Received August 15, 2019, accepted September 10, 2019, date of publication September 13, 2019, date of current version September 26, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2941441*

# Vegetation Cover Estimation Using Convolutional Neural Networks

**MOHAMMED ASAAD GHAZAL[1], (Member, IEEE), ALI MAHMOUD[2], (Member, IEEE), ALI ASLANTAS[2], (Member, IEEE), AHMED SOLIMAN[2], (Member, IEEE), AHMED SHALABY[2], (Member, IEEE), JÓN ATLI BENEDIKTSSON[3], (Fellow, IEEE), AND AYMAN EL-BAZ[2], (Senior Member, IEEE)**

[1]Electrical and Computer Engineering Department, Abu Dhabi University, Abu Dhabi, United Arab Emirates
[2]Bioengineering Department, University of Louisville, Louisville, KY 40292, USA
[3]Electrical and Computer Engineering Department, University of Iceland, 101 Reykjavik, Iceland

Corresponding author: Mohammed Asaad Ghazal (mohammed.ghazal@adu.ac.ae)

**ABSTRACT** Vegetation is an important parameter in all bio- and ecosystems, and it should be monitored to conserve and restore the environment. This paper presents a design and implementation of a compact system for vegetation cover monitoring, which consists of a small unmanned aerial vehicle (UAV) equipped with four different cameras (RGB, NIR, NDVI, and red). These cameras simultaneously record videos and wirelessly send them to a ground station that applies on them a set of algorithms to construct a mosaic representing the covered area and segment vegetative regions. Mosaicing is performed using a fast multi-threaded approach based on binary descriptors. For segmenting vegetation, two scenarios are tested by using convolutional neural networks (CNNs). The first scenario trained a CNN using the images obtained from the four cameras, and then used it with the constructed mosaics. In the second scenario, we trained individual CNNs using images from each of the four cameras individually. The performance of the former scenario exceeded the others from the perspective of accuracy. Moreover, it proved to be highly comparable to previous approaches utilizing level sets, while the processing time is reduced. The proposed approach obtained high accuracy in terms of the Dice similarity coefficient (97 %), which demonstrates its favorable performance.

**INDEX TERMS** CNN, segmentation, UAV, vegetation cover estimation.

## I. INTRODUCTION

Classifying and mapping vegetation is significantly important for studying the environmental fluctuations worldwide including the climate, as well as the way it affects various ecosystems, hydrologic processes, and geochemical cycles [1]. Vegetation regions, such as forests, are responsible for protecting biodiversity, sustaining land productivity, regulating water resources, and keeping up economic activities [2] mainly by providing natural habitats for plants and animals, which are significant for nourishment [3]. They also help to preserve energy stability in the world [4] in addition to wood supplies [5]. As the economy developed rapidly with the human population growth, urban areas increased which dramatically affected the vegetative resources [6] with more than a third of the natural territory has been destroyed by humans [7]. This triggered countries to significantly examine the progress and construction actions in order to effectively protect water and ground resources [8], which change rapidly in vegetation. Such rapid alteration makes developing new methods monitoring the changes and estimation of vegetation of high importance [3], [9].

Traditionally, acquiring information on green area cover was obtained by field sampling method which was time-consuming and of high cost. Therefore, this information was usually updated infrequently, e.g. in 10-year intervals [3], especially for the purpose of city planning to grasp land cover and vegetation in urban areas. Recently, remote-sensing techniques [10] have shown to be powerful in understanding land and vegetation covering in the territories, in addition to monitoring climate changes. Significant research efforts have been devoted to develop techniques for mapping and estimating forest cover and carbon emissions using remote sensing technology over large scales. These techniques offer

The associate editor coordinating the review of this manuscript and approving it for publication was Dusmanta Kumar Kumar Mohanta.

  

an easy [11] and cost-effective solution for monitoring vegetation areas [2], [12]. The practicability of remote sensing data has been shown in published articles by taking advantage of either optical or microwave data [13]. Remote sensing can be classified into two parts, airborne and spaceborne. Spaceborne remote sensing collects satellite images e.g., the remote sensing image classification framework developed by Romero *et al.* [14] using a deep extraction technique. The pre-training they used was unsupervised greedy layerwise linked with an effective sparse feature learning algorithm. On the other hand, airborne remote sensing collects images from airplanes and Unmanned Aerial Vehicles (UAVs). There are some restraints on satellite remote sensing. Satellite based data cannot often provide the needed measurements. In addition, it is hard to get optical images of a particular area on demand because of unpredictable cloud coverage. These limitations could be handled by using airborne remote sensing platforms instead, which will be the focus of this paper.

UAV-based remote sensing made use of novel data sources to efficiently identify vegetation metrics spatially at a significantly high resolution with adjustable revisit frequencies [15]. Having the ability to operate at low altitudes, UAVs allow data collection at a high resolution in a convenient and controllable manner [16], [17]. Recently, several UAV-based studies have been employed for vegetation estimation and monitoring [4], [16], [18]–[20]. Berni and colleagues [16] mounted affordable thermal and multispectral, narrowband image sensors on a UAV. These sensors that ranged from surface reflectance to temperature imagery were used for the construction of quantitative remote sensing products to help with evaluation and vegetation monitoring. Also, a fixed wing UAV platform was utilized in southern New Mexico for the assessment of the rangelands, with a 5–6 cm/pixel resolution [18]. Another study was conducted later, over the same area [19], by mounting a multispectral camera with six bands on a UAV to capture high-resolution data in the near infrared (NIR) spectrum. Furthermore, a study was conducted in Cordoba, Spain [20] using a color infrared camera mounted on the UAV for assessing and quantifying orchard tree heights. This study showed that the use of cheaper products and systems can produce similar results to expensive sensors such as LIDAR systems. Cunliffe *et al.* [21] presented an inexpensive approach that utilized structure-from-motion (sfm) to characterize the biotic structure from aerial images captured by a light-weight hexacopter. Ishida *et al.* [22] presented a hyperspectral imaging UAV-based classification approach for vegetation. In their work, they used a liquid crystal tunable filter to avoid using a large number of optical filters for capturing the data,and their classification was based on support vector machine (SVM). Wan *et al.* [23] used a UAV to acquire aerial multispectral images for oilseed rape, and they used K-means approach to estimate the flower number. Lu *et al.* [24] used aerial images captured by an octocopter to estimate the the chlorophyll content and the leaf area index (LAI) for grassland.

Ghazal *et al.* [4] presented a UAV based vegetation estimation system that was equipped with a Normalized Difference Vegetation Index (NDVI). In their system, the captured NDVI images were stitched using a scale invariant feature transform (SIFT) algorithm-based mosaicing technique [25], and the vegetation regions were segmented using level sets. Since UAVs are commonly constrained by the strength of their controlling unit signals, such platforms cannot cover a big area compared to satellites. Therefore, the data obtained by UAVs usually cover small areas, which calls for a recording of the video of the area instead of taking photographs. For this purpose, the video is processed using video mosaicing techniques, which are convenient for real-time surveillance, disaster management, and military inspection. All of these applications require a precise and fast formation of the mosaic, especially when it comes to real-time airborne remote sensing such as the evaluation of different vegetation indices [26]. Thus, developing an accurate mosaic of the recorded area is the preliminary step in vegetation estimation. The majority of the proposed techniques use photogrammetric methods and various complex processes, which is time-consuming, as well as semi-automatic due to the necessity for human intervention [27]. The obtained mosaic is usually subjected to various segmentation methods for identifying areas of interest. Most popular segmentation methods involve active contours, which are defined as a set of points that aim to enclose a target feature that is to be extracted. These are used in many applications that require object segmentation [28]–[30] but are rarely used with segmentation of photosynthetically active objects.

The motivation of this paper is to employ machine learning algorithms to study high-resolution ground vegetation cover videos captured by a small-platform UAV equipped with four different special cameras (NIR, NDVI, Red, and RGB). The UAV was wirelessly connected to a ground station for processing. The captured videos were then stitched individually to develop accurate mosaics of the recorded area, which is the preliminary step in vegetation estimation. Finally, the obtained mosaics were fed into a segmentation framework to find vegetation areas. Throughout this work, the vegetation segmentation was performed using a convolutional neural network (CNN) trained with the images obtained from the four cameras. We investigated two scenarios for training the CNN: using the images obtained from the four cameras to train a single CNN and using the images obtained from each camera to train an individual CNN. Moreover, the mosaicing was done using a fast multi-threaded approach based on binary descriptors, that we modified its implementation to give mosaics for the four cameras. We evaluated the performance of the segmentation technique for detecting and outlining the photosynthetically-active areas from the perspective of accuracy as well as the computational time of the video mosaicing. The results obtained for the proposed system were compared to [4]. The proposed system can be beneficial in the analysis and tracking of important vegetative metrics, as well

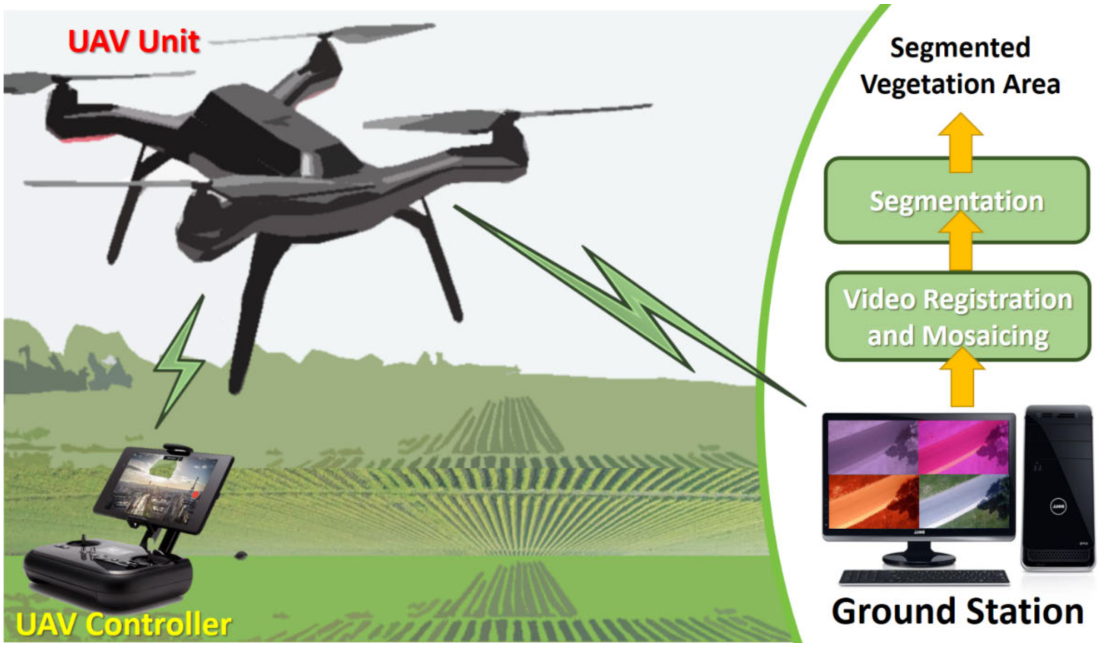


**FIGURE 1.** Overall system design.

**FIGURE 2.** Study sites in the city of Louisville in the state of Kentucky in the United States.

as changes in ground cover, in areas of precision agriculture, soil erosion control, land degradation and desertification and vegetation health assessment. The remainder of this paper is organized as follows: In Section II, we present the materials and methods; in Section III, we present and discuss our results; and in Section IV, we conclude the paper.

## II. MATERIALS AND METHODS

In this section, we discuss our study site, as well as the data acquisition setup, which includes a UAV platform equipped with cameras to record videos which are sent to a ground station over a wireless connection for processing. The processing techniques which are mainly for mosaicing and segmentation are also discussed in this section. The overall system including the general processing steps is shown in Fig. 1.

### A. DATA ACQUISITION

In this study, our fieldwork was done in the city of Louisville in the state of Kentucky in the United States. Three different study sites were used. The exact location of each study area is presented in Fig. 2. The positions of the areas where the acquisition of the data was performed is illustrated by red. The areas contain a combination of vegetation and dry land, water or pavement. For reference, the mean annual temperature of Kentucky varies from 56 °F in the extreme north to 59 °F in southwestern counties. July is usually the warmest month and January is the coldest. Extreme summer temperatures usually reach 100 °F at most locations in the state, but only infrequently. Louisville is losing trees at a rate of about 54,000 a year, according to a recent, comprehensive assessment of trees in Jefferson County [31].

Throughout this work, the data was acquired using a UAV with four different cameras mounted on (NIR, NDVI, Red and RGB). The UAV platform used is a Solo 3DR, which packs over 35 sensors and 20 microprocessors into a rigid, lightweight unibody shell. It has a maximum flight time of 16 minutes with the four cameras being attached with a maximum altitude of 400 ft (122 m). Moreover, it weighs 4.4 pounds and is capable of communication over a range

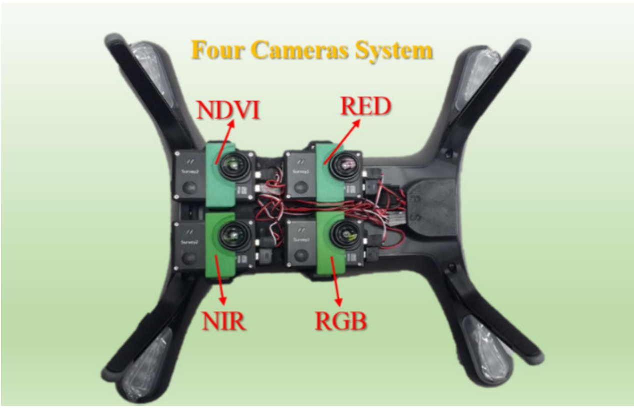


**FIGURE 3.** Camera arrangement on the UAV.

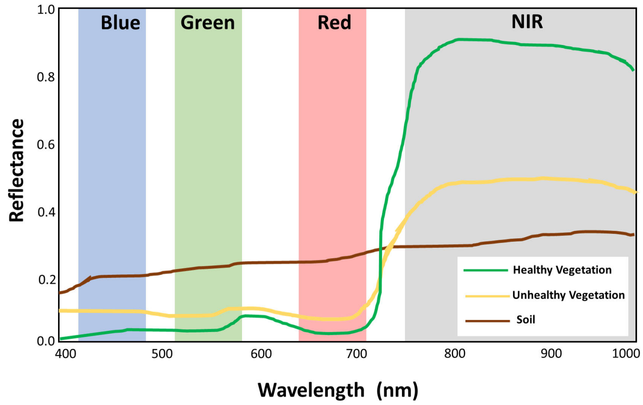


**FIGURE 4.** Vegetation examination using different spectral ranges.

of 0.5 mile (0.8 km). In addition, the cameras were mounted on the drone using a stable mount, which eliminates vibrations occurring during the flight. The cameras were arranged as shown in Fig. 3. The four cameras used were of brand Survey2 Mapir, covering different spectral ranges: visible RGB, 850 nm NIR, 650 nm Red, and Red + NIR (NDVI).

Healthy vegetative regions mainly absorb a large portion of the incident visible light and reflect most of the NIR light; as opposed to the sparse unhealthy vegetation, where the reflection of the visible light is higher and that of the NIR is lower. With respect to bare soils, the reflection occurring is moderate in the infrared and red bands [32]. The visible RGB camera receives visible light RGB color. NDVI cameras have been used to determine whether the object being imaged has green live vegetation content or not. This is done by dividing the subtraction of the values of the reflectance in the red band from those in the NIR band by the addition of the reflectance values in both bands (NDVI = (NIR − Red)/(NIR + Red)). However, the resulting contrast will not be so correct as mere Red and NIR cameras. Fig. 4 illustrates the wavelengths of the electromagnetic spectrum captured at each spectral range to examine the vegetation [33].

The lens of each camera has a horizontal field of view (HFOV) of 82° and $f/2.8$ aperture. In addition, the sensor size is 4675 pixel × 3500 pixel, where the pixel size is 1.34 $\mu$m. The maximum shutter speed is 0.5 ms. The cameras are directed towards the ground. At a given UAV altitude $A$, the horizontal ($H$) and vertical ($W$) dimensions covered by the camera can be calculated in meters (or yards) to be used in area mapping using the field of view (FOV) of the cameras as:

$$H = 2A \tan(\frac{VFOV}{2}), \quad (1)$$

$$W = 2A \tan(\frac{HFOV}{2}), \quad (2)$$

where $HFOV$ and $VFOV$ are the horizontal and vertical field of views, respectively.

During the data acquisition, the flight paths were chosen in a way to cover the area efficiently and enable enough overlap between the frames for an accurate mosaicing process. Example of selected flight paths for the chosen study areas are marked by red in Fig. 5. The UAV is wirelessly connected to a ground station equipped with a PC with 2.6 GHz Core i7 Quad processor, and 64 GB RAM. Algorithms were implemented using MATLAB (The MathWorks, Natick, Massachusetts) and C++.

### B. VEGETATION SEGMENTATION

The vegetation segmentation is performed using a deep 3D-CNN, which is a premium type of deep learning that is doing way better than other approaches in neural network [34]. It is worth mentioning that the inspiration for this particular deep neural architecture is neuroscience. The 3D-CNN is based on the sensitivity of the cortex cells in the vision system of human to small regions. Such cells strongly respond inside the receptive fields in a manner that makes them capable of making full use of the spatial local correlation of images [35]. The performance of remote sensing image scene classification has been progressed since the powerful features acting learned through CNNs [36]. We used a 3D-CNN (Fig. 6) that is fully connected to a conditional random field (CRF) model [37] in order to utilize all the captured images for each scene, (namely: RGB, NDVI, NIR, and red channels), to segment the vegetation area within that scene.

Each image obtained from one of the four cameras is considered as an input channel to the CNN. To obtain output a final labeled region map of the segmented input object (vegetation regions), the deep 3D-CNN generates soft segmentation maps followed by a post-processing step using fully-connected classifier based on a 3D CRF to effectively remove false positives. Sequential convolution is carried out for the input with multiple filters at the network cascaded layers, where each consists of multiple channels, with each channel being corresponding to the three layers of a single captured feature (output image from specific camera). At the first layer, the input 3D feature volumes represent the input channels of the input. This could be considered as performing convolution of 4D kernels with 4D volumes (concatenated feature 3D volumes).

FIGURE 5. Flight paths for the study sites.

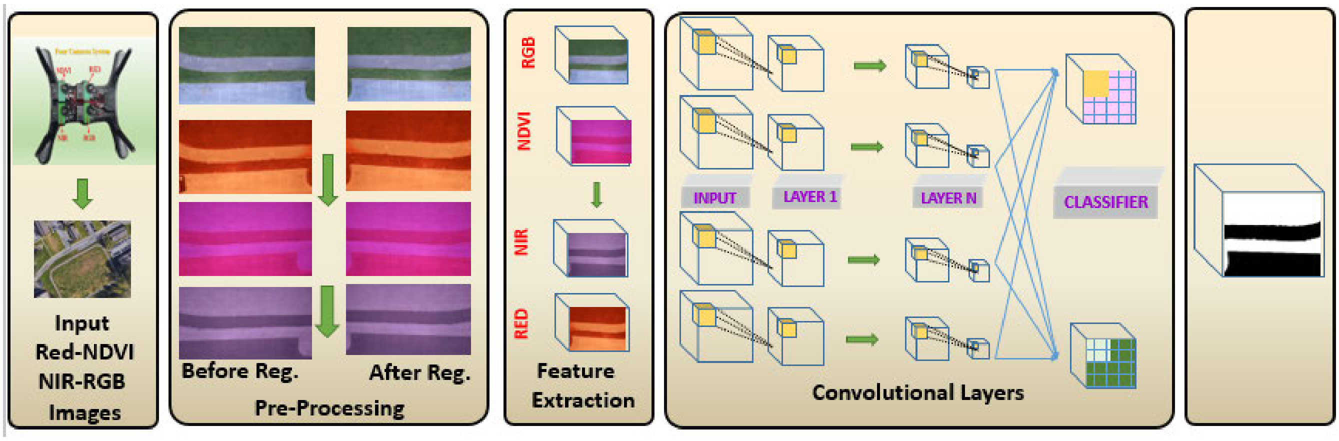


FIGURE 6. Convolutional neural network based framework.

The deep 3-D CNN architecture is composed of seven layers to provide a more power to discriminate between the different classes. The proposed network takes the adjacent image patches and in a dense way train them into one pass of the network using an effective dense training scheme that takes into consideration the adaptation to inherent the class imbalance issues that exist in the data automatically. The network utilizes parallel convolutional pathways to process the input image at different scales simultaneously. This structure incorporates a wider range of both local and contextual information to enhance and enrich the segmentation process. A post-processing step, for the generated soft segmentation maps by the network, were done using a 3-D fully connected CRF to remove any false positives. In the baseline network, the used layers for the feature extraction process consists of $5^3$ kernels each. The size of the neighbourhood system for the input voxel that influence the neuron activation (receptive field) was $17^3$. The classification layer has a size $1^3$ kernel. Adding extra layers to the baseline scheme can add more depth that has the advantage of more discriminatory power. However, this can result in more suffering from the expensive computation and the need for additional trainable parameters. We solved this issue by using a set of smaller kernels to reduce the number of element-wise multiplication (convolutional operations) in addition to the training parameters by factor of 5. In this architecture, a two successive layers with $3^3$ kernels replaced the $5^3$ kernels layers in the baseline without changing the receptive field. This change has reduced the computation cost in addition to the number of weight parameters. This architecture is advantageous for its capability of fusing and capturing contextual 3D information from the input feature images. The configuration parameters were heuristically selected.

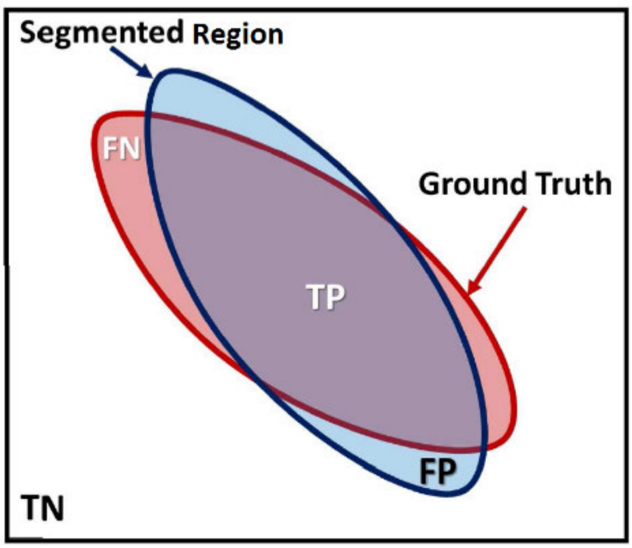


FIGURE 7. 2D illustration of DSC measurement for segmentation evaluation.

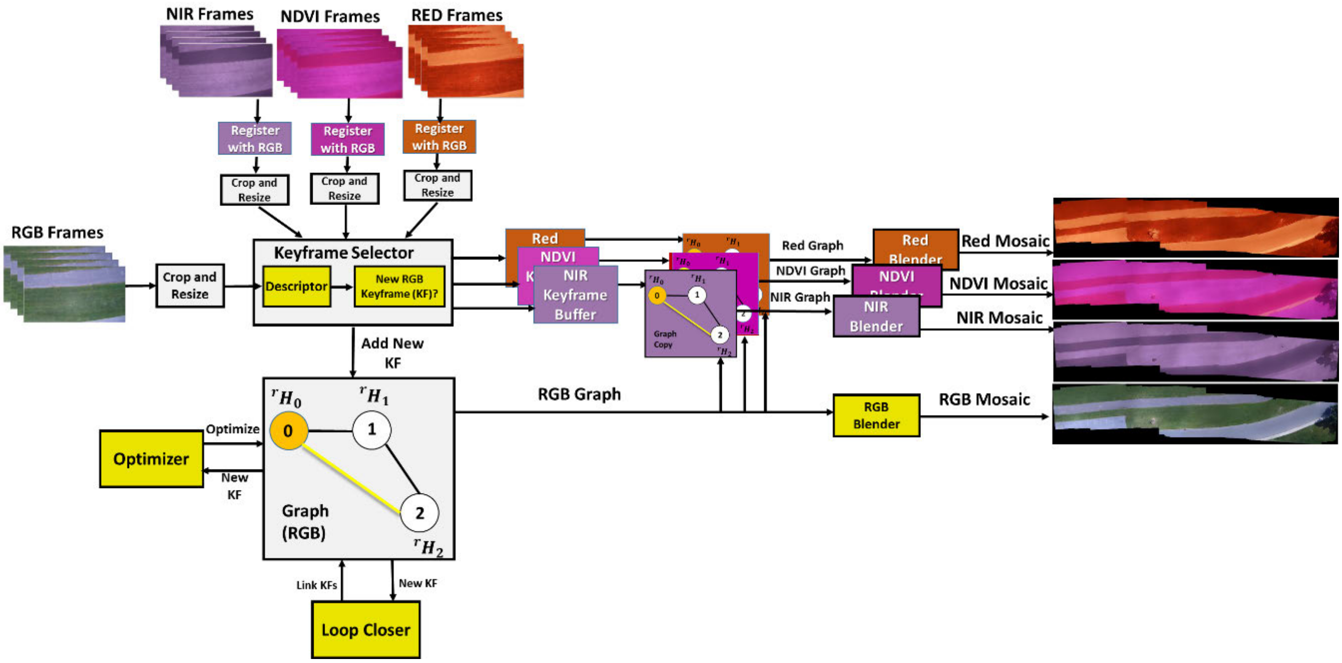


**FIGURE 8.** Mosaicing framework: most of the processing is done to establish the RGB graph which is then copied to be used with the images from the other three cameras. The nodes and the black links in the RGB graph are established by the keyframe selector while the yellow links are established by the loop closer.

#### 1) EVALUATION METRIC

The performance of the segmentation was quantitatively evaluated using the Dice similarity coefficient (DSC), which characterizes spatial overlaps between the segmented and ground truth regions (Fig. 7).

Let **G** and **S** denote sets of ground truth and segmented pixels, respectively. The similarity surface evaluates an overlap between these sets and accounts for cardinalities (i.e., pixel numbers) $c_i = |V_i|$ of true positive (tp), false negative (fn) subsets $V_i$; $i \in \{tp, fp, fn\}$. The subsets contain true positive pixels for each class, background pixels labeled as class, and true class pixels labeled as background, respectively:

$$\begin{aligned} V_{tp} &= \{v : v \in \mathbf{G}, v \in \mathbf{S}\}; \quad c_{tp} = |V_{tp}| \\ V_{fp} &= \{v : v \notin \mathbf{G}, v \in \mathbf{S}\}; \quad c_{fp} = |V_{fp}|, \\ V_{fn} &= \{v : v \in \mathbf{G}, v \notin \mathbf{S}\}; \quad c_{fn} = |V_{fn}| \end{aligned} \tag{3}$$

where $\mathbf{G} = V_{tp} \cup V_{fn}$; $\mathbf{S} = V_{tp} \cup V_{fp}$; $V_{tp} = \mathbf{G} \cap \mathbf{S}$; and $V_{tp} \cup V_{fp} \cup V_{fn} = \mathbf{G} \cup \mathbf{S}$. Moreover, $\cup$ and $\cap$ denote the set union and intersection, respectively. Therefore, it holds that $|\mathbf{G}| = c_{tp} + c_{fn}$; $|\mathbf{S}| = c_{tp} + c_{fp}$, and $|\mathbf{G} \cup \mathbf{S}| = c_{tp} + c_{fp} + c_{fn}$. The DSC [38] is defined as:

$$DSC = 100\frac{2c_{tp}}{2c_{tp} + c_{fp} + c_{fn}} \equiv 100\frac{2|\mathbf{G} \cap \mathbf{S}|}{|\mathbf{G}| + |\mathbf{S}|}. \tag{4}$$

### C. MULTI-THREADED VIDEO MOSAICING USING BINARY DESCRIPTORS

An image mosaic can be defined as a composition of a sequence of images that have matching features from which geometric relationships can be established [39]. Image mosaics are used in cases when a single frame of camera cannot represent a desired view. In general, mosaicing process utilizes a series of images that have overlapping regions, according to which each image is stitched with the next image in series. After obtaining the recorded video from the data acquisition system, the video frames are filtered in order to extract the keyframes that are consequently propagated for mosaicing. For a certain camera, using all consecutive frames obtained from the recorded video is inefficient, increases the complexity of the algorithm, and is time-consuming. Moreover, due to similarities between each of the consecutive frames, using all frames introduces redundancy while producing a mosaic of the same quality as the mosaic formed by utilizing the keyframes only.

In this paper, we perform mosaicing for the images from the four cameras mounted on the UAV based on the multi-threaded approach presented by Garcia-Fidalgo [40], which gives fast seamless mosaics. Firstly, the images from the NIR, NDVI, and red cameras are roughly registered to the RGB camera in order to account for the translation occurring from the spatial arrangement of the cameras in the space. The resulting images, in addition to the RGB images are then cropped and resized to a smaller size to speed up the processing. A graph is used to describe the relations between the keyframes for each camera. In this mosaic graph, each keyframe is represented by a node while the links between nodes represent the overlap between nodes. The first keyframe in the video is considered the reference to which other keyframes are aligned to form the output mosaic. In order to do this alignment, the absolute homography of each keyframe needs to be computed, where the absolute homography ${}^{r}H_i$ of a certain keyframe $i$ represents

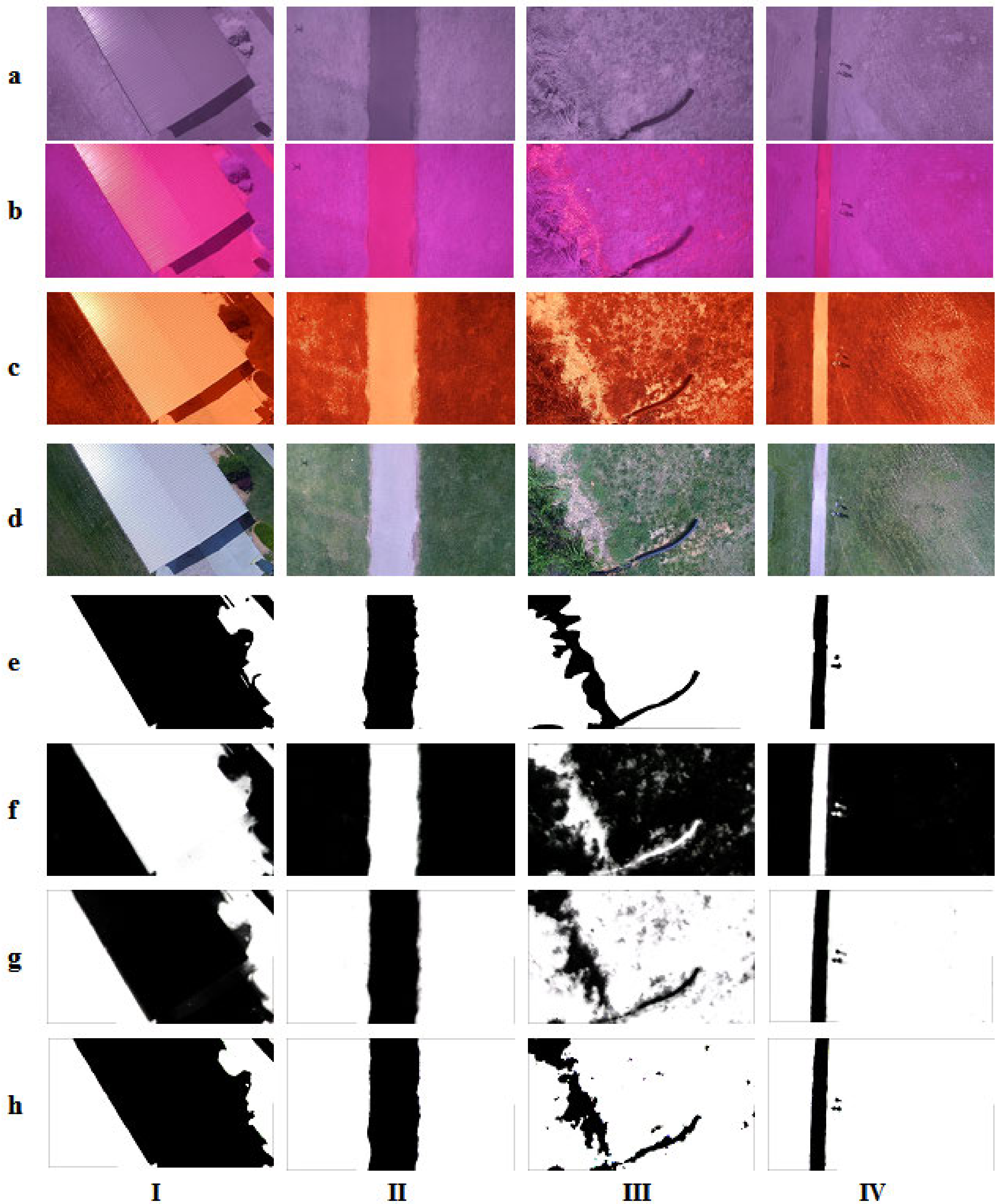


**FIGURE 9.** Segmentation results. (Columns) input and output for 4 samples. (Rows) from top to bottom: NIR images, NDVI images, Red images, RGB images, ground truth mask, probability map for class 0 (non-vegetation), probability map for class 1 (vegetation), and segmentation output.

its homography relative to the reference keyframe $r$ whose absolute homography is the identity matrix. Also a relative homography is given to each link to be used in the pose-graph optimization. The homography ${}^{i}H_{j}$ that relates the two overlapping frames $i$ and $j$ is a linear transformation that relates the point $p_i$ in frame $i$ with its corresponding point $p_j$ in frame

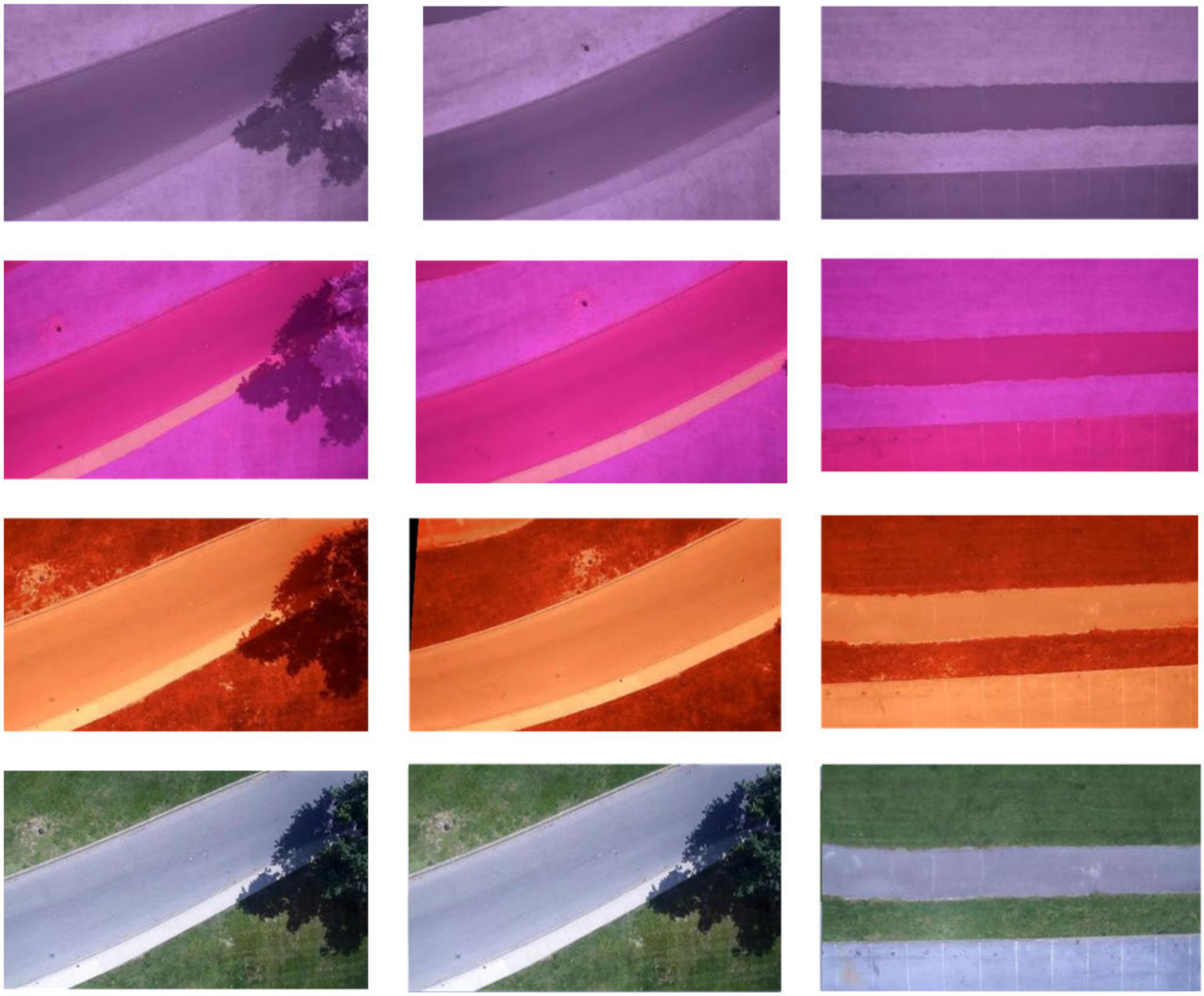

**FIGURE 10. Samples of images used to build the mosaics in Fig. 12. (Rows) from top to bottom: NIR images, NDVI images, Red images, and RGB images.**

$j$ as follows:

$$p_i = {}^{i}H_j p_j, \tag{5}$$

$${}^{i}H_j = \begin{pmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix}, \tag{6}$$

where $s$ is the scale, $\theta$ is the angle of rotation, and $t_x$ and $t_y$ are the translations in the $x$ and $y$ directions, respectively. Having four cameras, means four mosaic graphs. However most of the processing is done to establish the RGB graph, and the resulting graph is transferred to be used with the images from the other three cameras. The RGB mosaic graph is updated online by the following modules of the system which work on separate threads: The keyframe selector, the loop closer, and the optimizer. The final mosaic graph is fed into the blender module that gives the final output.

The keyframe selector describes the current frame and makes the decision whether to include this frame in the graph as a new node or not. This is done in six steps: 1) Detecting and describing the key points of the current frame of the RGB camera using the rotation invariant Oriented FAST and Rotated BRIEF (ORB) [41] binary descriptor which is based on the Features from Accelerated Segment Test (FAST) fast feature detector [42] and the Binary Robust Independent Elementary Features (BRIEF) fast feature point descriptor [43]. The ORB is selected because it has performance close to SIFT but much faster which makes it more efficient for use with applications that are time sensitive; 2) Using the features extracted from the RGB camera current frame and the features extracted from the last keyframe in the graph to compute the homography between the current frame and the last keyframe in the graph; 3) Drawing an upright minimal bounding rectangle around the inliers in the current frame

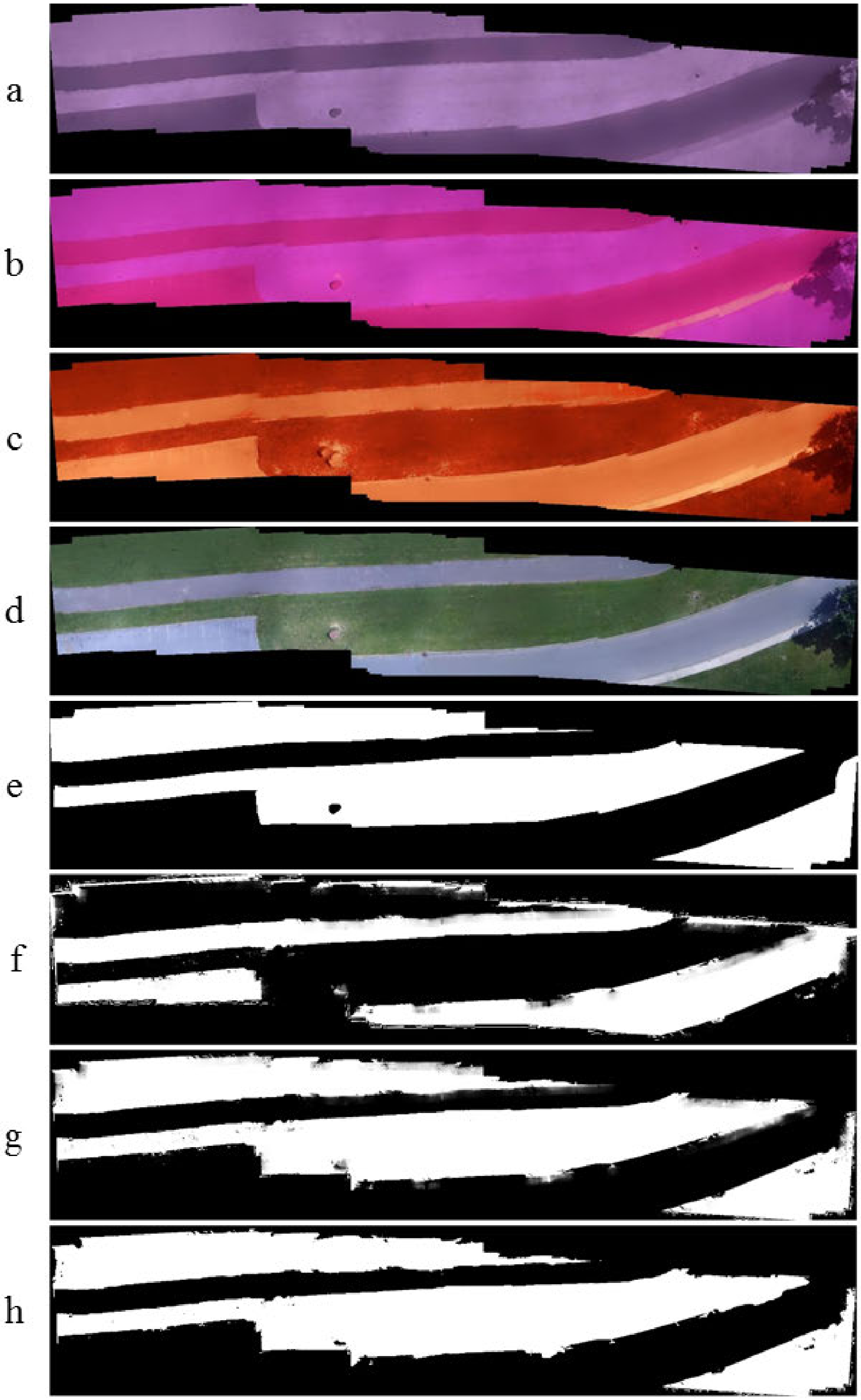


**FIGURE 11.** Constructed mosaic and its segmentation using our approach. (Rows) from top to bottom: NIR images, NDVI images, Red images, RGB images, ground truth mask, probability map for class 0 (non-vegetation), probability map for class 1 (vegetation), and segmentation output.

and another one around the inliers in the last keyframe. The normalized areas of the two rectangles are compared and the smaller is considered as the overlap between the current frame and the last keyframe in the graph, taking into account that the normalized area of a rectangle is equal to its area divided by of the area of the corresponding image; 4) The

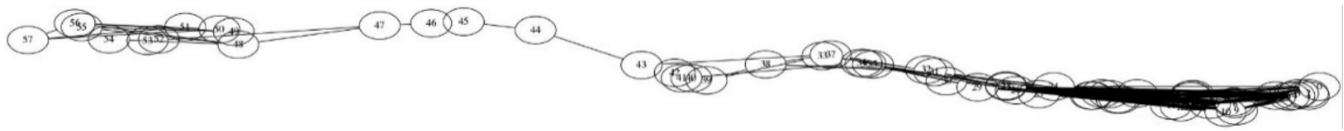

**FIGURE 12.** Mosaic graph used for constructing the mosaics in Fig. 11.

values of the calculated overlap and the number of inliers are compared respectively to two thresholds. If both values are higher than the corresponding thresholds, then the current frame is considered as a valid candidate to be a keyframe; 5) Steps 1 through 4 are repeated until no more valid keyframe candidate is found. The last valid keyframe candidate is then considered as a keyframe and is added as a new node in the graph. This approach guarantees that the minimum overlap required between successive keyframes is established while having the minimum number of nodes in the mosaic graph; 6) The images from the other three cameras corresponding to the keyframes constituting the nodes of the RGB camera mosaic graph are selected as keyframes to be used for building mosaics for images from these three cameras.

The loop closer is responsible for determining the RGB keyframes that close a loop with keyframes which were added earlier. It is based on computing the visual similarity between the current keyframe of concern and other keyframes. The values of visual similarity are arranged in descending order using the approach presented in [44] and the homographies are computed between the corresponding keyframes and the current keyframe. Whenever the number of inliers between two keyframes exceed a certain threshold, a corresponding link is added in the graph.

The optimizer is responsible for minimizing the RGB keyframes alignment errors, in order to avoid global inconsistency in the mosaics. In order to do this, the error function $\mathcal{L}$ between the keyframes $i$ and $j$ that are related by a link was formulated in [40] as:

$$\mathcal{L} = \sum_i \sum_j \sum_{k=1}^{n} \|p_i^k - ({}^rH_i)^{-1}\, {}^rH_j p_j^k\| + R({}^rH_j) + \|p_j^k - ({}^rH_j)^{-1}\, {}^rH_i p_i^k\| + R({}^rH_i), \quad (7)$$

where $n$ is the number of inliers, $p_i^k$ is an inlier point in keyframe $i$ and $p_j^k$ is its corresponding inlier point in keyframe $j$. ${}^rH_i$ and ${}^rH_j$ are the absolute homographies of keyframe $i$, and $j$, respectively. The terms $R({}^rH_i)$ and $R({}^rH_j)$ are used for regularization and are given by:

$$R({}^rH_i) = \beta((s_i \cos\theta_i)^2 + (s_i \sin\theta_i)^2 - 1), \quad (8)$$

$$R({}^rH_j) = \beta((s_j \cos\theta_j)^2 + (s_j \sin\theta_j)^2 - 1), \quad (9)$$

where $s_i$ and $s_j$ are the scales resulting from the homography computation for keyframes $i$ and $j$, respectively; $\theta_i$ and $\theta_j$ are the corresponding orientations; and $\beta$ is a regularization factor. The Levenberg-Marquardt algorithm [45] was used for optimizing the error Huber function $h(\mathcal{L})$, where:

$$h(\mathcal{L}) = \begin{cases} \mathcal{L}^2 & \text{if } |\mathcal{L}| \le 1 \\ 2|\mathcal{L}| - 1 & \text{if } |\mathcal{L}| > 1. \end{cases} \quad (10)$$

Up to this point, the topology between the keyframes of the RGB camera has been estimated, and the keyframes are ready to be blended using the blender module which is responsible for creating the final mosaic seamlessly using the multi-band blending algorithm [46]. This topology is directly transferred to each of the other three cameras (650 nm red, NDVI, 850 nm NIR) by copying the RGB graph to be used with the frames of the corresponding camera. The final mosaics for the frames of the 650 nm red, NDVI and 850 nm NIR cameras are created using the blender modules. Fig. 8 illustrates the mosaicing framework, while Algorithm 1 summarizes its main steps.

**TABLE 1.** Comparison of segmentation results, (using the proposed approach using different image channels in addition to the fusion between all of them, and two recent vegetation segmentation approaches), in terms of Dice similarity coefficient (DSC).

| Method | | DSC |
|---|---|---|
| Proposed (CNN) | NIR | 92.52% |
| | NDVI | 93.68% |
| | RED | 91.69% |
| | RGB | 95.08% |
| | Fused | 97.29% |
| Level sets [4] | | 90.71% |
| Vegetative mapping [47] | | 79.59% |

## III. EXPERIMENT AND DISCUSSION

In this work, we conducted several experiments in which aerial images were collected using our 3DR solo UAV equipped with four cameras (NIR, NDVI, Red and RGB). All cameras were setup to capture images of resolution 2560 pixels × 1440 pixels at a framerate of 30 fps at altitudes of 25 ft, 50 ft, and 75 ft. We selected 315 groups of images, where each group consists of the corresponding NIR, NDVI, RGB, and red images. These image groups were randomly split into training, validation and test sets of equal size. The captured images were manually segmented to generate the reference images that will be used in both training and testing phases. Full segmentation of the validation data sets was periodically performed. Samples of the test images are shown in Fig. 9 with the outer two columns captured at an altitude of 75 ft and the inner two columns at 25 ft.

The DSC for columns I, II, III, IV (Fig. 9) was evaluated using Eq. 4 to be 98.24%, 99.31%, 97.03%, and 99.55% respectively with respect to the reference images. Moreover, the mean DSC was calculated to be 97.29%. To demonstrate the effect and contribution of different image channels to the

**Algorithm 1:** Main Steps of the Mosaicing Approach

1) Preparing the Frames for Processing:
   - Rough registration of the NIR, NDVI, Red cameras to the RGB camera.
   - Cropping the NIR, NVDI, Red and RGB images.
   - Resize the NIR, NVDI, Red and RGB images into a smaller size.
2) Mosaic Graphs Initialization:
   - For each camera, select the first frame to be the reference keyframe (node).
   - Detect and extract the ORB features of the reference keyframes.
3) RGB Mosaic Graph Update: For each RGB current frame ($\geq$ 2nd frame), do the following:
   - Detect and extract the features (ORB).
   - Compute the homography between the current frame and the previous keyframe.
   - Find the overlap between the current frame and the previous keyframe.
   - If Overlap $> \mu_o$ and number of inliers $> \mu_i$, consider the current frame as a valid keyframe candidate.
   - Else:
     a) Insert the last valid keyframe candidate as a new node in the graph.
     b) Select the corresponding frames in the NIR, NDVI, and Red camera as keyframes.
     c) Find the RGB keyframes that close a loop with last inserted RGB keyframe.
     d) Minimize the alignment errors using the optimizer.
4) Find the NIR, NDVI and Red Mosaic Graphs:
   - For each camera, copy the RGB final mosaic graph with replacing the keyframes while the corresponding previously found keyframes.
5) Find the Final Mosaics:
   - Feed the four mosaic graphs into four corresponding blenders to obtain the final mosaics.

**TABLE 2.** Comparison of mosaicing algorithms.

| Video duration | Computation time [4] | Computation time of proposed approach |
|---|---|---|
| 3 sec | 44.78 sec | 29.32 sec |
| 4 sec | 103.09 sec | 37.46 sec |
| 5 sec | 167.25 sec | 41.48 sec |
| 6 sec | 262.41 sec | 52.97 sec |

proposed framework, the segmentation is performed using each separate channel in addition to the four fused channels (Table 1). It is clear that the CNN fusion for all channels achieves the highest DSC. Moreover, Table 1 compares the proposed segmentation performance to that with two other segmentation methods, (the level sets method [4] and the vegetative cover mapping software [47]), and outperformed both of them.

We also compared the performance of the mosaicing approach, with respect to that presented in [4] for different video durations at a resolution of 481 pixels $\times$ 287 pixels as illustrated in Table 2.

Fig. 10 shows samples of the individual images of 22.5 sec long videos that were captured by the UAV at an altitude of 75 ft and stitched using the technique discussed in Section II to give the mosaics in Fig. 11. Such figure shows the four captured images /channels, the reference images segmentation, the probability for each class, and the final segmentation. The corresponding mosaic graph, which has 58 nodes, is shown in Fig. 12. The total duration for constructing the mosaic was 165.5 sec.

Furthermore, we used Eq. 1 and Eq. 2, in addition to the image resizing information that occurred during the mosaicing process to estimate the area of the vegetation region segmented in Fig. 11, which was found to be 1331.39 ft$^2$. For future work, we plan to investigate the effect of including mid-wave and long-wave infrared images in our system on the performance.

## IV. CONCLUSION

In this paper, we presented and implemented a design of a monitoring system for estimating the vegetative regions in an area recorded by a UAV equipped with four different cameras (RGB, NIR, NDVI, and red). The UAV can carry out flights over a certain area through communication with a dedicated flight controller and mobile application. Our approach began by automatically acquiring videos and sending them to a ground station over a wireless connection for processing. A fast multi-threaded mosaicing approach based on binary descriptors was modified to give mosaics for the captured videos simultaneously. In addition, we utilized a CNN based approach for segmenting the photosynthetic areas in these mosaics. We tested two scenarios for training the CNN: The first one fuses the images from the four cameras, while the second deals with each camera individually. The performance of the former scenario that used the four cameras was the best. On the average, our segmentation approach obtained high accuracy in term of the Dice similarity coefficient (higher than 97 %), which makes our approach promising in the analysis of ground cover.

## REFERENCES

[1] J. Chen, S. Yi, Y. Qin, and X. Wang, "Improving estimates of fractional vegetation cover based on UAV in alpine grassland on the qinghai–tibetan plateau," *Int. J. Remote Sens.*, vol. 37, no. 8, pp. 1922–1936, 2016.
[2] A. Karpatne, M. Blank, M. Lau, S. Boriah, K. Steinhaeuser, M. Steinbach, and V. Kumar, "Importance of vegetation type in forest cover estimation," in *Proc. Conf. Intell. Data Understand.*, Oct. 2012, pp. 71–78.
[3] J. Mohammadi, S. Shataee, M. Namiranian, and E. Næsset, "Modeling biophysical properties of broad-leaved stands in the hyrcanian forests of Iran using fused airborne laser scanner data and ultraCam-D images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 61, pp. 32–45, Sep. 2017.
[4] M. Ghazal, Y. Al Khalil, and H. Hajjdiab, "UAV-based remote sensing for vegetation cover estimation using NDVI imagery and level sets method," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2015, pp. 332–337.
[5] K. Kandare, H. O. Ørka, M. Dalponte, E. Næsset, and T. Gobakken, "Individual tree crown approach for predicting site index in boreal forests using airborne laser scanning and hyperspectral data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 60, pp. 72–82, Aug. 2017.
[6] Z. Chao and P. Zhang, "Assessing the impact of urbanization on vegetation change and arable land resources change in Shandong province," in *Proc. 3rd Int. Conf. Agro-Geoinform.*, Aug. 2014, pp. 1–5.
[7] X. Zhang, Z. Xiong, X. Zhang, Y. Shi, J. Liu, Q. Shao, and X. Yan, "Simulation of the climatic effects of land use/land cover changes in eastern china using multi-model ensembles," *Global Planet. Change*, vol. 154, Jul. 2017, pp. 1–9.
[8] G. Jin, X. Deng, Y. Yuan, S. Ma, Z. Li, and S. Shi, "Scenario simulation of land exploitation and risk assessment of soil erosion in the low-slope hilly area of the erhai basin," *Phys. Chem. Earth, Parts A/B/C*, vol. 101, pp. 78–85, Oct. 2017.
[9] D. Gudex-Cross, J. Pontius, and A. Adams, "Enhanced forest cover mapping using spectral unmixing and object-based classification of multi-temporal Landsat imagery," *Remote Sens. Environ.*, vol. 196, pp. 193–204, Jul. 2017.
[10] Y. Amano, N. Takagi, and A. F. Getz, "A study on the classification of urban region using Hyper-spectrum data at AVIRIS," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 4, Jul. 2008, pp. 687–690.
[11] R. C. Grecchi, R. Beuchle, Y. E. Shimabukuro, L. E. Arag ao, E. Arai, D. Simonetti, and F. Achard, "An integrated remote sensing and GIS approach for monitoring areas affected by selective logging: A case study in northern mato grosso, brazilian amazon," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 61, pp. 70–80, Sep. 2017.
[12] Q. Chen and R. McRoberts, "Statewide mapping and estimation of vegetation aboveground biomass using airborne lidar," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 4442–4444.
[13] R. Fieuzal and F. Baup, "Forecast of wheat yield throughout the agricultural season using optical and radar satellite images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 59, pp. 147–156, Jul. 2017.
[14] A. Romero, C. Gatta, and G. Camps-Valls, "Unsupervised deep feature extraction for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1349–1362, Mar. 2016.
[15] J. Tian, L. Wang, X. Li, H. Gong, C. Shi, R. Zhong, and X. Liu, "Comparison of UAV and WorldView-2 imagery for mapping leaf area index of mangrove forest," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 61, pp. 22–31, Sep. 2017.
[16] J. A. J. Berni, P. J. Zarco-Tejada, L. Suarez, and E. Fereres, "Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 3, pp. 722–738, Mar. 2009.
[17] E. R. Hunt, W. D. Hively, S. J. Fujikawa, D. S. Linden, C. S. T. Daughtry, and G. W. McCarty, "Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring," *Remote Sens.*, vol. 2, no. 1, pp. 290–305, Jan. 2010.
[18] A. Rango, A. Laliberte, J. E. Herrick, C. Winters, K. Havstad, C. Steele, and D. Browning, "Unmanned aerial vehicle-based remote sensing for rangeland assessment, monitoring, and management," *Proc. SPIE*, vol. 3, no. 1, Aug. 2009, Art. no. 033542.
[19] A. S. Laliberte, M. A. Goforth, C. M. Steele, and A. Rango, "Multispectral remote sensing from unmanned aircraft: Image processing workflows and applications for rangeland environments," *Remote Sens.*, vol. 3, no. 11, pp. 2529–2551, 2011.
[20] P. J. Zarco-Tejada, R. Diaz-Varela, V. Angileri, and P. Loudjani, "Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods," *Eur. J. Agronomy*, vol. 55, no. 2, pp. 89–99, Apr. 2014.
[21] A. M. Cunliffe, R. E. Brazier, and K. Anderson, "Ultra-fine grain landscape-scale quantification of dryland vegetation structure with drone-acquired structure-from-motion photogrammetry," *Remote Sens. Environ.*, vol. 183, pp. 129–143, Sep. 2016.
[22] T. Ishida, J. Kurihara, F. A. Viray, S. B. Namuco, E. C. Paringit, G. J. Perez, Y. Takahashi, and J. J. Marciano, Jr., "A novel approach for vegetation classification using UAV-based hyperspectral imaging," *Comput. Electron. Agricult.*, vol. 144, pp. 80–85, Jan. 2018.
[23] L. Wan, Y. Li, H. Cen, J. Zhu, W. Yin, W. Wu, H. Zhu, D. Sun, W. Zhou, and Y. He, "Combining UAV-based vegetation indices and image classification to estimate flower number in oilseed rape," *Remote Sens.*, vol. 10, no. 9, p. 1484, Sep. 2018.
[24] B. Lu, Y. He, and H. T. Liu, "Mapping vegetation biophysical and biochemical properties using unmanned aerial vehicles-acquired imagery," *Int. J. Remote Sens.*, vol. 39, nos. 15–16, pp. 5265–5287, 2018.
[25] J. Zhu and M. Ren, "Image mosaic method based on SIFT features of line segment," *Comput. Math. Methods Med.*, vol. 2014, Jan. 2014, Art. no. 926312.
[26] J. Wang, J. Fang, X. Liu, D. Zhao, and Q. Xiao, "A fast mosaic method for airborne images: The new template-convolution speed-up robust features (TSURF) algorithm," *Int. J. Remote Sens.*, vol. 35, no. 16, pp. 5959–5970, Mar. 2014.
[27] D. Turner, A. Lucieer, and C. Watson, "An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (UAV) imagery, based on structure from motion (SfM) point clouds," *Remote Sens.*, vol. 4, no. 5, pp. 1392–1410, May 2012.
[28] Y. Zheng, G. Li, X. Sun, and X. Zhou, "Fast edge integration based active contours for color images," *Comput. Electr. Eng.*, vol. 35, no. 1, pp. 141–149, Jan. 2009.
[29] H. Le, S.-H. Kim, and I.-S. Na, "Object segmentation based on adaptive contour initialization for level set method," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2012, pp. 000067–000072.
[30] N. Xu, N. Ahuja, and R. Bansal, "Object segmentation using graph cuts based active contours," *Comput. Vis. Image Understand.*, vol. 107, no. 3, pp. 210–224, 2007.
[31] *Louisville's Tree Canopy*. Accessed: Mar. 5, 2019. [Online]. Available: https://louisvilleky.gov
[32] A. Holm, D. G. Burnside, and A. A. Mitchell, "The development of a system for monitoring trend in range condition in the arid shrublands of western australia," *Austral. Rangeland J.*, vol. 9, no. 1, pp. 14–20, 1987.
[33] *What is NDVI Index?* Accessed: Mar. 8, 2019. [Online]. Available: http://physicsopenlab.org/2017/01/30/ndvi-index/
[34] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, Feb. 2017.
[35] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
[36] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2811–2821, May 2018.

[37] K. Kamnitsas, C. Ledig , V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glockera, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, Feb. 2017.

[38] A. C. S. Chung and J. A. Noble, "Statistical 3D Vessel Segmentation Using a Rician Distribution," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, C. Taylor and A. Colchester, Eds. Berlin, Germany: Springer, 1999, pp. 82–89.

[39] M. D. Patidar and A. Jain, "Automatic image mosaicing: An approach based on fft," *Int. J. Sci. Eng. Technol.*, vol. 1, no. 1, pp. 1–4, Dec. 2011.

[40] E. Garcia-Fidalgo, A. Ortiz, F. Bonnin-Pascual, and J. P. Company, "Fast image mosaicing using incremental bags of binary words," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2016, pp. 1174–1180.

[41] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.

[42] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2006, pp. 430–443.

[43] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Binary robust independent elementary features," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2010, pp. 778–792.

[44] E. Garcia-Fidalgo and A. Ortiz, "On the use of binary feature descriptors for loop closure detection," in *Proc. IEEE Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2014, pp. 1–8.

[45] J. J. Moré, "The Levenberg-Marquardt algorithm: Implementation and theory," in *Numerical Analysis*. Berlin, Germany: Springer, 1978, pp. 105–116.

[46] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Trans. Graph.*, vol. 2, no. 4, pp. 217–236, 1983.

[47] L. J. Wiles, "Software to quantify and map vegetative cover in fallow fields for weed management decisions," *Comput. Electron. Agricult.*, vol. 78, no. 1, pp. 106–115, Aug. 2011.

**MOHAMMED ASAAD GHAZAL** received the B.Sc. degree in computer engineering from the American University of Sharjah, where he received the President's Cup and the Ministry of Education Shield for Creative Thinking, and the M.Sc. and Ph.D. degrees in electrical and computer engineering (ECE) from Concordia University, Canada, in 2010 and 2006, respectively. He is currently an Associate Professor of ECE and the Chair of the Department at Abu Dhabi University (ADU). He received multiple awards, including the Distinguished Faculty Award, in 2016 and 2014, and the University Teaching Award, in 2013. He lead the B.Sc. programs in ECE to achieve ABET and WASC accreditation. Under his leadership, the department has grown in the number of students, laboratories, faculty, and programs and is home to two bachelors, two minors, and two master's programs serving more than 300 students. Under his supervision, ECE students have received major national and international awards and secured top places in competitions, including a Special Innovation Award of 1,000,000 from HH Sheikh Muhammad Bin Rashid Al Maktoum as part of the Best m-Government Service Award Competition during the Government Summit. He also leads ADU E-Learning initiative. He has authored more than 50 publications in technical journals and international conferences including Elsevier's *Renewable and Sustainable Energy Reviews*, the IEEE Transactions on Image Processing, the IEEE Transactions on Circuits and Systems for Video Technology, and the IEEE Transactions on Consumer Electronics. His current research interests include biomedical signal processing, mobile computing, and embedded systems. He is a member of the ACM and BMES. He received the Alexander Graham Bell Fellowship from NSERC and the QFRNT from the Government of Quebec.

**ALI MAHMOUD** received the B.S. and M.S. degrees in electrical engineering from Alexandria University, Egypt, in 2005 and 2009, respectively, and the Ph.D. degree in electrical engineering from the University of Louisville, Louisville, USA, in 2014, where he is currently a Postdoctoral Research Associate with the Bioengineering Department. His research interests include computer vision and image processing, object detection and tracking, medical imaging, and facial biometrics.

**ALI ASLANTAS** was born in Ankara, Turkey, in 1979. He graduated in computer system education from Suleyman Demirel University, in 2001, and received the M.S. degree from Suleyman Demirel University, in 2006, and the Ph.D. degree from Sakarya University, in 2015. His research interests include image processing and neural networks.

**AHMED SOLIMAN** received the B.Sc. and M.S. degrees in computer engineering from Mansoura University, Mansoura, Egypt. He is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department, University of Louisville, Louisville, KY, USA, where he joined the BioImaging Laboratory, in December 2011, as a Research Assistant. He has actively been working on design of computer aided-diagnostic systems and parallel algorithms development for processing N-dimensional images in the field of medical image analysis.

**AHMED SHALABY** received the B.S. and M.S. degrees in electrical engineering from Alexandria University, Egypt, in 2003 and 2009, respectively, and the Ph.D. degree in electrical engineering from the University of Louisville, Louisville, USA, in 2014, where he is also a Postdoctoral Research Associate with the Bioengineering Department. His research interests include computer vision and image processing, object detection and tracking, medical imaging, and facial biometrics.

**JÓN ATLI BENEDIKTSSON** received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, Iceland, in 1984, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1987 and 1990, respectively. From 2009 to 2015, he was the Pro Rector of science and academic affairs and a Professor of electrical and computer engineering with the University of Iceland, where he became the President and the Rector, in July 2015. He is a Co-Founder of the biomedical startup company Oxymap, Reykjavik. His research interests include remote sensing, biomedical analysis of signals, pattern recognition, image processing, and signal processing. He has many publications extensively in those fields.

Dr. Benediktsson was a member of the 2014 IEEE Fellow Committee. He is a member of the Association of Chartered Engineers in Iceland (VFI), Societas Scinetiarum Islandica, and Tau Beta Pi. He is a Fellow of SPIE. He was a recipient of the Stevan J. Kristof Award from Purdue University, in 1991 for being the Outstanding Graduate Student in remote sensing, the Icelandic Research Council's Outstanding Young Researcher Award, in 1997, the IEEE Third Millennium Medal in 2000, the Yearly Research Award from the Engineering Research Institute, University of Iceland, in 2006, the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society (GRSS), in 2007, the IEEE/VFI Electrical Engineer of the Year Award, in 2013, and the OECE Award from the School of Electrical and Computer Engineering, Purdue University, in 2016. He was a co-recipient of the University of Iceland's Technology Innovation Award, in 2004, the 2012 IEEE Transactions on Geoscience and Remote Sensing (TGRS) Paper Award, the IEEE GRSS Highest Impact Paper Award, in 2013, and the *International Journal of Image and Data Fusion* Best Paper Award, in 2014. He has been on the GRSS AdCom, since 2000. He was the 2011–2012 President of the IEEE GRSS. He was the Editor-in-Chief of the IEEE TGRS, from 2003 to 2008. He has been an Associate Editor of the IEEE TGRS, since 1999, the IEEE Geoscience and Remote Sensing Letters, since 2003, and IEEE Access, since 2013. He was the Chairman of the Steering Committee of IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, from 2007 to 2010. He is on the Editorial Board of the Proceedings of the IEEE, the International Editorial Board of the *International Journal of Image and Data Fusion*, and the Editorial Board of *Remote Sensing*.

**AYMAN EL-BAZ** received the bachelor's and master's degrees in electrical engineering, in 1997 and 2001, respectively, and the Ph.D. degree in electrical engineering from the University of Louisville, in 2006. He is currently a Professor, a University Scholar, and a Chair of the Bioengineering Department, University of Louisville, KY, USA. In 2009, he was named a Coulter Fellow for his contributions to the field of biomedical translational research. He has 17 years of hands-on experience in the fields of bio-imaging modeling and non-invasive computer-assisted diagnosis systems. He has authored or coauthored more than 450 technical articles.