Flood Extent Mapping: An Integrated Method Using Deep Learning and Region Growing Using UAV Optical Data

Leila Hashemi-Beni and Asmamaw A. Gebrehiwot

***Abstract*—Flooding occurs frequently and causes loss of lives, and extensive damages to infrastructure and the environment. Accurate and timely mapping of flood extent to ascertain damages is critical and essential for relief activities. Recently, deep-learning-based approaches, including convolutional neural network (CNN) has shown promising results for flood extent mapping. However, these methods cannot extract floods underneath vegetation canopy using the optical imagery. This article attempts to address this problem by introducing an integrated CNN and region growing (RG) method for the mapping of both visible and underneath vegetation flooded areas. The CNN-based classifier is used to extract flooded areas from the optical images, whereas, the RG method is applied to estimate the extent of floods underneath vegetation that are not visible from imagery using the digital elevation model. A data augmentation technique is applied for training the CNN-based classifier to improve the classification results. The results show that the data augmentation can enhance the accuracy of image classification and the proposed integrated method efficiently detects floods in both the visible and the areas covered by vegetation, which is essential to supporting effective flood emergency response and recovery activities.**

***Index Terms*—Convolutional neural network (CNN), flood mapping, LiDAR, region growing (RG), remote sensing.**

I. Introduction

FLOODING is one of the catastrophic and frequently occurring natural disasters that cause extensive damages to life, infrastructure, and the environment. In many countries, the severity and frequency of flooding have increased in recent years due to extreme weather such as hurricanes, and the expansion of urbanization. Generating accurate and timely inundation maps is essential for regional and federal agencies to manage rescue operations and assess damages effectively [1]–[2].

Manuscript received May 19, 2020; revised August 28, 2020 and November 18, 2020; accepted December 31, 2020. Date of publication January 14, 2021; date of current version February 3, 2021. This work was supported in part by the U.S. National Science Foundation under Grant 1800768 (Remote Sensing for Flood Modeling and Management) and Grant 1832110 (Developing a Robust, Distributed, and Automated Sensing and Control System for Smart Agriculture), and in part by the North Carolina Collaboratory (Flood Resilience: Data Collection and Data Analytics in support of Flood Management). *(Corresponding author: Leila Hashemi-Beni.)*

Leila Hashemi-Beni is with the Department of Built Environment, North Carolina A&T State University, Greensboro, NC 27411 USA (e-mail: lhashemibeni@ncat.edu).

Asmamaw A. Gebrehiwot is with the program of Applied Science and Technology, North Carolina A&T State University, Greensboro, NC 27411 USA (e-mail: aagebrehiwot@aggies.ncat.edu).

Digital Object Identifier 10.1109/JSTARS.2021.3051873

In the past decades, remote sensing has been used as a powerful tool for flood extent mapping. High-resolution satellite imagery (e.g., Sentinel-1, Sentinel-2, and Landsat 8) can provide much of the useful information for detecting and extracting flood-affected areas, assessing the damage (e.g., road, bridge, and infrastructures), and feeding models that can predict the vulnerability to flooding of inland and coastal areas [3]. However, the satellite track may not always coincide with cloud-free conditions or with flood peak, which is related to the maximum inundation area due to revisit limitations. Thus, flood extent mapping is still challenging using satellite optical imagery due to dense cloud cover, complex urban landscapes, and satellite revisit and viewing angle limitations. With technological advances, unmanned aerial vehicles (UAVs) have been considered as effective platforms for flood management applications [4]. Compared to the conventional remote sensing platforms, UAVs have several advantages. UAVs can acquire multiview and high-resolution imagery that are less affected by cloud shadows, and are relatively low cost and flexible in the frequency and time of data acquisition [5], [6]. Therefore, UAVs offer a critical edge and potential in remote sensing including flood mapping. However, in spite of recent developments in UAV data collection technologies, more sophisticated processing is required to compensate for less controlled geometry and vibration of these platforms and sensors. In addition, UAVs' limited payload, short flight endurance, and small-scale coverage remain areas of weakness for large-scale remote sensing projects.

Processing of a vast number of high-resolution UAV imagery using traditional remote sensing methods are time-consuming and sometimes inconvenient for time-sensitive applications like flood management [4]. In the past decades, conventional machine learning classifiers have been used for remote sensing mapping applications [7]–[8]. These methods did not yield significant results in terms of accuracy owing to the complexity of the textual information, especially in highly vegetated and urban areas. Deep convolutional neural networks (CNNs) [9] overcome these problems and have shown to yield promising performance in many tasks, including image classification [10], object detection [11], scene labeling [12], and object recognition [13]. Unlike other machine learning classifiers, CNNs provide a hierarchical representation of the data using various convolutions and can automatically learn feature representation from big datasets [14]. This allows more extensive learning capabilities and, thus, higher performance and precision. Based on that

context, the remote-sensing community has recently started shifting its focus from traditional classifiers such as support vector machine [15] and random forest [16] to deep learning, for mapping and classification tasks such as land cover classification [17], scene classification [18], and crop deep species classification [19]. For inundation mapping, Gebrehiwot *et al.* [20] applied a CNN for flood extent mapping. In this article, the inundation extent is extracted by fine-tuning the Visual Geometry Group-16 (VGG-16) based pretrained fully convolutional model (FCN) [21] using very high-resolution UAV optical imagery. The research was conducted using 100 training samples and resulted in more than 95% accuracy on extracting the surface floods, which were visible on the imageries. However, the FCN method, similar to other image segmentation algorithms, is unable to extract the flood extent underneath vegetation canopies from 2-D optical imagery. This becomes a serious issue in flood-prone areas covered with dense vegetation using remote sensing methods. In addition to this problem, the FCNs were trained and tested on a relatively small data set (100 UAV images). However, the performance of these deep learning algorithms is heavily reliant on the volume of the training data. Although more massive training datasets can help machines to learn model parameters and improve the optimization process and imparts generalization, annotating a large amount of remote sensing data is challenging. Remote sensing images are rich in features including buildings, roads, grasses, vegetation, dry land, etc. and labeling these features is time-consuming and costly.

This research is an attempt to address the above problems. To extract the flood extents from both the visible areas and underneath vegetation, we integrated the visible flooded areas obtained using a CNN-based approach from the UAV data, and the flood extents in dense vegetation using a region growing (RG) method from the Digital Elevation Model (DEM) and water level data. RG, which was proposed by Adams *et al.* [22], is a pixel-based image segmentation method that involves the selection of initial points or seed points. Many researchers have used the RG method for various tasks, including for mapping land use [23], urban environment [24], and detection of water from remote sensing images [25]. Bins *et al.* [23] present an RG method to segment images to assess land-use changes in the Amazon region. The RG was implemented using the geographic information system and tested on forest and agricultural images and achieved satisfactory segmentation results. Vo *et al.* [24] introduced an RG method for the surface patch segmentation of 3-D point clouds for the urban environment. Pan *et al.* [25] presented a seeded RG method to effectively detect water bodies from aerial images based on the texture of the feature. In this article, the RG method is applied to estimate the flood extent underneath dense canopy using the topography information obtained from Lidar data collected before the flood events. Thus, the performance of the RG method highly depends on the accuracy of the topography information (e.g., DEM) and seed points. We implement an RG method considering two possible scenarios: 1) water level measurement of one or more locations is available in the study area (e.g., gauge stations), thus these points are used as seed points; 2) there is no water level measurement or gauge stations in the study areas, in this case, the FCN classification results are overlaid on DEM and slope map of the area, and some seed points are selected from the interface of water/no water classes using spatial analysis.

To address the issue related to the small data size to train a deep learning model, the research examines data augmentation to artificially enlarge the training dataset by either data warping or oversampling—without collecting new data. Data warping techniques use transformations (e.g., geometric transformations and color transformations) to augment training data while preserving labels. Oversampling techniques augment training data by creating synthetic images using methods such as generative adversarial network. Shorten *et al.* [26] provide an extensive overview of data augmentation algorithms for deep learning. Data augmentation using geometric transformations have shown a significant effect on deep learning on a small-scale dataset using several tasks [27]–[30]. Zhou *et al.* [28] applied data augmentation methods such as mirroring and rotation to a small volume aircraft training data. Based on their experimental results, the data augmentation significantly improved test accuracy. Zhong *et al.* [29] proposed an augmentation approach for classifying scenes in remote sensing images. Their method was evaluated with the IKONOS, the UC Merced, and the SIRI-WHU datasets and effectively enhanced the classification accuracy. Zach *et al.* [30] used an augmentation technique to improve semantic segmentation results for medical applications. They used linear combinations of training images and labels to augment the dataset using the mix-up algorithm and achieved significant increases in segmentation performance.

The objectives of the article are as follows.

1) To investigate the effect of data augmentation for improving the accuracy of flood extent extraction when a small dataset is available for deep learning methods. To the best of our knowledge, there is no study on data augmentation for flood extent mapping using high-resolution optical data including UAV data.
2) To develop a new method to extract and map flood extent in visible areas on imagery as well as underneath vegetation by deep learning and RG methods.

II. Data and Study Area

We selected three flood-prone areas in North Carolina (USA) for this research including the town of Princeville, Lumberton, and Fair Bluff. The Town of Princeville is located along the Tar River in Edgecombe County and has been highly affected by several flood events many times. Hurricane Floyd (1999) and Hurricane Matthew (2016) caused widespread devastation when the Tar river flooded the town of Princeville. These flood events caused massive damages to human life and property in this town. Lumberton is the largest city in Robeson County, North Carolina. It is located on the Lumber River in the coastal plains region of North Carolina. Hurricane Matthew and Florence hit Lumberton with major flooding in 2016 and 2018, respectively. Fair Bluff is a town in Columbus County that was also devastated by flooding from Hurricane Matthew in 2016 and Hurricane Florence in 2018.

TABLE I
INFORMATION ABOUT UAV AND AERIAL IMAGERY DATASETS USED FOR THE STUDY

		Bands	Spatial Resolution	Land coverage	Flood event	Data collection date
UAV Imagery	Princeville, NC	3 bands (RGB)	2.6 cm	0.64 km^2	Hurricane Matthew	10/16/2016
	Lumberton, NC	3 bands (RGB)	1.5 cm	0.52 km^2	Hurricane Florence	9/23/2018
	Fair Bluff, NC	3 bands (RGB)	1.5 cm	0.49 km^2	Hurricane Florence	9/26/2018
Aircraft Imagery	Princeville, NC	3 bands (RGB)	50 cm	4.6 km^2	Hurricane Matthew	10/10/2016
	Lumberton, NC	3 bands (RGB)	30 cm	5.2 km^2	Hurricane Florence	09/19/2018

The datasets used for the research include UAV imagery, high-resolution (manned) aerial imagery, and Lidar data. The Lidar and UAV data were acquired by North Carolina Emergency Management (NCEM), whereas the aerial imagery was collected by NOAA. Table I provides the information about the datasets.

The UAV data and aerial imagery are used to study the effect of data augmentation for improving the accuracy of flood extent extraction on these two high-resolution datasets when a small dataset is available for deep learning methods.

To implement RG method, preflood LiDAR data were available in the town of Princeville. This LiDAR with two pulses per square meter (pls/m^2) and an accuracy of 9.25-cm RMSE [31] were collected by NCEM in 2014. This data was used to create a DEM of the area at 10-cm resolution for developing our RG method to estimate the flood flows underneath vegetation. We also collected the information of the USGS Surface Water Gauge station [Fig. 1(b)] that is located near the Princeville study area (#02083500 Tar River) from the NOAA website. USGS Gauge Stations collect time-series data that describe stream levels, streamflow or discharge, reservoir and lake levels, surface–water quality, and rainfall. The elevation of the gauge station selected for this article is 42.7 ft. NAVD88, which was taken from stream levels on October 15, 2016, at 14:30 EDT.

Finally, we used the inundation map created by NOAA for the town of Princeville during the hurricane Matthew flood event. This inundation map was generated using hydrodynamic numerical models and used for validation of the proposed integrated method.

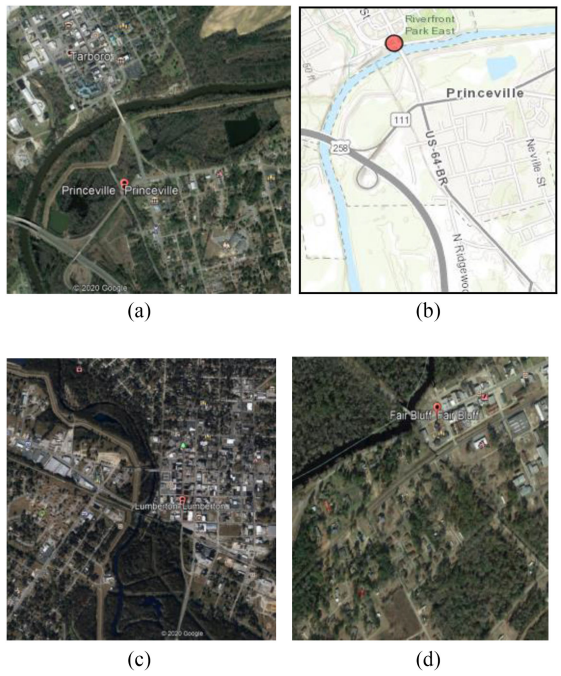


Fig. 1. Study areas. (a) Princeville. (b) USGS surface-water gauge station in Princeville. (c) Lumberton, NC. (d) Fair Bluff, NC.

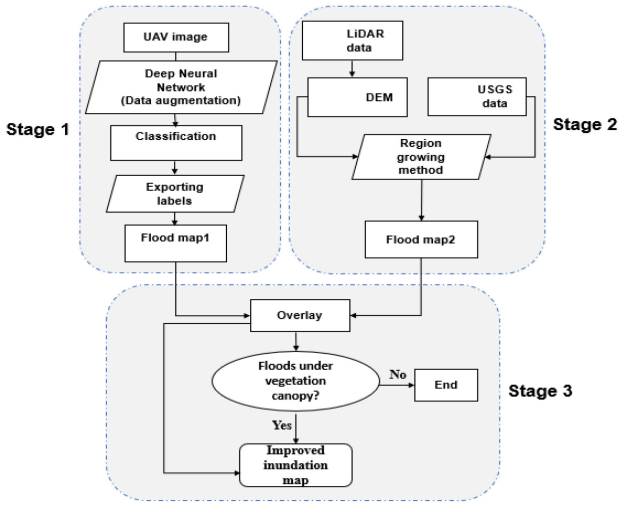


Fig. 2. Proposed approach for inundation mapping.

III. MATERIALS AND METHODS

To extract and delineate flood extents in both open and vegetated areas, the research method consists of three stages (Fig. 2). Stage 1 extracts flood extents using a deep neural network (FCN-8s) approach from high-resolution images (flood map 1). A data augmentation method is applied to increase the training dataset and improve the classification results. Stage 2 delineates the flood extent using an RG method using a DEM/topography data and water level information in one or more locations in the area (flood map 2). In stage 3, Using flood map 2, the FCN-based flood extent is modified and improved for vegetated areas where the flooded areas under canopies are not visible on the images.

A. Flood Extent Mapping Using Deep Convolutional Network

The first stage of the method involves classifying the UAV images by applying a data augmentation technique and georeferencing the classified images. Many deep learning methods have been successfully developed and used for various applications such as AlexNet, Unet, GoogleNet, and FCNs. Among these methods, FCNs have been adopted by several researchers for classification and segmentation tasks in remote sensing applications and showed promising results [20], [32], [33]. Specifically, FCN-8s showed better performance on extracting flooded areas from high-resolution UAV imagery compared to other FCNs

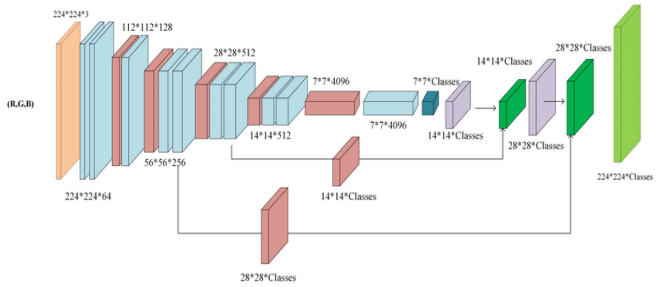


Fig. 3. FCN-8s architecture [35].

variants such as FCN-16s and FCN-32 [20]. Thus, in this article, we utilized FCN-8s to investigate the effect of data augmentation for improving the accuracy of flood extent extraction when a small dataset is available for deep learning methods.

The FCN-8s is composed of locally connected layers, such as convolution, pooling, and upsampling, without having any dense layer (Fig. 3) that allows reducing the number of parameters and computation time. Given that all connections are local, FCN-8s can work on any image size. In this model, VGG-16 fully connected based classification layers were replaced by convolutional layers to maintain the 2-D structure of images. VGG-16 is a CNN architecture proposed by Simonyan *et al.* [34] to investigate the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. The gradual upsampling of the scoring layer and merging of features from earlier layers results in a fine label map using the FCN-8s.

1) Annotating Training Images: 150 UAV images (4000 by 4000 pixels) were manually labeled to train the FCN-8s including 90 images from the Princeville dataset and 60 images from the Lumberton and Fair Bluff dataset [20]. We used the Image Labeler tool in MATLAB [36] to annotate each pixel into four classes: water, building, vegetation, and road classes. This manual classification served as ground truth for training and validation of the FCN-8s model results.

2) Training the Network: Training deep neural networks generally need large datasets (with plenty of diverse training data), which are not always available. In our previous study [20], we only used 150 images to train the FCN-8s classifier. Data augmentation can be an alternative method to increase the amount of training data and improve the segmentation accuracy by reducing the overfitting of CNN models caused by limited training samples. Data augmentation operations such as randomly cropping, translation, and random rotation are commonly used to artificially generate new training data from existing training data. These operations are applied to images in the input space. Random cropping is done by randomly sampling a section from the original image and resizing this section to the original image size. Translation is implemented by moving all pixels of the images in one direction such as horizontally (left or right) or vertically (up or down). When shifting an image, the remaining space is either filled by 0s, 255s, or filled with random or Gaussian noise, thus preserving the original spatial dimension of the image. Rotation is done by rotating the image right or left on an axis between 1° and 359°. To investigate the effect of data augmentation for improving the accuracy of flood extent extraction, the FCN-8s network was trained by applying data augmentation (including random cropping, rotation, reflection, and translation) and the classification results were compared with the results of the network trained without employing data augmentation technique.

To train the FCN-8s network, the 10-fold cross-validation strategy was used to avoid overfitting the data and improve the performance of the FCN-8s model. In a k-fold cross-validation method, we divided the training UAV images (150 images) into k equal subsets called folds (in our case 10 subsets). At each run, the union of k-1 folds (9 folds or 135 images) is put together to form a training set, and the remaining 1-fold (15 images) is used as a testing or validation set to evaluate the performance of the FCN model. This is an iterative algorithm using a different fold as the testing set at each time. In other words, the images included in the first testing fold would never be part of a testing fold again. Finally, the average error from all 10 folds was used to evaluate the performance of the model. The importance of this technique is to give a less biased estimate of the FCN model on unseen data. The network is trained using stochastic gradient descent for six epochs with a learning rate of 0.001, and the maximum batch size of 4. In the classification stage, the performance of the network is tested using the unseen testing images during cross-validation.

For the data augmentation experiment, the FCN-8s was trained by applying random cropping, random rotation, random left/right reflection, and random X/Y translation of data augmentation techniques. We extracted a total of 32 patches (512 *512) per image. The patches were inserted to the network with a batch size of 4. Random translation and rotation data augmentation techniques were implemented by randomly translating the images up to 10 pixels horizontally and vertically and rotating the images with an angle up to 20°. During the training, the datastore performed image augmentation without saving any images to the memory.

The FCN-8s was trained using this augmented dataset and applied to input images to predict the water, road, building, vegetation, and dry features. Then, the accuracy of the classification was assessed using a confusion matrix. Finally, the classification result obtained using data augmentation and without using data augmentation was compared to investigate the effect of data augmentation techniques to increase the classification performance on the small training dataset. To handle the class misbalancing issue, we applied the median frequency balancing approach during training [37]. In this approach, the weight assigned to each category (*ac*) in the loss function is the ratio of the median of the class frequencies (*median_freq(c)*) computed on the entire training set divided by the class frequency (*freq(c)*). The class frequency was calculated by dividing the number of pixels for each class by the total number of pixels in the image. Therefore, the dominant labels are assigned with the lowest weight, which balances the training process.

The research also evaluated the performance of the trained deep-learning-based method (with and without using data augmentation) on other high-resolution aerial imagery captured from an aircraft during three different flood events during

Hurricane Matthew (Lumberton), Hurricane Harvey (Houston), and Hurricane Florence (Lumberton) flood events that occurred in 2016, 2017, and 2018, respectively. These images were not used to train the FCN-8s model in this article.

3) Georeferencing the Classified Images: The classification results were exported to a categorical raster as well as a vector data model (polygons). These data models are suitable for spatial analysis, including area calculation, spatial overlay, and spatial join. The results are georeferenced using eight ground control points available in the study area for geospatial data integration and visualization purposes.

4) Performance Evaluations: In this research, the Confusion matrix, also known as the error matrix, was used to evaluate the performance of the FNC-8s model. A confusion matrix is a table applied to assess the performance of classifiers on a set of test data for which the actual values are known. In addition, the kappa coefficient was used in this article to summarize the information provided by the confusion matrix.

B. Flood Extent Mapping Using Region Growing

The second stage of the workflow (Fig. 2) involves producing a flood map using an RG method. The RG is a pixel-based image segmentation method that involves the selection of the initial points or seed points whose water level is known. This approach examines neighboring pixels of the initial seed points using DEM or topography information of the area and determines whether the pixel neighbors should be added to the (water) region. The performance of the RG approach highly relies on the accuracy of DEM or topography information as well as the seed points. A DEM can be generated from ground survey data, existing topographic maps, or Lidar data. Recently, LiDAR has been widely used to create DEMs, since it can map large spatial areas by less manpower at a lower cost. In addition, it can capture data for the areas hidden by trees and greenery, penetrating through the forest canopy. The ability to record more than one return signal per emitted pulse allows LiDAR to generate an accurate DEM of the landscape. This makes LiDAR highly valuable for detecting flooded areas underneath vegetation canopy [38]. However, LiDAR data from flooded areas are not always available. For this article, we used LiDAR data collected by NCEM in 2014 prior to the flooding events.

In this article, we implemented an RG method to determine the flooded (water) areas in the areas covered with vegetation (i.e., vegetation class resulted from the FCN algorithm) using two scenarios for seed points selection: 1) the water level is available in one or more points in the relatively small area (0.5 km^2); a gauge station is available in the study area; b) the water level information is not available; thus, we propose a new method to determine the seed points for the RG method. For the first case, the flood level recorded at the USGS gauge station near the TAR River was used to generate an RG-based flood extent map. The raster DEM was analyzed to classify flood and nonflood pixels starting from the USGS Gauge station location as the seed point. The flood region was grown from the seed pixel (with the water level of 42.7 ft.) by comparing the elevation of the neighboring pixels to the gauge water level. The pixels with elevations lower than the water level are assigned to the flood class. This process was applied recursively for all the flood pixels in the area to generate the flood extent map. This is a simple and fast method to identify flooded areas in a small study area. However, the method has many limitations. The simplifying assumption of the constant water level does not apply to a large study area for flowing water conditions. Since the extension of a flood depends on the areas hydrologic conditions such as flood speed and magnitude, however, these hydrologic conditions are omitted in the calculation of water level for each pixel using the method [39]. These limitations may lead to either underestimation or overestimation of the flood extent map.

To address these limitations, many studies calculate a variable water level for each pixel by calculating a geomorphic flood index (GFI) [40]–[41]. The GFI is a morphological descriptor and depends on river depth and the difference in elevation between each pixel under examination and the pixel of the river network closest to that pixel. These methods are relatively simple for understanding the floodplain geometry without the need for complex hydraulic models [40]. However, gauge stations are usually unavailable for most of the areas; if they exist, the results of flood extent mapping heavily depends on the distance between the selected gauge station/stream element and the study area and local topography of the area where the gauge is located and the study area. Moreover, the gauge data are usually in-channel measurements, which is sometime not hydrologically connected with urban areas unless there is huge overbank flooding. For these cases, we propose applying the FCN classification results directly to guide the RG. In this article, we manually selected some seed points on the interface between the dry and water areas (wet pixel with a water depth $= 0$) near to the vegetated areas. However, by considering some criteria for spatial proximity between water polygons and vegetation polygons, and the distribution and region of influence of seed points, the locations of seed points can be automatically identified. The elevations of the seed points and water levels for flooded areas were then extracted by intersecting the water polygons (FCN-8s-based flood/water class) with the LiDAR-based DTM (bare earth) generated before the flood event. Slope, mean, and standard deviation of each water polygon were calculated and analyzed to remove noises and verify the seed points elevation information. It should be noted that the seed points must not be located on the water-vegetation or water-building interfaces where there is an ambiguity in the water level estimation. Line segments resulting from the intersection of water and vegetation polygons or water and building polygons do not necessary represent the interfaces between dry and water areas. Thus, assigning a water depth of zero to any seed point located on those line segments could be inaccurate. Using the seed points, another RG flood map was developed.

C. Flood Maps Integration

Although the FCNs provide promising results for flood mapping in the areas where the water surface is visible on the optical images, they are unable to extract the flood extent in the areas covered with dense vegetation canopies. Thus, we applied an

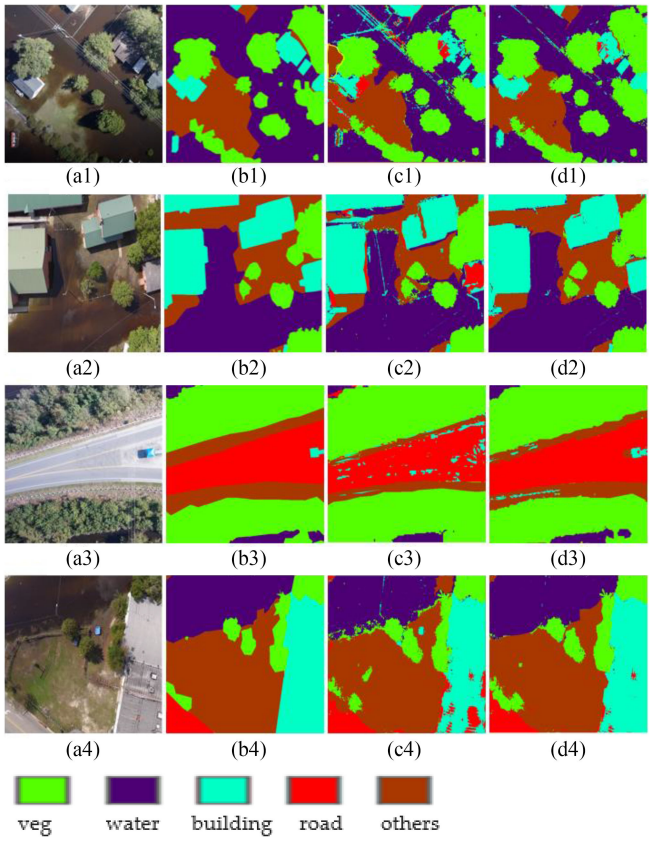


Fig. 4. FCN-8s classification results using UAV images. (a) Test images. (b) Labeled images (ground truth). (c) FCN-8s results without data augmentation. (d) FCN-8s results by applying data augmentation.

integrated method to improve the FCN-based flood map for underneath vegetation canopy, which is the final stage of the workflow (Fig. 2). In this stage, the FCN-based flood map and RG-based flood map are integrated. To combine these two raster data, having same resolution, Map Algebra was used to modify the FCN-based flood extent for dense vegetated areas. If a cell is classified as vegetation in the FCN flood map, its label was updated based on the flood raster created by the GR flood map. This integrated method provides detailed information about the floods in both open areas and under vegetation canopy areas.

IV. Results and Discussion

In this section, the results of classification and flood maps obtained from FCN-8s, RG, and the integrated approach are presented.

A. Classification Results of FCN-8s Approach

To investigate the effect of data augmentation on image classification, the results of the FCN-8s generated with and without applying data augmentation are compared. All the training parameters for the two FCN-8s experiments are the same. The classification results of FCN-8s with and without implementing data augmentations are shown in Fig. 4. Fig. 4(a) shows the testing UAV images from the Lumberton study area during hurricane

TABLE II
Confusion Matrix of FCN-8s Using Data Augmentation (Unit: Percentage)

	Water	Building	Vegetation	Road
Water	**98.63**	0.84	0.88	0.06
Building	0.51	**95.21**	0.80	3.48
Vegetation	1.09	0.41	**98.43**	0.07
Road	0.29	1.87	0.86	**97.98**

TABLE III
Classification Results of FCN-8s for Individual Classes With and Without Data Augmentation Techniques Using High-Resolution UAV Dataset (Unit: Percentage)

	FCN-8s	FCN-8s (data augmentation)
Water	96.52	98.63
Building	87.05	95.21
Vegetation	97.01	98.43
Road	95.10	97.98

Florence that were not used to train the FCN-8s model; the model was trained using UAV imagery from Princeville and Lumberton during hurricane Matthew and Hurricane Florence, respectively. Each test image has a size of 4000 × 4000 pixels. Table II describes the confusion matrix or the detailed information on the FCN-8s model performance using data augmentation techniques for classification.

The confusion matrix was calculated to assess the overall effectiveness of the network. The cells of a confusion matrix show the percentage of true and false predictions for all the possible correlations between the validation and segmented image. The cell in the *i*th row and *j*th column means the percentage of the *i*th class samples, which classified into the *j*th class. The above table shows as 98.63% of class water samples were classified correctly, but 0.84%, 0.88%, and 0.06% of class water samples were incorrectly classified as the building, vegetation, and road classes, respectively. The diagonal cells of the matrix represent the percentage of correctly classified pixels for each class. Our primary goal in this article was to extract flooded areas (water class) from the UAV imagery. The FCN-8s trained by the augmented data achieved 98.63% accuracy in extracting the inundation areas (excluding underneath vegetation canopy or in vegetation class).

The overall accuracy was also calculated to provide a comprehensive assessment of the proposed approach. The overall accuracy is calculated as the total number of correctly classified pixels (diagonal elements) divided by the total number of test pixels. The overall accuracy achieved for the FCN-8s with and without implementing the data augmentation technique were about 97.56% and 95.2%, respectively. Applying data augmentation improved the segmentation overall accuracy by 2.4%.

Using data augmentation techniques, the accuracy of extracting the flooded areas increased from 96.52% to 98.63%. The results also show that data augmentation has improved the classification accuracies of other infrastructures such as buildings and

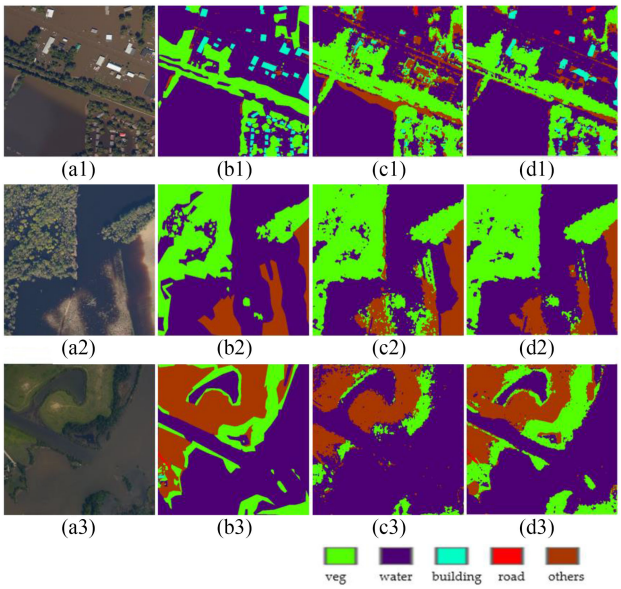


Fig. 5. FCN-8s classification results using the aerial high-resolution images. (a) Test images. (b) Labeled images (ground truth). (c) FCN-8s results without data augmentation. (d) FCN-8s results by applying data augmentation.

TABLE IV
OVERALL CLASSIFICATION RESULTS OF FCN-8S FOR INDIVIDUAL CLASSES WITH AND WITHOUT DATA AUGMENTATION TECHNIQUES USING HIGH-RESOLUTION AERIAL IMAGES (UNIT: PERCENTAGE)

Study Area	FCN-8s	FCN-8s (data augmentation)
Lumberton(Matthew)	76.4%	84.5%
Houston(Harvey)	79.8%	86.74%
Lumberton (Florence)	83.3%	90.5%

roads, which are critical infrastructures when studying flooding issues in urban areas. The classification results for buildings were enhanced by about 7%.

We also tested the FCN trained by UAV imagery using and without data augmentation on the high-resolution imagery captured by an aircraft from Lumberton (during Hurricane Matthew in 2016 and Hurricane Florence in 2018) and Houston (Hurricane Harvey). Fig. 5 and Table IV show the results of the classification.

Table IV describes the classification results to extract flood extent from high-resolution aerial images. The FCNs achieved 76.4%, 79.8%, and 83.3% of overall accuracy for the datasets, which is less accurate in comparison with the results obtained by the high-resolution UAV dataset (96.52%). However, with the augmented data, the overall accuracy of the classification results improved by 8.1%, 6.94%, and 7.2%, respectively. This experiment emphasizes the importance of data augmentation for flood extent mapping using the limited number of images.

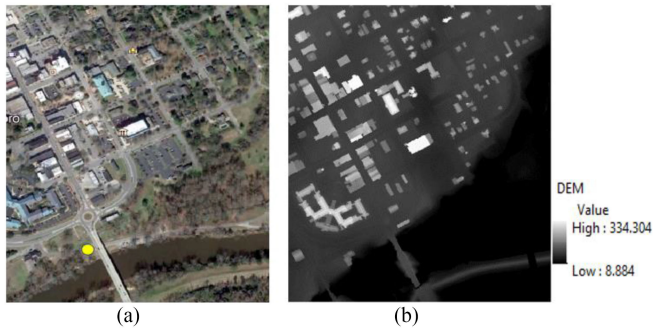


Fig. 6. (a) Test image (Princeville, NC). (b) DEM created by LiDAR data (Unit: Foot).

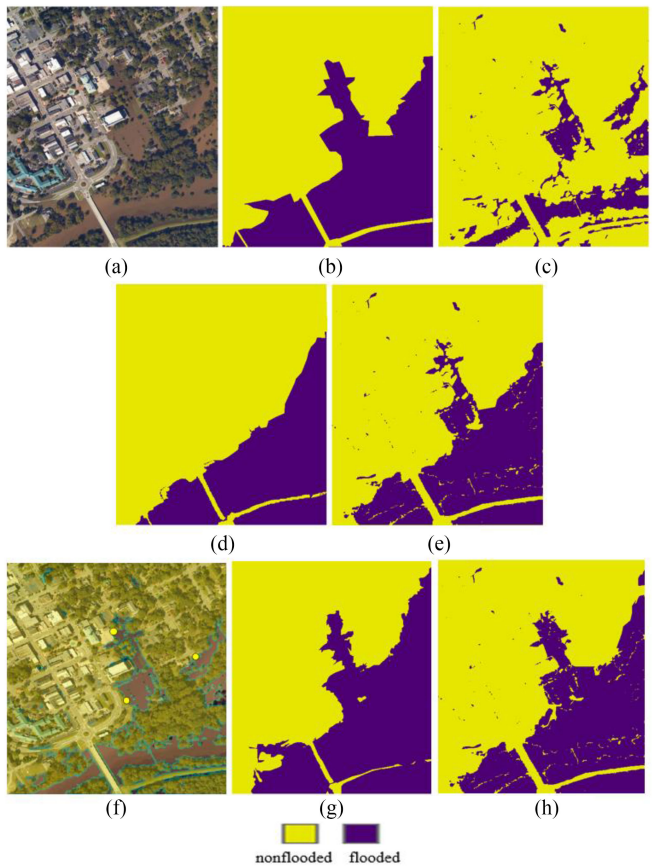


Fig. 7. (a) Test image. (b) Validation data-NOAA inundation map. (c) FCN-8s result. (d) RG classification result for Scenario 1. (e) Flood map using the integrated method for Scenario 1. (f) Seedpoints location on the FCNs-based flood-non flood interface. (g) RG classification result for Scenario 2. (h) Flood map using the integrated method for Scenario 2.

B. Classification Results of the Proposed Integrated Approach

The USGS gauge station 02083500 Tar River [circled with yellow color in Fig. 6(a)] located inside our study area. Fig. 6(b) shows the DEM created by LiDAR data of the study area.

Fig. 7 shows the results generated by FCN, RG, and the integrated method for two scenarios for seedpoints selection. Fig. 7(b) shows the NOAA inundation map for this area during the hurricane Matthew flood event and Fig. 7(c) illustrates

TABLE V
FLOOD EXTENT MAPPING: COMPARISON OF THE RG AND INTENERATED CLASSIFICATION RESULTS (UNIT: PERCENTAGE).

Method	Non-Veg	Veg	Overall
RG (Scenario 1)	70.5	81.7	77.3
RG (Scenario 2)	90.7	92.1	91.8
Integrated method (Scenario 1)	98.6	81.7	88.4
Integrated method (Scenario 2)	98.6	92.1	92.4

the flood map generated using the FCN classification using the UAV optical images. For this, the multiclass classification (water, building, road, and other classes) was converted to a binary classification, flooded, and nonflooded classes. The flood extent map in Fig. 7(d) was generated using the RG-based approach from the DEM and the flood level recorded on the gauge station (43.7 ft) in the area at the time that the UAV data was captured (Scenario 1). Fig. 7(e) illustrates the result of the integrated method discussed in Section III-C. The map was generated by modifying the flood extent results of the FCN classification for underneath vegetation canopies using the RG method. Comparing the results of the integrated method with the NOAA inundation map [Fig. 7(b)], it shows that the integrated approach [Fig. 7(e)] has significantly improved the FCN results for mapping of both visible and vegetation floods.

As Scenario 2, we conducted another analysis for the integrated method when the FCN classification results lead the selection of the seedpoints [Fig. 7(f) and (g)]. The seedpoints can be automatically selected from the flood and nonflood interface in locally flat areas (using slope analysis) considering their proximity to the dense vegetation areas (using spatial proximity analysis). In this article, we manually selected three seed points (considering both flatness and proximity criteria) to investigate the performance of the integrated method [Fig. 7(f)]. The flood extent map was generated using the RG method from DEM [Fig. 7(g)]. The results of the integrated method [Fig. 8(h)] show the promising performance of detecting both visible and underneath vegetation canopy floods.

Table V summarizes the results of the flood classification assessment for the RG and the integrated method for both Scenario 1 and 2 in vegetated, nonvegetated areas, as well as the overall accuracy for the classification. As the results show the FCN-8s has better classification potential compared to the RGs for open areas, i.e., nonvegetated areas. This is because the FCN-8s classification is a data-driven method extracting spatial features and patterns automatically, whereas the performance of the DEM-based RG approach depends on the accuracy of topography info, seed points, and the assumptions for environmental and hydrological parameters. However, the FCN-8s segmentation, similar to other image classification methods, does not have the ability to detect the floods underneath the dense vegetation canopies (flooded vegetation) from the optical imagery, since these flooded areas are simply not visible on the imagery. For flood-prone areas that are mostly covered by vegetation (e.g., North Carolina), it is essential to detect these areas to estimate the extent of the floods and to avoid the unseen floods that come from these areas protecting both human life and property. The integrated method helped to overcome the problems of the occlusion to obtain a classification of the entire area.

V. CONCLUSION

In this article, we proposed an integrated method for mapping the flood extent using FCN deep learning and RG. The deep-learning-based (FCN-8s) model was used to extract the surface flood extent from high-resolution UAV imagery. A data augmentation method was applied during training to improve the classification results of FCN-8s. The results show that applying data augmentation to the FCN during training can outperform those methods when a small dataset is available. Although the FCNs provide promising results for flood mapping in the areas with visible water surface from imagery, they fail to extract the flood extent in the areas covered with dense canopies. To resolve this problem, we implemented an RG approach to detect the floods underneath vegetation canopy using DEM and water level information. The flooded areas extracted under vegetation canopy from the RG method were combined to the flood map generated by the deep learning method using Map Algebra. The experimental results show the integrated approach can efficiently detect floods in both the visible and hidden areas, which is essential to supporting effective flood emergency response and recovery activities.

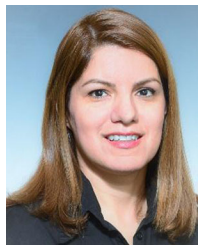
ACKNOWLEDGMENT

The authors would like to thank the North Carolina Emergency Management for providing the UAV data, GCPs, and support data for the project. They would also like to thank Dr. Majid Mohammadian from Department of Civil Engineering, University of Ottawa, Canada, and the anonymous reviewers whose comments/suggestions helped improve and clarify this manuscript.

REFERENCES

[1] Q. Feng, J. Liu, and J. Gong, "Urban flood mapping based on unmanned aerial vehicle remote sensing and random forest classifier—A case of yuyao," *China Water*, vol. 7, pp. 1437–1455, 2015.
[2] P. Boccardo, F. Chiabrando, F. Dutto, F. Tonolo, and A. Lingua, "UAV deployment exercise for mapping purposes: Evaluation of emergency response applications," *Sensors*, vol. 15, pp. 15717–15737, 2015.
[3] V. Klemas, "Remote sensing of floods and flood-prone areas: An overview," *J. Coastal Res.*, vol. 31, no. 4, pp. 1005–1013, 2015.
[4] L. Hashemi-Beni, J. Jones, G. Thompson, C. Johnson, and A. Gebrehiwot, "Challenges and opportunities for UAV-based digital elevation model generation for flood-risk management: A case of Princeville," *North Carolina Sensors*, vol. 18, no. 11, 2018, Art. no. 3843.
[5] H. Yao, R. Qin, and X. Chen, "Unmanned aerial vehicle for remote sensing applications—A review," *Remote Sens.*, vol. 11, 2019, Art. no. 1443.
[6] V. V. Klemas, "Coastal and environmental remote sensing from unmanned aerial vehicles: An overview," *J. Coastal Res.*, vol. 31, pp. 1260–1267, 2015.
[7] P. Boccardo, F. Chiabrando, F. Dutto, F. Tonolo, and A. Lingua, "UAV deployment exercise for mapping purposes: Evaluation of emergency response applications," *Sensors*, vol. 15, pp. 15717–15737, 2015.
[8] G. Ireland, M. Volpi, and G. P. Petropoulos, "Examining the capability of supervised machine learning classifiers in extracting flooded areas from landsat TM imagery: A case study from a Mediterranean flood," *Remote Sens.*, vol. 7, pp. 3372–3399, 2015.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, Dec. 2012, pp. 1097–1105.
[10] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, 2015.
[11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *Adv. Neural Inf. Process. Syst.*, pp. 91–99, 2015.
[12] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1915–1929, Aug. 2013.
[13] Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1701–1708.
[14] T. N. Sainath, A. R. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for LVCSR," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 8614–8618.
[15] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer, 1995.
[16] T. K. Ho, "Random decision forests," in *Proc. IEEE 3rd Int. Conf. Document Anal. Recognit.*, vol. 1, 1995, pp. 278–282.
[17] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.
[18] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep learning-based feature selection for remote sensing scene classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2321–2325, Nov. 2015.
[19] Y. Y. Zheng, J.-L. Kong, X.-B. Jin, X.-Y. Wang, T.-L. Su, and M. Zuo, "CropDeep: The crop vision dataset for deep-learning-based classification and detection in precision agriculture," *Sensors*, vol. 19, 2019, Art. no. 1058.
[20] A. Gebrehiwot, L. Hashemi-Beni, G. Thompson, P. Kordjamshidi, and T. E. Langan, "Deep convolutional neural network for flood extent mapping using unmanned aerial vehicles data," *Sensors*, vol. 19, 2019, Art. no. 1486.
[21] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
[22] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 6, pp. 641–647, Jun. 1994.
[23] L. S. A. Bins, L. G. Fonseca, G. J. Erthal, and F. M Ii, "Satellite imagery segmentation: A region growing approach," *Simpósio Brasileiro de Sensoriamento Remoto*, vol. 8, pp. 677–680, 1996.
[24] A. V. Vo, L. Truong-Hong, D. F. Laefer, and M. Bertolotto, "Octree-based region growing for point cloud segmentation," *ISPRS J. Photogramm. Remote Sens.*, vol. 104, pp. 88–100, 2015.
[25] J. Pan and M. Wang, "Improved seeded region growing for detection of water bodies in aerial images," *Geo-Spatial Inf. Sci.*, vol. 19, no. 1, pp. 1–8, 2016.
[26] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, 2019, Art. no. 60.
[27] J. Wang and L. Perez, "The effectiveness of data augmentation in image classification using deep learning," 2017, *arXiv: 1712.04621*.
[28] M. Zhou, Z. W. Shi, and H. P. Ding, "Aircraft classification in remote-sensing images using convolutional neural networks," *J. Image Graph.*, vol. 22, pp. 702–708, 2017.
[29] Y. Zhong, F. Fei, and L. Zhang, "Large patch convolutional neural networks for the scene classification of high spatial resolution imagery," *J. Appl. Remote Sens.*, vol. 10, 2016, Art. no. 025006.
[30] Z. Eaton-Rosen, F. Bragman, S. Ourselin, and M. J. Cardoso, "Improving data augmentation for medical image segmentation," 2018.
[31] "QL2/QL1 LiDAR collection," North Carolina Emergency Management. Accessed: Aug. 2020. [Online]. Available: https://sdd.nc.gov/SDD/docs/LidarSummary.pdf
[32] Z. Zhong, J. Li, W. Cui, and H. Jiang, "Fully convolutional networks for building and road extraction: Preliminary results," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2016, pp. 1591–1594.
[33] H. L. Yang, D. Lunga, and J. Yuan, "Toward country scale building detection with convolutional neural network using aerial images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2017, pp. 870–873.
[34] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[35] S. Piramanayagam, E. Saber, W. Schwartzkopf, and F. W. Koehler, "Supervised classification of multisensor remotely sensed images using a deep learning framework," *Remote Sens.*, vol. 10, 2018, Art. no. 1429.
[36] Get started with the image labeler - MATLAB & Simulink, Dec 2020. [Online]. Available: https://www.mathworks.com/help/vision/ug/get-started-with-the-image-labeler.html
[37] D. Eigen and R. Fergus, "Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, 2650–2658.
[38] M. W. Lang and G. W. McCarty, "Lidar intensity for improved detection of inundation below the forest canopy," *Wetlands*, vol. 29, pp. 1166–1178, 2009.
[39] R. Romanowicz and K. Beven, "Estimation of flood inundation probabilities as conditioned on event inundation maps," *Water Resour. Res.*, vol. 39, no. 3, 2003.
[40] F. Nardi, E. R. Vivoni, and S. Grimaldi, "Investigating a floodplain scaling relation using a hydrogeomorphic delineation method," *Water Resour. Res.*, vol. 42, no. 9, 2006.
[41] S. Manfreda and C. Samela, "A digital elevation model based method for a rapid estimation of flood inundation depth," *J. Flood Risk Manage.*, vol. 12, 2019, Art. no. e12541.



Leila Hashemi-Beni (Member, IEEE) received the B.Sc. degree in civil-surveying engineering (geomatics) from Isfahan University, Isfahan, Iran, the M.Sc. degree in photogrammetry engineering from Tehran University, Tehran, Iran, and the Ph.D. degree in geomatics and GIS from Laval University, Quebec, QC, Canada.

From 2009 to 2013, she was a Postdoc Research Associate with Agriculture and Agrifood Canada, Ottawa, ON, Canada, and a Geomatics Project Lead and GIS Manager in industry between 2013 and 2016. She is currently an Assistant Professor of geomatics and the Director of Geospatial Science and Remote Sensing Laboratory, Department of Built Environment, College of Science and Technology, North Carolina A&T State University, Greensboro, NC, USA. Her research experience and interests include multisource, multitemporal, and multimodal data fusion, 3-D geospatial data modeling, deep learning, remote sensing image classification, and developing GIS and RS for different applications including environmental management, transportation, hydrology, agriculture and agriFood, and urban planning.

Asmamaw A. Gebrehiwot received the B.Sc. degree in surveying and the M.Sc. degree in geodesy and remote sensing. He is currently working toward the Ph.D. degree in applied science & technology Ph.D. program (Remote Sensing & Machine Learning) with North Carolina A&T State University, Greensboro, NC, USA.

He is currently a Research Assistant with North Carolina A&T State University, where he is working on an NSF-funded project entitled "UAV remote sensing for flood mapping and management." His research focuses on automated inundation mapping in 3-D using deep learning using remote sensing data. His research interests include machine learning and deep learning, remote sensing, and photogrammetry. He completed an internship on disaster mapping in section of radar science & engineering at NASA Jet Propulsion in 2019.