology modelling. *Biosystem Engineering, 125*, 80-97. DOI: http://doi.org/10.1016/J.BIOSYSTEMSENG.2014.06.015

Bao, X., Javanbakhti, S., Zinger, S., Wijnhoven, R., & De With, P. H. N. (2013). Context modeling combined with motion analysis for moving ship detection in port surveillance. *Journal of Electronic Imaging, 22*(4), 041114. DOI: http://doi.org/10.1117/1.JEI.22.4.041114

Chen, S. W., Shivakumar, S. S., Dcunha, S., Das, J., Okon, E., Qu, C., ... Kumar, V. (2017). Counting apples and oranges with deep learning: a data-driven approach. *IEEE Robotics and Automation Letters, 2*(2), 781-788. DOI: http://doi.org/10.1109/LRA.2017.2651944

Gulmezoglu, M. B., Dzhafarov, V., & Barkana, A. (2001). The common vector approach and its relation to principal component analysis. *IEEE Trasnsactions on Speech and Audio Processing, 9*(6), 655-662. DOI: http://doi.org/10.1109/89.943343

Gulmezoglu, M. B., Dzhafarov, V., Edizkan, R., & Barkana, A. (2007). The common vector approach and its comparison with other subspace methods in case of sufficient data. *Computer Speech and Language, 21*(2), 266-281. DOI: http://doi.org/10.1016/j.csl.2006.06.002

Guo, W., Rage, U. K., & Ninomiya, S. (2013). Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model. *Computers and Electronics in Agriculture, 96*, 58–66. DOI: http://doi.org/10.1016/j.compag.2013.04.010

Hamuda, E., Glavin, M., & Jones, E. (2016). A survey of image processing techniques for plant extraction and segmentation in the field. *Computers and Electronics in Agriculture, 125*, 184–199. DOI: http://doi.org/10.1016/j.compag.2016.04.024

Hamuda, E., Mc Ginley, B., Glavin, M., & Jones, E. (2017). Automatic crop detection under field conditions using the HSV colour space and morphological operations. *Computers and Electronics in Agriculture,133*(5), 97–107. DOI: http://doi.org/10.1016/S0031-3203(00)00149-7

Hassanein, M., Lari, Z., & El-Sheimy, N. (2018). A new vegetation segmentation approach for cropped fields based on threshold detection from hue histograms. *Sensors, 18*(4), 1253. DOI: http://doi.org/10.3390/s18041253

Haug S., & Ostermann J. (2015). A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. *Lecture Notes in Computer Science*, vol 8928. Springer, Cham. DOI:https://doi.org/10.1007/978-3-319-16220-1_8

Kaur, D., & Kaur, Y. (2014). Various image segmentation techniques: a review. *International Journal of Computer Science Mobile Computing, 3*(5), 809–814.

Kolda, T. G. & Bader, B. W. (2009). Tensor decompositions and applications. *Society for Industrial and Applied Mathematics - SIAM Review, 51*(3), 455-500. DOI: http://doi.org/10.1137/07070111X

Kumar, A. D., & Prema, P. (2016). A novel wrapping curvelet transformation based angular texture pattern (WCTATP) extraction method for weed identification. *ICTACT Journal on Image and Video Processing, 6*(3), 1192-1206. DOI: http://doi.org/10.21917/ijivp.2016.0174

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps. *Biosystems Engineering, 114*(4), 358–371. DOI: http://doi.org/10.1016/j.biosystemseng.2012.08.009

Nex, F. C, & Remondino, F. (2014). UAV for 3D mapping applications: a review. *Applied Geomatics, 6*(1), 1–15. DOI: http://doi.org/10.1007/s12518-013-0120-x

Osunmadewa, B. A., Wessollek, C., & Karrasch, P. (2015). Linear and segmented linear trend detection for vegetation cover using GIMMS normalized difference vegetation index data in semiarid regions of Nigeria. *Journal of Applied Remote Sensing, 9*(1), 096029. DOI: http://doi.org/10.1117/1.JRS.9.096029

Parra, A., Boutin, M., Delp, E. J. (2017). Automatic gang graffiti recognition and interpretation. *Journal of Electronic Imaging, 26*(5), 051409. DOI: http://doi.org/10.1117/1.JEI.26.5.051409

Potena, C., Nardi, D., Pretto, A. (2017). Fast and accurate crop and weed identification with summarized train sets for precision agriculture. In A. Rodic & T. Borangiu (Eds.), *Advances in Robot Design and Intelligent Control* (p. 105–121). Cham, SW: Springer.

Prema, P., & Murugan, D. (2016). A novel angular texture pattern (ATP) extraction method for crop and weed discrimination using curvelet transformation. *ELCVIA Electronic Letters on Computer Vision and Image Analysis, 15*(1), 27–59. DOI: http://doi.org/10.5565/rev/elcvia.826

Rahnemoonfar, M., & Sheppard, C. (2017). Deep count: fruit counting based on deep simulated learning. *Sensors, 17*(4), 905. DOI: http://doi.org/10.3390/s17040905

Riegler-Nurscher, P., Prankl, J., Bauer, T., Strauss, P., & Prankl, H. (2018). A machine learning approach for pixel wise classification of residue and vegetation cover under field conditions. *Biosystem Engineering 169*, 188–198. DOI: http://doi.org/10.1016/j.biosystemseng.2018.02.011

Ruiz-Ruiz, G., Gómez-Gil, J., & Navas-Gracia, L. M. (2009). Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *Computers and Electronics in Agriculture, 68*(1), 88–96. DOI: http://doi.org/10.1016/j.compag.2009.04.009

Sattari, M., Toroslu, I. H., Karagoz, P., Symeonidis, P., & Manolopoulos, Y. (2015). Extended feature combination model for recommendations in location-based mobile services. *Knowledge and Information Systems, 44*(3), 629-661. DOI: http://doi.org/10.1007/s10115-014-0776-5

Shen, F., & Zeng, G. (2019). Semantic image segmentation via guidance of image classification. *Neurocomputing, 330*, 259-266. DOI: http://doi.org/10.1016/j.neucom.2018.11.027

Slaughter, D. C., Giles, D. K., & Downey, D. (2008). Autonomous robotic weed control systems: a review. *Computers and Electronics in Agriculture, 61*(1), 63–78. DOI: http://doi.org/10.1016/j.compag.2007.05.008

Wang, A., Zhang, W., & Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. *Computers and Electronics in Agriculture, 158*, 226-240. DOI: http://doi.org/10.1016/j.compag.2019.02.005

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: a review. *Precision Agriculture, 13*(6), 693–712. DOI: http://doi.org/10.1007/s11119-012-9274-5

Zheng, L., Shi, D., & Zhang, J. (2010). Segmentation of green vegetation of crop canopy images based on mean shift and Fisher linear discriminant. *Pattern Recognition Letters, 31*(9), 920–925. DOI: http://doi.org/10.1016/j.patrec.2010.01.016