

AI-ForestWatch: semantic segmentation based end-to-end framework for forest estimation and change detection using multi-spectral remote sensing imagery

Annus Zulfiqar^{a,†}, Muhammad M. Ghaffar^{b,*}, Muhammad Shahzad^{a,c},
Christian Weis^b, Muhammad I. Malik^{a,c}, Faisal Shafait^{a,c}
and Norbert Wehn^b

^aNational University of Sciences and Technology, School of Electrical Engineering and Computer Science, Islamabad, Pakistan

^bTechnische Universität Kaiserslautern, Department of Electrical and Computer Engineering, Microelectronic Systems Design Research Group, Kaiserslautern, Germany

^cNational Center of Artificial Intelligence, Deep Learning Laboratory, Islamabad, Pakistan

Abstract. Forest change detection is crucial for sustainable forest management. The changes in the forest area due to deforestation (such as wild fires or logging due to development activities) or afforestation alter the total forest area. Additionally, it impacts the available stock for commercial purposes, climate change due to carbon emissions, and biodiversity of the forest habitat estimations, which are essential for disaster management and policy making. In recent years, foresters have relied on hand-crafted features or bi-temporal change detection methods to detect change in the remote sensing imagery to estimate the forest area. Due to manual processing steps, these methods are fragile and prone to errors and can generate inaccurate (i.e., under or over) segmentation results. In contrast to traditional methods, we present AI-ForestWatch, an end to end framework for forest estimation and change analysis. The proposed approach uses deep convolution neural network-based semantic segmentation to process multi-spectral spaceborne images to quantitatively monitor the forest cover change patterns by automatically extracting features from the dataset. Our analysis is completely data driven and has been performed using extended (with vegetation indices) Landsat-8 multi-spectral imagery from 2014 to 2020. As a case study, we estimated the forest area in 15 districts of Pakistan and generated forest change maps from 2014 to 2020, where major afforestation activity is carried out during this period. Our critical analysis shows an improvement of forest cover in 14 out of 15 districts. The AI-ForestWatch framework along with the associated dataset will be made public upon publication so that it can be adapted by other countries or regions. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/JRS.15.024518](https://doi.org/10.1117/JRS.15.024518)]

Keywords: deep neural networks; semantic segmentation; multi-spectral remote sensing; multi-temporal forest change detection.

Paper 210135 received Mar. 8, 2021; accepted for publication May 11, 2021; published online May 31, 2021.

1 Introduction

Earth's forests are under significant pressure of degradation due to human development activities and natural disasters. Owing to this, the deforestation remains one of the most intractable environmental problems of today. About one third of the size (around 13 million hectares) of the world's global forest area has disappeared.¹ The forest area reduction has various adverse effects. First, forests are carbon dioxide sinks that decrease greenhouse gases in the atmosphere.

*Address all correspondence to Muhammad M. Ghaffar, ghaffar@eit.uni-kl.de

[†]Both authors contributed equally to this research.

Second, deforestation affects biological diversity, forest habitat, and increases soil erosion. Despite continuous efforts by the world community to curb deforestation, it continues unrelenting in the world, posing a serious threat to human livelihoods, eco-systems, and global climate. There is an urgent need to curb deforestation using automated techniques and tools to detect changes more frequently in the forests for sustainable management of forest inventory.

Historically, the forests are monitored by manual and/or aerial field surveys. These surveys are time-consuming and requires a lot of financial and manpower resources. Due to this, these surveys are conducted after multiple years and in case of developing or under-developed countries after several years. To automate the forest inventory and data management processing, the Food and Agriculture Organization of the United Nations developed an open source tool Open Foris.² However, field workers are still required to manually collect the data from forest locations, which are later processed automatically. In the context of automating the monitoring process, there are number of existing approaches that provide a solution to monitor forests either globally³⁻⁵ or regionally⁶⁻⁸ using remote sensing imagery. However, these approaches are limited to bi-temporal change detection or uses hand crafted feature extraction for statistical machine learning techniques to detect changes. The statistical machine learning algorithms are not designed to process high-dimensionality data and huge amount of training samples generated by the remote sensors. In contrast to statistical machine learning algorithms, deep learning-based algorithms can easily manage this huge amount of data due to parallelism offered by the graphical processing units. Furthermore, with the advances in remote sensing and artificial intelligence fields, the whole forest estimation and change detection pipeline can be completely automated. Specifically, this automated chain will be very useful for countries with limited resources as they do not have to rely on expensive tools and surveys to get initial estimates of forest area.

In order to resolve the above-mentioned problems, in this work, we present AI-ForestWatch, an end to end framework for continuous monitoring of the forests by utilizing advancements in the remote sensing and neural network (NN) domains. Our approach is aimed at reducing the financial and manpower burden on the foresters and provides a framework for initial estimations on multi-temporal changes in the forest cover. Specifically, we aim to exploit the available spatiotemporal multi-spectral Landsat satellite data to detect the forest area and changes. The main contributions of this work are as follows.

- An end-to-end semantic segmentation-based framework has been proposed for detailed analysis of forests cover change by utilizing the open-source frameworks (i.e., Pytorch for implementation), tools (i.e., QGIS and Google Maps for dataset preparation), and remote sensing data (i.e., Landsat-8).
- The analysis is completely data driven and is done by employing deep learning-based semantic segmentation on extended Landsat-8 imagery to achieve pixel-wise labeling of spatiotemporal multi-spectral images.
- Detailed exploration and comparison of our semantic segmentation-based approach with other statistical machine learning algorithms (i.e., random forest, decision tree, and logistic regression) for land cover classification has been carried out using Landsat-8 imagery as input and generated digitized maps as ground truth.
- Finally, all the analysis including the source code and the associated dataset will be made public on GitHub (<https://github.com/dll-ncai/AI-ForestWatch>) upon publication.

The remainder of this paper is organized as follows. The related literature in terms of land cover classification and forest change detection is presented in Sec. 2. Section 3 gives a detailed explanation on the study area, background, and proposed approach for forest change detection using semantic segmentation. Section 4 summarizes the obtained results on basis of our implementation. Section 5 analyzes the obtained results and presents detailed discussion on the AI-ForestWatch framework. Finally, conclusions and outlook to the future work are presented in Sec. 6.

2 Related Work

The methods related to forest change detection can be divided into two categories:⁹⁻¹¹ (1) low-level feature-based methods and (2) object-based methods. The former employ pixel-based

statistical inference uses multi-spectral images.^{12,13} They perform pixel-wise contrast comparison to detect changes in temporal sequence of images. Few other approaches use spectral mixture analysis,⁶ vegetation index,¹⁴ and change vector analysis¹⁵-based statistical indices for analyzing Landsat time-series imagery for forest change detection. Alternatively, object-based methods^{9,10,12} rely on incorporating contextual knowledge by processing similarity-based (e.g., spatial and appearance) homogeneous pixels together.¹⁶ Typically, this is achieved by employing unsupervised machine learning techniques for object-based approaches via prior segmentation to obtain pixel clusters. These approaches perform better as they overcome spatial and spectral variations due to noise or georeferencing effects but suffer from inaccurate (under or over) segmentation results.⁹ Shimada et al.⁵ and Hansen et al.¹⁷ provided global forest cover change statistics while using ALOS PALSAR data and Landsat data, respectively. Hansen et al. provided forest loss statistics at 30 m per pixel spatial resolution. On the other hand, Shimada et al. provided forest cover maps at a spatial resolution of 25 m until 2018.

2.1 Machine Learning-Based Forest Change Detection

Machine learning approaches based on classifiers, such as support vector machines (SVM),¹⁸ random forest classification,¹⁹ and decision trees²⁰ have also been employed for land cover classification in both pixel and object-based categories. For instance, Zhao et al.²¹ compared four machine learning classifiers including classification and regression tree, SVM, artificial NN, and random forest to estimate forest parameters of black locust plantations on the Loess Plateau. Qamer et al.²² performed a study for analysis of forest cover change in Western Himalaya, Pakistan. In this study, six forest and eight non-forest classes are labeled using GPS-based ground truth measurements, which are used to train a semiautomatic supervised classifier. They considered forest to non-forest transition to be deforestation and dense to sparse forest to be degradation or cover loss of the forests. The results depicted a relative loss in forest cover in the Khyber Pakhtunkhwa province regions with annual forest cover change rate of -0.42% from 1990 to 2010. Similarly, forest cover change study has been carried out by Souza et al.⁶ for the Brazilian Amazon for the 2000 to 2010 period using Landsat imagery. Image registration and spectral mixture analysis were performed as preprocessing steps. With the main aim of generating large scale maps of forest change harnessing the power of cloud computing, Lehmann et al.¹³ defined forest as tree cover with at least 20% of canopy reaching at least 2-m height and used National Carbon Accounting System and National Forest Trend datasets for forest cover change detection in the Australian continent. More recently, Voight et al.²³ analyzed forest cover change using Landsat-8 imagery using a multi-layer perceptron for training and using annotations by manual labeling of images via Google Earth Engine to forecast magnitude and spatial patterns of deforestation.

2.2 Forest Change Detection Using Deep Learning

In the context of deep learning, the problem of forest cover classification is divided into two categories: one is patch-based classification that divides the full size multi-spectral images into spatially equal sized patches, i.e., 64×64 pixels and a deep learning convolution NN model assigns a single class to the whole patch based on the most dominant distribution of the class.^{24,25} Deep neural networks (DNNs), such as ResNet,²⁶ GoogleNet,²⁷ and other similar topologies have been shown to perform well on this problem but because a whole patch of multiple pixels is labeled with just one class, they produce low-resolution under-segmented output maps. Another method to detect forest change is pixel-wise classification using encoder-decoder topologies.²⁸ A similar approach is used by de Bem et al.²⁸ to map deforestation in the Brazilian Amazon using Landsat data. However, the study focused only on mapping the deforestation by channel wise fusing 1 year apart images between 2017 and 2018 and between 2018 and 2019. This technique is different from our approach as we use only 1 year images with extended vegetation indices for estimation of the forest cover and generate change maps for all the other years (i.e., 2014 and 2016 to 2020). Furthermore, in this work, we use UNet-based semantic segmentation on patches of the input multi-spectral Landsat-8 imagery to perform pixel-wise classification. The advantage of using this approach is that the high resolution of the image can be fully exploited to generate detailed forest cover maps at the same spatial resolution.

3 Materials and Methods

3.1 Study Area

In this study, we selected the KP province, which initiated the Billion Tree Afforestation Project (BTAP) in 2014 to restore 350,000 hectares of deforested land in northern Pakistan. BTAP project is also a winner of the international Bonn challenge²⁹ due to its environmental and social impact. Before this project, the forest area of Pakistan was <2% of its total land area. This percentage is dangerously low when compared with other parts of the world and 25% recommendation of the United Nations. This afforestation project has considerably improved the forest cover of Pakistan, which will help to reduce the chances of floods, increase the precipitation, provide a guard against soil erosion, and help to control the impacts of the climate change.

We performed our analysis for 15 out of 25 districts of the BTAP regions where most of the afforestation drive was carried out. Figure 1 shows the districts used for this work as shaded regions. The remaining districts showed <1% forest cover for the year 2015 and were not considered for this study. Additionally, Chitral and Upper-Dir districts were also not considered for this study, since most parts of these districts are covered with snow and we were unable to label any useful forest data points in this region. Although, the paper maps labeled forest in these regions. However, it would only introduce noise in the training dataset and can produce wrong segmentation results. The BTAP study regions along with their coordinates are listed in Table 1.

3.2 Methods

Figure 2 shows a detailed overview of the complete framework for forest estimation and change detection. Landsat-8 satellite mission has a temporal resolution of 16 days. As already mentioned that we only have the data available for year 2015 only. Hence, we start with the time series Landsat-8 images of year 2015 [Fig. 2-(1)]. The images with <10% cloud cover are filtered, and pixel-wise median values are calculated for each band to generate a composite image or a clean image [Fig. 2-(2)]. The paper printed maps [Fig. 2-(3)] from the BTAP regions are taken

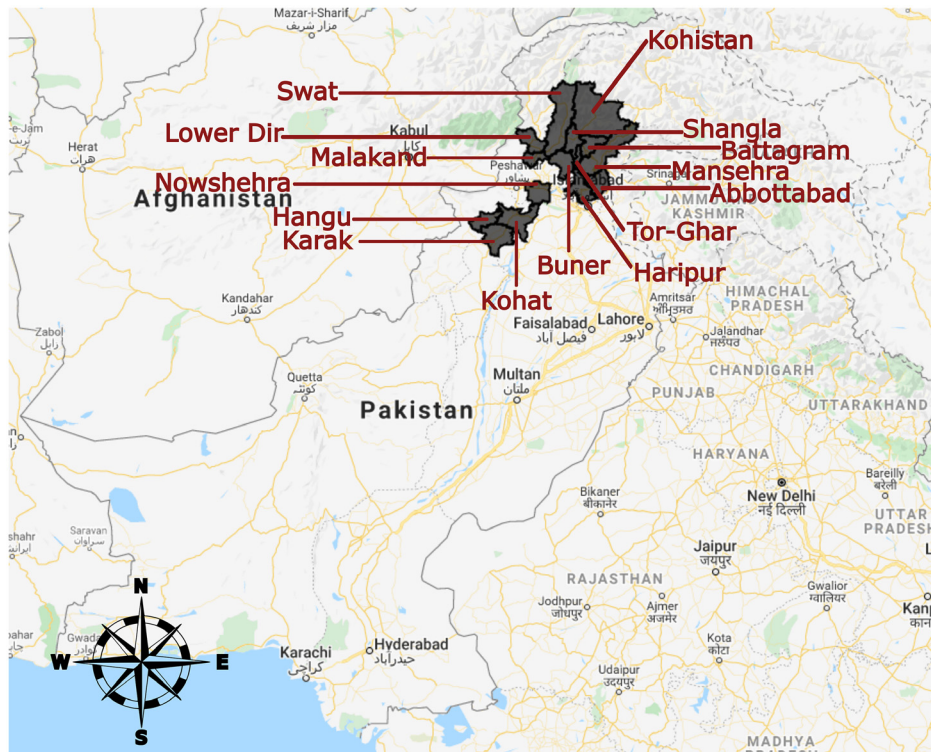


Fig. 1 Geographic location of the districts of BTAP regions under study. The shaded regions indicate the 15 districts considered for this study.

Table 1 The coordinates of the study area.

BTAP locations	Coordinates (longitude, latitude)
Hangu	(70.49 deg, 33.21 deg), (71.24 deg, 33.61 deg)
Karak	(70.75 deg, 32.79 deg), (71.49 deg, 33.38 deg)
Kohat	(71.05 deg, 33.75 deg), (72.03 deg, 33.05 deg)
Nowshehra	(71.68 deg, 34.14 deg), (72.26 deg, 33.69 deg)
Battagram	(72.85 deg, 34.97 deg), (73.51 deg, 34.55 deg)
Abbottabad	(72.96 deg, 34.37 deg), (73.52 deg, 33.81 deg)
Kohistan	(72.68 deg, 35.89 deg), (73.97 deg, 34.894 deg)
Haripur	(72.47 deg, 34.45 deg), (73.27 deg, 33.75 deg)
Tor-Ghar	(72.71 deg, 34.78 deg), (72.94 deg, 34.33 deg)
Mansehra	(72.81 deg, 35.18 deg), (74.17 deg, 34.20 deg)
Buner	(72.20 deg, 34.72 deg), (72.78 deg, 34.14 deg)
Lower Dir	(71.50 deg, 35.07 deg), (72.20 deg, 34.62 deg)
Malakand	(71.63 deg, 34.66 deg), (72.24 deg, 34.36 deg)
Shangla	(72.52 deg, 35.17 deg), (73.02 deg, 34.52 deg)
Swat	(72.08 deg, 35.89 deg), (72.86 deg, 34.56 deg)

as a reference and forest/non-forest data points are labeled on the clean Landsat-8 image using a georeferencing software [Fig. 2-(4)]. As a result a digitized map is generated [Fig. 2-(5)] consisting of green, red and black pixels. The digitization techniques is discussed in detail in the Sec. 3.2.1. After creating the ground truth image, training of the model is performed using an extended (with vegetation indices explained in Sec. 3.2.2) Landsat-8 image as input. A UNet model is trained [Fig. 2-(6)] with an 18-channel input image and the generated digitized forest cover map [Fig. 2-(5)] of year 2015 as ground truth. After a model is trained successfully, inference is performed on the clean Landsat-8 district images of years 2014 to 2020 [Fig. 2-(7)], yielding a temporal series of forest cover maps. In the final step, pixel-wise difference [Fig. 2-(8)] of the multiple maps is carried out to generate the change map. Each step of the AI-ForestWatch framework is discussed in detail in the following sections.

3.2.1 Datasets

Satellite images are available in analysis-ready formats from multiple sensors, i.e., MODIS, Landsat that allow full coverage of the globe at very high spatial and temporal resolution. The Landsat program was launched in 1972 and provides the longest continuous coverage of the whole world through its eight satellite programs.³⁰ In this study, we have used Landsat-8 top of atmosphere imagery, which is available from 2013 and has a temporal resolution and spatial resolution of 16 days and 30 m per pixel, respectively. For our explorations, we selected images for each year with <10% cloud cover [Fig. 2-(1)] and the set of these less cloudy images were then passed through a median filter that calculated the median value for each pixel location for each band of the image to generate clean composites representing the land cover for 1 year. The resulting image served as input [Fig. 2-(2)] to our framework for forest estimation and change detection.

The foresters of the KP province maintain the land cover maps of the BTAP districts in paper printed form. These maps are publicly available for the year 2015³¹ with 10 labeled classes. For our analysis, binary maps with only forest and non-forest classes are created by fusing all the

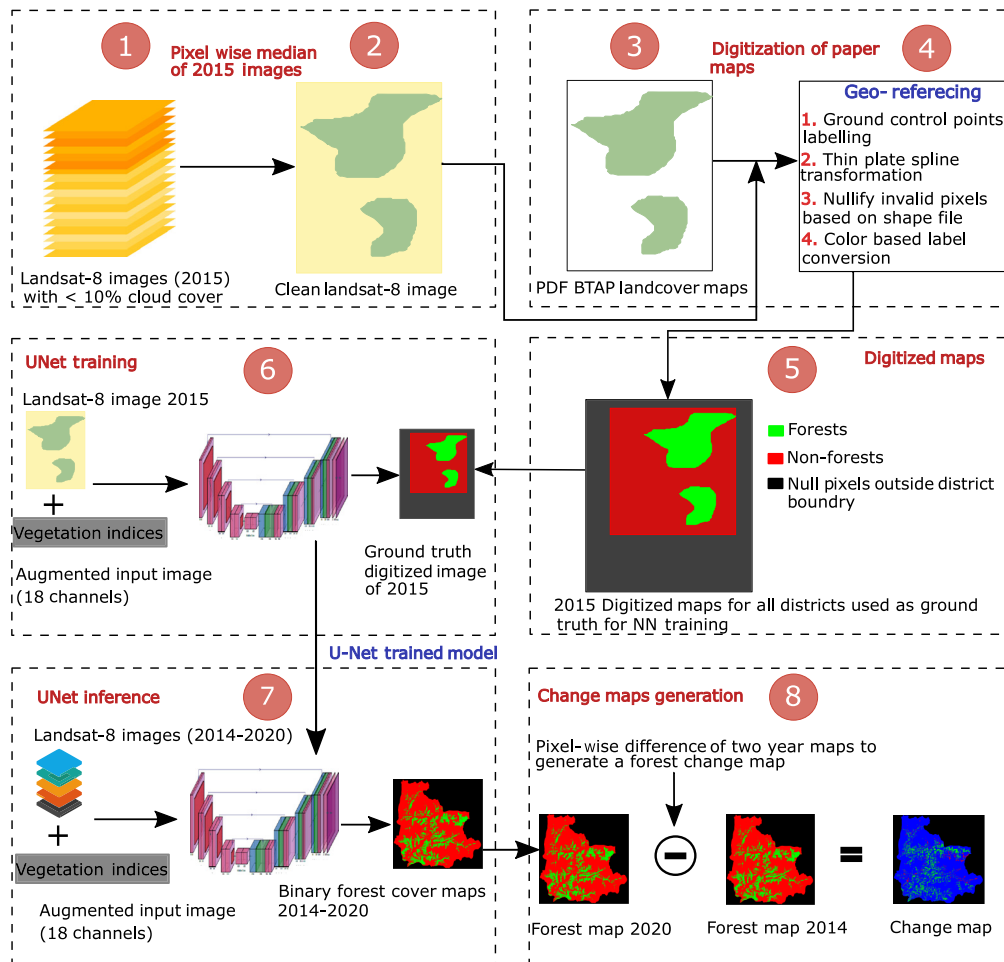


Fig. 2 Detailed flow diagram of the AI-ForestWatch framework. Step-(1) to step-(5) depict the digitization process of the available year 2015 maps. Step-(6) shows the UNet training and the step-(7) and step-(8) show the generation of forest estimation and change detection, respectively.

non-forest classes together. Since, no digital version of the maps are available, we used georeferencing technique to digitize these maps. This technique is discussed in detail in the next section.

Digitization of available land cover maps for BTAP districts. The BTAP land cover maps of year 2015 have 10 labeled classes such as alpine pasture, shrub and bushes, forests, agriculture land, and others. Since our aim is to detect forests in these maps, all classes but forest are fused together to form a single non-forest class. We utilized Google Earth Engine³² for extracting Landsat-8 imagery and an open-source tool, QGIS, for performing this digitization task [Fig. 2-(3)]. The digitization process consists of the following steps.

1. The paper printed BTAP district land cover maps were used as reference for digitization. A sample publicly available land cover map for Battagram district, KP, is depicted in Fig. 3. We only labeled two classes for digitization, forest and non-forest classes, by combining all non-forest classes into one.
2. The available land cover maps were in the form of images that were georeferenced against Landsat-8 imagery by precisely selecting distinct features such as rivers bends and lakes, as our ground control points (GCP) in the land cover maps and Landsat-8 images.
3. Thin plate spline was utilized as the transformation to map the land cover map on top of Landsat-8 imagery via the labeled GCPs. For all the maps digitized, the georeferencing error was $<1 \times 10^{-13}$.

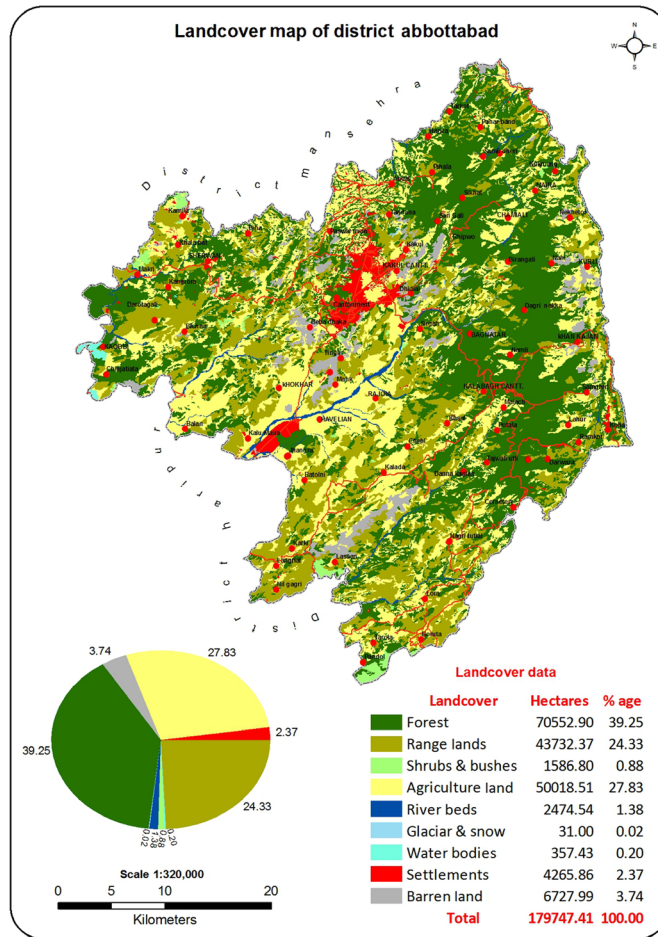


Fig. 3 Abbottabad land cover map of the year 2015 provided by the KP Government.³¹

4. The landcover maps were already color coded with dark green color used for forest pixels. We converted all dark-green pixels in the digitized maps to forest labels and all the other pixels to non-forest. This resulted in a georeferenced digitized map.
5. Since shapefiles of districts for which the maps were digitized were available, we multiplied all the pixels outside the district boundaries with zero to create NULL or invalid pixels outside district boundaries. This was necessary since the labeled maps only contained valid forest/non-forest pixel labels for all locations inside district boundaries, so outer pixels must be considered invalid in the digitized map as well. It is important to mention at this point that the actual Landsat-8 imagery is at 30 m per pixel spatial resolution which makes district images large in size (several thousands by several thousands pixels in every district) but the available labeled land cover maps are low-resolution images, which are georeferenced on the satellite images. This results in some differences in the forest cover percentages (FC%) labeled in the forest cover maps and our digitized maps and cannot be avoided due to the limitation of our ground truth.

The results at different steps of the above-mentioned process are shown in Fig. 4. These digitized maps are in the form of binary images with each pixel labeled as 2 or 1 to indicate forest/non-forest class for each pixel while 0 indicates an invalid/NULL pixel. NULL pixels are taken for training but test scores are reported only for the forest/non-forest classes as a binary classification task. A similar procedure is adopted for all the selected fifteen districts of the KP BTAP districts. These digital maps are used as target annotations for the rest of our research. It is important to mention here that no numerical assessment of the accuracy of these maps is performed since the available ground truth, the actual land cover maps are paper printed maps.

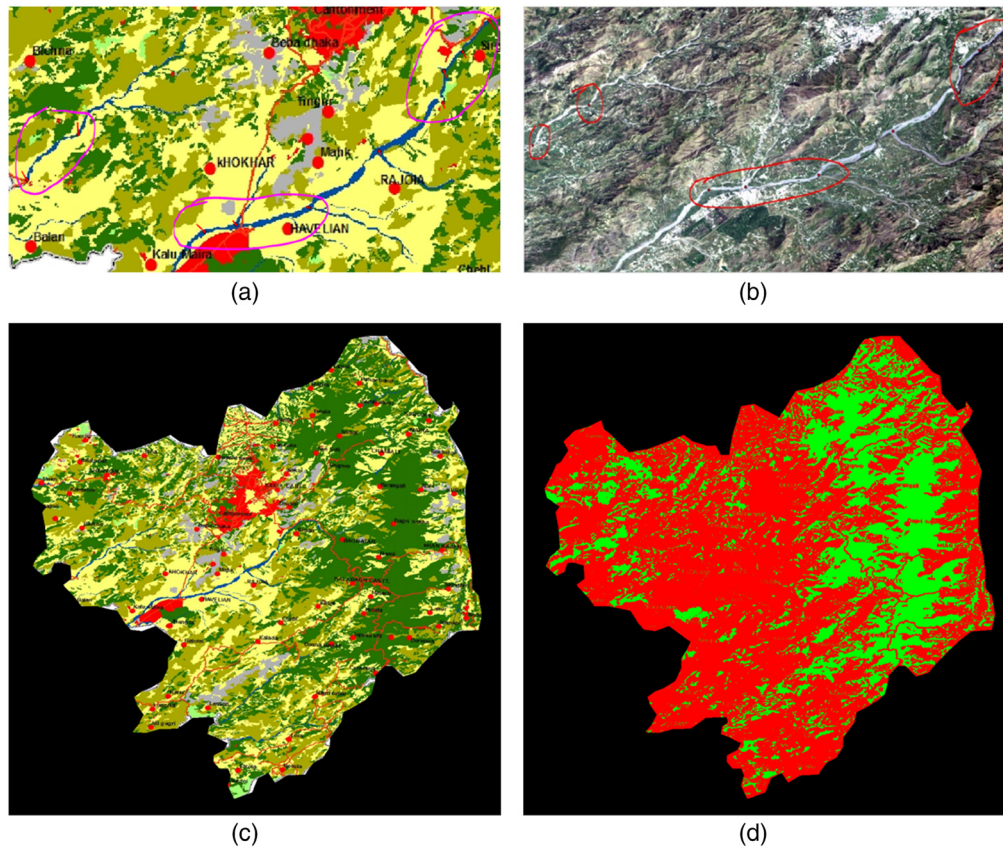


Fig. 4 The digitization steps for BTAP land cover maps into binary forest cover maps. (a) A portion of the ground truth land cover map of district Abbottabad. GCPs being labeled are encircled in pink color. (b) The same portion of district Abbottabad as it appears in its corresponding Landsat-8 image of 2015. Same GCPs are encircled in red color for comparison. (c) The land cover map of Abbottabad after georeferencing the map onto the Landsat-8 image of 2015. It has also been processed using its shape file to nullify all pixels outside the district boundary in order to extract only valid pixels for training. The dark green pixels are color coded in the ground truth map indicating forest pixels. (d) Final digitized map of district Abbotabad after converting all dark green pixels to forest label, shown in green, and all the other non-forest classes into non-forest label, shown in red. The NULL pixels are shown in black which includes all pixels outside the district boundary.

3.2.2 Data preprocessing

Landsat-8 satellite has a temporal resolution of 16 days, which means that every location on Earth is captured by its multi-spectral sensor camera roughly 22 times in a year. Each of these images has a cloud cover score, which can be used to filter them for least cloud cover. For our analysis, we filtered images for each year with <10% cloud cover [Fig. 2-(1)] and the set of these less cloudy images were then passed through a median filter that calculated the median value for each pixel location for each band of the image to generate clean composites representing the land cover for 1 year. For example, for year 2015, we first filtered the images that had <10% cloud and then for each of the 11 original bands of these images, for every single pixel, median value was calculated across these images to generate clean composites. The indices explained above were calculated on these composites and appended with them afterward. This median filter allows us to minimize the atmospheric and seasonal variation effects and generate single images that represent one whole year. This step was performed to generate clean images for all years from 2014 to 2020 throughout this work [Fig. 2-(2)]. For some districts, 10% cloud cover criteria was relaxed to 20% if clear images were not available for one whole year.

Landsat-8 provides imagery in 11 spectral bands³⁰ as presented in Table 2. Considering the high temporal resolution of 30 m, the available spectral bands are appended with seven more

Table 2 Landsat-8 spectral bands.

Band	Description	Wavelength (μm)
B1	Coastal/aerosol	0.433 to 0.453
B2	Visible blue	0.450 to 0.515
B3	Visible green	0.525 to 0.600
B4	Visible red	0.630 to 0.680
B5	NIR	0.845 to 0.885
B6	SWIR-1	1.56 to 1.66
B7	SWIR-2	2.10 to 2.30
B8	Panchromatic	0.50 to 0.68
B9	Cirrus	1.36 to 1.39
B10	Thermal InfraRed (TIR-1)	10.3 to 11.3
B11	TIR-2	11.5 to 12.5

vegetation indices in the form of seven bands that are well-known to provide useful information for land cover classification.³³ Details of these appended seven indices are given below. The B_x represents band number of the Landsat-8 image ($x = 1, 2, \dots, 11$), which are required to calculate the appended index value.

Normalized difference vegetation index. Normalized difference vegetation index (NDVI) is a commonly used index for vegetation detection.³⁴ It exploits the chlorophyll's absorption property of red wavelength and reflectance of infrared wavelength. Its value is always between -1 and $+1$, where close to -1 means water, around 0 means barren land and close to $+1$ means healthy vegetation. It is calculated as given under

$$\text{NDVI} = \frac{B5 - B4}{B5 + B4}. \quad (1)$$

Enhanced vegetation index. Enhanced vegetation index (EVI) allows calculation of enhanced vegetation index³⁵ calculation by introducing other factors into the equation to reduce background/atmospheric noise and saturation. It is particularly useful in areas of high canopy or leaf density. Its range is also between -1 and $+1$ with values between 0.2 and 0.8 indicating healthy vegetation.

In terms of Landsat-8 bands, it can be calculated as

$$\text{EVI} = \frac{2.5 * (B5 - B4)}{B5 + 6 * B4 - 7.5 * B2 + 1}. \quad (2)$$

Soil adjusted vegetation index. Soil adjusted vegetation index (SAVI) helps to rectify the effects of surface reflectance in cases where there is low vegetation cover in the area of interest ($<40\%$).³⁶ The reflectances of red and infrared wavelengths of light affect calculation of NDVI and SAVI modifies the original NDVI for Landsat-8 as given as follows:

$$\text{SAVI} = \frac{1.5 * (B5 - B4)}{B5 + B4 + 0.5}. \quad (3)$$

Modified soil adjusted vegetation index. Modified soil adjusted vegetation index (MSAVI) accommodates for the soil brightness in NDVI in areas of low vegetation cover.³⁷ Although, similar to SAVI, MSAVI also depends on the red ($B4$) and near-infrared (NIR)

(B5) bands of the multi-spectral imagery. However, the MSAVI minimizes the effects of bare soil on the SAVI and in the red-NIR reflectance space, the characteristics of MSAVI are different from SAVI.³⁸ MSAVI can be calculated using the equation given as follows:

$$\text{MSAVI} = 0.5 * \left(2 * B5 + 1 - \sqrt{(2 * B5 + 1)^2 - 8 * (B5 - B4)} \right). \quad (4)$$

The work from Barati et al.³⁹ presents the accuracy of different vegetation indices for estimation of vegetation cover including SAVI and MSAVI. Olmos-Trujillo et al.⁴⁰ used vegetation indices (both SAVI and MSAVI) to perform a geostatistical analysis of rainfall and temperature.

Normalized difference moisture index. Normalized difference moisture index (NDMI) is a ratio calculated using the NIR and short-wavelength infrared (SWIR) bands. It is used for calculating the water content in vegetation⁴¹ using the following equation:

$$\text{NDMI} = \frac{B5 - B6}{B5 + B6}. \quad (5)$$

Normalized burn ratio. Normalized burn ratio (NBR) is used for detecting the presence and severity of burned areas.⁴² The following expression is used for calculating NBR:

$$\text{NBR} = \frac{B5 - B7}{B5 + B7}. \quad (6)$$

Normalized burn ratio-2. Normalized burn ratio-2 (NBR2) modifies NBR to estimate water content in burned areas and is helpful for burned areas recovery studies.⁴³ It is calculated as follows:

$$\text{NBR2} = \frac{B6 - B7}{B6 + B7}. \quad (7)$$

For detailed exploration, different configurations with respect to number of input channels were explored. Machine learning and deep learning algorithms were trained and tested with different numbers of input channels to select an ideal model for the task of forest estimation and change detection. These configurations are as follows.

- **RGB.** Only three red (R), green (G), and blue (B) channels of the Landsat-8 image are provided as input. For instance, in case of RGB input configuration, the input image size for our network is $128 \times 128 \times 3$.
- **Full-spectrum.** Complete 11 bands (presented in Table 2) of the LandSat8 input are used as input to the models for training and test. This configuration extends the number input channels to 11 extending our input image size to $128 \times 128 \times 11$.
- **Vegetation indices.** Seven indices presented in Sec. 3.2.2 are used as input to the model. This configuration uses only the vegetation indices hence the input image size to the network is $128 \times 128 \times 7$.
- **Augmented.** Vegetation indices are stacked along the full-spectrum configuration to create an 18-channel input image. The input image size considered for this configuration is $128 \times 128 \times 18$.

3.2.3 Forest estimation using semantic segmentation

The AI-ForestWatch framework has a two-stage forest change detection pipeline. First, pixel-wise segmentation is performed on a trained model from Fig. 2-(6) to generate the binary forest/non-forest cover maps [Fig. 2-(7)]. Second, pixel-wise difference [Fig. 2-(8)] between two consecutive binary forest/non-forest cover maps generates a change map.

For pixel-wise segmentation, a UNet topology as shown in Fig. 5 with encoder and decoder⁴⁴ stages is used. An encoder is a fully convolution feature extractor network. The operators used in

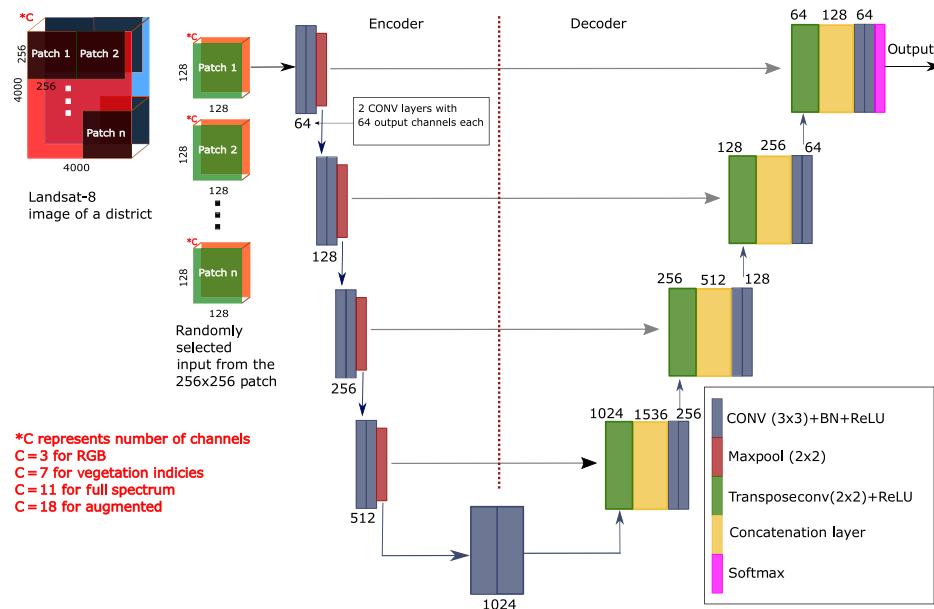


Fig. 5 UNet topology for the AI-ForestWatch framework to detect the forest change. The arrows from encoder toward decoder indicate tensor outputs (feature maps) from the encoder being copied and concatenated with the transposed convolution decoder output along the depth dimension.

this encoder are convolution, batch normalization, activation, and pooling. Four of such modules comprise the encoder of our UNet topology, and it is known as a four-stage encoder. The other part of the UNet, the decoder, is the mirror image of the first half, a four-stage decoder. Each module in a decoder consists of transposed convolution operation followed by a copy-and-fuse connection between the corresponding stages of encoder and decoder. It is followed by convolution, batch normalization, activation, and pooling layers. This concatenation operation of encoder and decoder outputs allows the decoder to utilize the features extracted by the encoder at subsequent stages. The detailed network topology is shown in Fig. 5. The encoder (dotted square on the left) is where the input tensor is down-sampled and encoded into a smaller dimensional vector. The decoder (dotted square on the right) decodes this vector and produces full resolution segmentation for the input image. Each decoder module concatenates its output with the encoder output at the corresponding stage and the final decoder tensor is passed through a convolution again followed by a softmax layer, which normalizes the output probabilities for both classes at each pixel.

Implementation and training. The deep learning part of the proposed AI-ForestWatch framework is implemented in Python using the PyTorch framework. The district images at 30 m per pixel spatial resolution are quite large in volume so we divide them into smaller patches of size $256 \times 256 \times C$ pixels, where C represents the number of channels of the input image, i.e., $C = 3$ for RGB, $C = 11$ for full spectrum, $C = 7$ for vegetation indices, and $C = 18$ for augmented input image. All district images are of different sizes, with an average image size of $4000 \times 4000 \times C$ pixels, a dataset of 3375 $\left(\left(\frac{4000}{256} \times \frac{4000}{256}\right) \times 15\right)$ image patches for training and testing are created. 80% of these patches were randomly chosen for training and while 10% for validation and 10% for testing. Table 3 presents the number of patches used for each of the aforementioned sets. The random selection of the patches for each set results in distinct parts of districts going into training, validation, and test sets, so that we can capture maximum distribution. The network itself accepts inputs of size $128 \times 128 \times C$, which are randomly cropped from these patches at training time [Fig. 2-(6)]. Furthermore, for fast training of the model, the encoder layers are initialized by pretrained weights of the VGG⁴⁵ layers. At inference, these patches are generated directly at $128 \times 128 \times C$ dimensions to cover the whole image in a raster scan manner.

Table 3 Number of patches in the train, validation, and test sets.

Band	Percentage (%)	Number of images
Training	80	2700
Validation	10	338
Test	10	337
Overall	100	3375

The network has been trained using back propagation with the RMSprop optimizer. The learning rate was scheduled to drop exponentially from an initial value of 1×10^{-06} . The model was trained for 200 epochs with batch size of 64. The loss gradients were clipped at 0.05 to avoid exploding gradients. The whole training and testing are performed for the year 2015 since only the annotated maps are available. We also experimented with different loss functions including cross-entropy loss function, focal loss,⁴⁶ and dice loss⁴⁷ in both weighted and unweighted settings. We found that the unweighted focal loss outperforms the other losses and hence it was adopted for this work.

3.2.4 Change detection statistics

Since the change is detected between consecutive years, our metrics are also based on two images under inspection. We combine these metrics temporally to assess the overall change trends. When two consecutive forest cover maps generated by our classification model are subtracted, we get a pixel-wise change map. Based on this, we can calculate the net change, gain, and loss in our regions of interest. These metrics are explained as follows.

Forest cover percentage. This is calculated on the individual forest cover maps generated by the classification model. This is the ratio of the number of forest pixels to the sum of forest and non-forest pixels in the classified image. This gives a percentage of the forest cover in the district under inspection but no information about change. It is calculated for all the years in the temporal series of images as follows:

$$FC\% = \frac{\text{forest pixels}}{\text{forest pixels} + \text{non-forest pixels}}. \quad (8)$$

Gain percentage. Gain percentage (GP%) is calculated on the change map and it is the ratio of the gain pixels (-1) to the total number of change map pixels (-1, 0, and 1):

$$GP\% = \frac{\text{gain pixels}}{\text{gain pixels} + \text{loss pixels} + \text{no change pixels}}. \quad (9)$$

Loss percentage. Loss percentage (LP%) is calculated on the change map and it is the ratio of the loss pixels (+1) to the total number of change map pixels (-1, 0 and 1):

$$LP\% = \frac{\text{loss pixels}}{\text{gain pixels} + \text{loss pixels} + \text{no change pixels}}. \quad (10)$$

Effective forest cover change percentage. Effective forest cover change percentage (ECP%) is also calculated on the change map to calculate the overall change in forest percentage with respect to the former year's forest percentage. It is calculated as follows:

$$ECP\% = \frac{\text{forest\% latter year} - \text{forest\% former year}}{\text{forest\% former year}}. \quad (11)$$

4 Results

In this section, we present detailed evaluations of the AI-ForestWatch framework. We present comparisons of the statistical ML-based classification algorithms against UNet semantic segmentation for all the configurations. Furthermore, we also present forest cover and change statistics in the study area from years 2014 to 2020.

4.1 Comparison with Classification Algorithms

As already discussed in Sec. 2, statistical machine learning algorithms are commonly employed to solve the land cover classification problem. In addition to the semantic segmentation-based approach, three machine learning algorithms, decision tree, logistic regression, and random forest were also explored in this study. For these machine learning algorithms, 5000 points were chosen for each classifier as training input and 5000 for testing. The Landsat-8 was used as input while the digitized data as the ground truth. Furthermore, extensive exploration was conducted on selection of ideal model by considering different input channel configurations for all above-mentioned machine learning and deep learning algorithms.

The results of explorations of the algorithms for RGB, full-spectrum, vegetation indices, and augmented are presented in Tables 4–7, respectively. Decision tree and logistic regression showed the worst results in most configurations. Quantitatively, semantic segmentation-based approach outperformed statistical machine learning algorithms in terms of accuracy. However, this measure should be carefully considered as we are performing binary classification and there exists a class imbalance between the forest and non-forest classes. Due to this reason,

Table 4 Test accuracy and per-class $F1$ scores comparison of the baseline ML-based land cover classification algorithms with UNet-based semantic segmentation approach for RGB channels.

Classifier	Precision		Recall		$F1$ score		Test accuracy (%)
	Forest	Non-forest	Forest	Non-forest	Forest	Non-forest	
Decision tree	0.41	0.82	0.50	0.75	0.45	0.78	68.78
Logistic regression	0.75	0.75	0.05	0.99	0.09	0.86	75.45
Random forest	0.57	0.84	0.50	0.87	0.53	0.85	77.60
Semantic segmentation (UNet)	0.69	0.86	0.41	0.95	0.52	0.90	83.88

Table 5 Test accuracy and per-class $F1$ scores comparison of the baseline ML-based land cover classification algorithms with UNet-based semantic segmentation approach for full-spectrum input image.

Classifier	Precision		Recall		$F1$ score		Test accuracy (%)
	Forest	Non-forest	Forest	Non-forest	Forest	Non-forest	
Decision tree	0.45	0.84	0.56	0.77	0.50	0.80	71.31
Logistic regression	0.75	0.75	0.05	0.99	0.09	0.86	75.45
Random forest	0.65	0.87	0.60	0.89	0.62	0.88	81.67
Semantic segmentation (UNet)	0.62	0.86	0.43	0.93	0.51	0.89	82.52

Table 6 Test accuracy and per-class $F1$ scores comparison of the baseline ML-based land cover classification algorithms with UNet-based semantic segmentation approach for vegetation indices.

Classifier	Precision		Recall		F1 Score		Test accuracy (%)
	Forest	Non-forest	Forest	Non-forest	Forest	Non-forest	
Decision tree	0.43	0.82	0.52	0.77	0.47	0.80	70.58
Logistic regression	0.68	0.82	0.39	0.94	0.50	0.87	79.89
Random forest	0.62	0.85	0.53	0.89	0.57	0.87	79.74
Semantic segmentation (UNet)	0.64	0.87	0.49	0.93	0.55	0.90	83.45

Table 7 Test accuracy and per-class $F1$ scores comparison of the baseline ML-based land cover classification algorithms with UNet-based semantic segmentation approach for augmented configuration.

Classifier	Precision		Recall		F1 score		Test accuracy (%)
	Forest	Non-forest	Forest	Non-forest	Forest	Non-forest	
Decision tree	0.45	0.84	0.56	0.77	0.50	0.80	71.45
Logistic regression	0.66	0.84	0.48	0.92	0.55	0.87	80.44
Random forest	0.65	0.87	0.60	0.89	0.62	0.88	81.74
Semantic segmentation (UNet)	0.66	0.87	0.48	0.93	0.55	0.90	83.79

combination of $F1$ score, precision, and recall along with accuracy are used as standards to evaluate the performance of the models.

Among semantic segmentation models, although, the RGB model yielded the best accuracy (83.88%). However, it lacked in terms of the $F1$ score (0.52) and recall (0.41) for the forest class. The vegetation indices model performed best in terms of $F1$ score (0.58) for and recall (0.53) for the forest class but lacked in precision. The augmented model (18 input indices) produced the best results with regard to every measure with the exception of recall for the forest class. Due to this reason, the augmented model was selected for the AI-ForestWatch framework to perform forest estimation and change detection task. The estimation and change detection results discussed from here onward are derived using this model.

4.2 Estimation and Change Detection Results

After selection of the augmented model on basis on accuracy and the $F1$ score, the forest area cover and change maps for years 2014 to 2020 of the 15 districts of the BTAP project are generated. As the model is only trained and tested on year 2015 data, inference is performed on clean images from year 2014 and 2016 to 2020 using the augmented input configuration. The detailed numerical statistics for all the 15 districts of the BTAP project in our case study are summarized in Table 8. Nowshehra is the only district to show a negative change in forest cover from 2014 to 2020. The rest of the districts show an increase in the forestation. The ECP% column in Table 8 is the percentage change in forest cover compared to the forest cover in 2014. Hence some changes are very higher (+764% for Haripur and +327.13% for Malakand) because forest cover in 2014 was quite low and majority of the afforestation drive was carried out in these districts. The 14 to 20 LP% and 14 to 20 GP% columns show the forest loss and GP%, respectively, according to the metrics given in Sec. 3.2.4. In all districts except Nowshehra, the GP% is more than the LP%, which explains the overall forest gain seen in most districts.

Table 8 Forest statistics for the 15 BTAP districts produced by our AI-ForestWatch framework from 2014 to 2020. All districts except Nowshehra have shown a positive change in forest cover. It must be noted here that the ECP% compares the forest cover of 2020 relative to 2014 and not with respect to the total land area of the district.

Region	2014 FC%	2015 FC%	2016 FC%	2017 FC%	2018 FC%	2019 FC%	2020 FC%	14 to 20 LP%	14 to 20 GP%	14 to 20 ECP%
Nowshehra	2.94	11.31	3.63	7.74	5.39	5.71	2.80	1.17	1.02	-4.92
Kohat	0.72	15.0	1.04	1.81	2.48	0.83	1.0	0.27	0.55	+38.93
Karak	1.08	13.84	2.85	5.21	3.45	3.56	2.51	0.35	1.78	+131.56
Hangu	0.15	4.03	1.62	4.53	1.62	1.76	0.58	0.11	0.54	+295.09
Battagram	17.84	23.92	23.96	25.90	27.81	28.18	33.80	1.92	17.88	+89.48
Abbottabad	17.22	29.33	15.84	24.95	30.68	17.16	23.53	0.72	7.03	+36.64
Haripur	0.75	21.78	1.27	12.40	17.14	5.00	6.50	0.21	5.87	+764.83
Kohistan	21.23	29.42	24.12	31.56	30.97	32.29	31.82	2.03	12.62	+49.88
Tor Ghar	15.95	25.43	16.89	24.55	35.67	22.04	25.76	0.48	10.29	+61.49
Mansehra	16.46	20.53	18.87	21.50	24.38	19.14	23.98	1.44	8.97	+45.72
Buner	2.97	25.63	5.63	21.40	18.50	8.83	12.19	0.24	9.46	+310.96
Lower Dir	4.29	11.94	10.54	12.53	11.88	4.43	7.71	1.06	4.49	+79.95
Malakand	0.94	13.15	11.01	15.11	8.63	6.62	4.03	0.24	3.33	+327.13
Shangla	17.68	31.91	27.81	32.02	31.45	25.30	32.17	0.83	15.32	+81.92
Swat	16.73	18.01	19.04	22.0	20.84	21.51	24.42	1.67	9.37	+46.0

5 Discussions

In order to solve the problem of forest estimation and change detection, we compared various machine learning algorithms with a semantic segmentation-based approach. Additionally, we also considered multiple input channel configurations for detailed exploration of the algorithms. The deep learning-based semantic segmentation using UNet outperformed the other statistical machine learning algorithms quantitatively. The validation of the algorithms (statistical machine learning and deep learning) is based on the year 2015 images provided by the forest department of the KP province, Pakistan. These images are a result of a field survey conducted by the government to estimate the effects of BTAP project on the forest areas in year 2015. After detailed exploration, a UNet model with augmented configuration was selected for the AI-ForestWatch framework and trained and tested on year 2015 data. The trained model was inferred on the Landsat-8 images of years 2014 and 2016 to 2020 to generate the detailed forest statistics.

Since we detect the change trend for the temporal period from 2014 to 2020, we generated the forest cover maps for all these years. Difference maps of consecutive years yield the pixel-wise change maps for years 2014 to 2015, 2015 to 2016, 2016 to 2017, 2017 to 2018, 2018 to 2019, and 2019 to 2020. These change maps indicate the pixels, in which forest changed to non-forest, vice-versa, and the pixels with no change. Such discrete series of images can help in analysis of how the forest cover is changing over the temporal period and it can be extrapolated to see how the forest maps would look in the coming years. Since the maps show pixel-wise change, precise locations of forest loss/gain can be determined at upto 30 m per pixel. This information is very crucial for the foresters as it can be used to determine the reasons for loss and devise policies for forest protection and afforestation.

One artifact seen in our estimations is that the FC% jumps from years 2014 to 2015, came down closer to year 2014 percentage cover again in year 2016, and then increased normally from

years 2016 to 2020. This artifact is clearly visible in Table 8, districts Karak, Haripur, and Buner. In our case, we are constrained by the fact that only the local land cover data for year 2015 are available. Such artifacts can be resolved by cross-temporal training on labeled Landsat-8 data of multiple years, i.e., 2014, and 2016 to 2020. Since for training the model, we are relying on the data from year 2015 only (because that is the only available data) and generating forest cover estimations and maps for rest of the years, then any kind of noise, such as seasonality effects or cloud cover taken from year 2015 data could have generated this anomaly. We surmise that this would normalize or die out as more training data becomes available since in that case there would be a greater chance of receiving images of the same district from different years that have different seasonality effects in them even when they were cleaned by cross-temporal median filtering within 1 year of data. The classification model would then learn these seasonality effects alongside the classification task and would be more robust.

It should be noted that such detailed forest cover change maps and statistics (as shown in Appendix A) are not available for the BTAP project and this is considered a first attempt to assess the effects of afforestation on the forest cover area of the BTAP project at a reasonably high spatial resolution (see images of selected regions in Figs. 6 and 7). Nazir et al.⁴⁸ presented a mathematical model which suggests that the BTAP project will increase the forest area of the KP province to 23.59 million hectares, which is a 3.29% improvement over the forest area. Although this mathematical model considers the complete area of the KP province, our study of 15 districts shows a similar trend (presented in Table 8). Along with improvement in the forest area, the BTAP project has a positive socioeconomic impact on the population of KP. Khan et al.⁴⁹ showed that social sustainability increased by 69% between years 2014 and 2018. Additionally, the economic condition of the rural households in the region has improved considerably. Although the BTAP project has numerous positive environmental and social impacts, it lacks human resources and funding⁵⁰ to continuously monitor the forest cover and change. In order to tackle these problems, the change maps generated by the AI-ForestWatch framework provide the foresters with an initial estimations of the BTAP project and its impacts on land cover and all the associated ecosystems while relying on our digitized maps accuracy and the assumption that our cross-temporal forest cover predictions are accurate. Furthermore, it will provide policy makers with not only essential information for long-term planning but also to devise afforestation strategies.

Although, global forest change cover maps are also presented by Shimada et al.⁵ and Hansen et al.¹⁷ until 2012 and 2018, respectively. However, these lack the local knowledge and context, i.e., forest tree species and local forest rules. Globally different countries and locations have different forest species/qualities with respect their climate and region, for instance, forests in Asia will have different specifications as compared to Europe and North America. Therefore, the approach presented herein takes into account the local knowledge using the local training and annotated data to generate detailed maps of the specific region as shown in Appendix A. As we aim to open source the AI-ForestWatch framework, the release will include the dataset (training, validation, and test), the methodology, and the pretrained models. In order to adopt our framework for any other location or region, the pretrained models can be easily adapted by using the transfer learning approach to fine-tune the model for that specific study area or location.

The investigations carried out in this study will lead to new opportunities for forest monitoring specifically for the countries with limited resources. This would surely help foresters to have initial estimations of the available forest inventory and associated ecosystems. Furthermore, multi-label land cover maps can allow the same approach to be provided with, for example, monthly multi-spectral image profiles from Sentinel-2 or Landsat-8 data to monitor the change in water bodies, glaciers, or yearly urban growth.

6 Conclusion

Forest estimation and change detection are essential parameters for maintenance of the forest inventory, forest ecosystems, and carbon emissions. In order to perform these estimations frequently, in this paper, we presented AI-ForestWatch, an end to end framework for forest

estimation and change detection using the open source tools and imagery. Our approach is based on semantic segmentation (UNet), which performs pixel-wise classification to detect the forest and non-forest areas in an extended (eleven Landsat-8 bands concatenated with seven vegetation indices) Landsat-8 imagery. We generate pixel-wise forest and non-forest change maps using multi-temporal change detection from the years 2014 to 2020. As a case study, we analyzed the proposed framework on the 15 BTAP project districts in the KP province of Pakistan. Our evaluations show a positive change with forest cover improvements in the 14 BTAP districts of the study area. In the future, we aim to explore newer CNN topologies such as SegNet and Mask RCNN for more accurate forest cover map predictions. Furthermore, more accurate comparisons will be made when yearly ground truth data for all the years are available. That would allow us to train the classification network cross-temporally for better map predictions.

7 Appendix A: Forest Cover Maps and Change Maps of Selected Regions for the BTAP Project Generated by Our AI-ForestWatch Framework

Although, detailed forest cover and change maps of all 15 districts under study were generated by the AI-ForestWatch framework. In Figs. 6 and 7 we show only forest cover and change maps of districts Abbottabad and Battagram respectively.

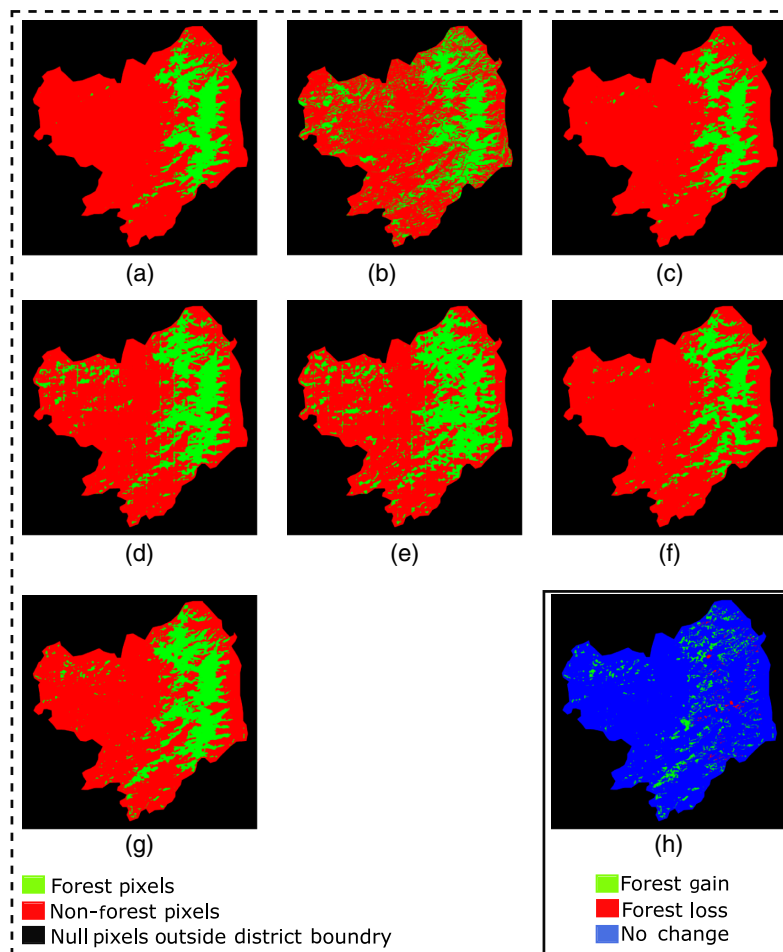


Fig. 6 Abbottabad district 2014 to 2020 statistics generated automatically by the AI-ForestWatch framework: (a)–(g) the forest cover maps generated by the UNet inference from years 2014 to 2020, respectively, and (h) the forest change map of 2020 with respect to 2014. Blue pixels indicate no change in the area, red pixels indicate forest loss, and green pixels indicate forest gain.

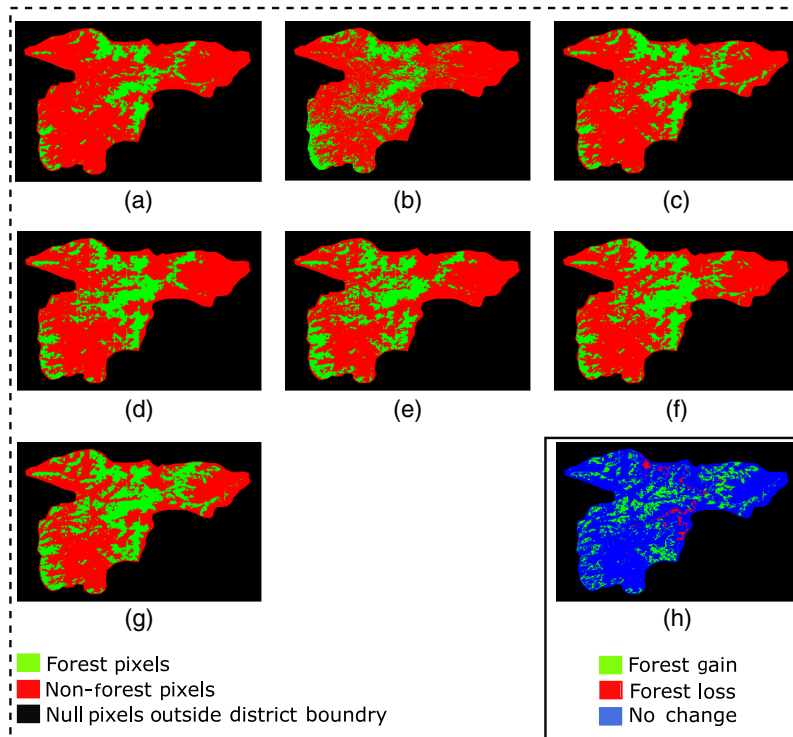


Fig. 7 Battagram district 2014 to 2020 statistics generated automatically by the AI-ForestWatch framework: (a)–(g) the forest cover maps generated by the UNet inference from years 2014 to 2020, respectively, and (h) the forest change map of 2020 with respect to 2014. Blue pixels indicate no change in the area, red pixels indicate forest loss, and green pixels indicate forest gain.

Acknowledgments

The authors would like to thank Forest Department of Khyber Pakhtunkhwa for providing the Billion Tree Afforestation Project data. The authors would also like to thank Tasleem Khan for providing with useful insights into land cover classification problem and its application and difficulties working with KP forest cover data. Furthermore, we would like to thank the anonymous reviewers for their valuable feedback. Part of this research was supported by the German Academic Exchange Service (DAAD) under Grant No. 57402923. The authors have no conflicts of interest to declare.

References

1. UN Sustainable Development Goals, Sustainably manage forests, combat desertification, halt and reverse land degradation, halt biodiversity loss, <https://www.un.org/sustainable-development/biodiversity/> (accessed December 2020).
2. L. Vesa, *Forest Inventory Data Management and Processing With Open Foris Tools*, Open Foris (2019).
3. D.-H. Kim et al., “Global, landsat-based forest-cover change from 1990 to 2000,” *Remote Sens. Environ.* **155**, 178–193 (2014).
4. M. C. Hansen et al., “Mapping tree height distributions in Sub-Saharan Africa using landsat 7 and 8 data,” *Remote Sens. Environ.* **185**, 221–232 (2016).
5. M. Shimada et al., “New global forest/non-forest maps from ALOS PALSAR data (2007–2010),” *Remote Sens. Environ.* **155**, 13–31 (2014).
6. C. Souza, Jr. et al., “Ten-year landsat classification of deforestation and forest degradation in the Brazilian Amazon,” *Remote Sens.* **5**(11), 5493–5513 (2013).
7. P. V. Potapov et al., “Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with landsat ETM+ data,” *Remote Sens. Environ.* **122**, 106–116 (2012).

8. X. Ni et al., "Forest height estimation and change monitoring based on artificial neural network using geoscience laser altimeter system and landsat data," *J. Appl. Remote Sens.* **14**(2), 022207 (2019).
9. M. Hussain, "Change detection from remotely sensed images: from pixel-based to object-based approaches," *ISPRS J. Photogramm. Remote Sens.* **80**, 91–106 (2013).
10. S. H. Khan et al., "Forest change detection in incomplete satellite images with deep neural networks," *IEEE Trans. Geosci. Remote Sens.* **55**(9), 5407–5423 (2017).
11. J. Liu, X. Wang, and T. Wang, "Classification of tree species and stock volume estimation in ground forest images using deep learning," *Comput. Electron. Agric.* **166**, 105012 (2019).
12. G. Xian and C. Homer, "Updating the 2001 National Land Cover Database Impervious Surface products to 2006 using landsat imagery change detection methods," *Remote Sens. Environ.* **114**(8), 1676–1686 (2010).
13. E. A. Lehmann et al., "Forest cover trends from time series landsat data for the Australian continent," *Int. J. Appl. Earth Obs. Geoinf.* **21**, 453–462 (2013).
14. C. Huang et al., "An automated approach for reconstructing recent forest disturbance history using dense landsat time series stacks," *Remote Sens. Environ.* **114**(1), 183–198 (2010).
15. O. A. Carvalho, Jr. et al., "A new approach to change vector analysis using distance and similarity measures," *Remote Sens.* **3**(11), 2473–2493 (2011).
16. T. Blaschke, "Object based image analysis for remote sensing," *ISPRS J. Photogramm. Remote Sens.* **65**(1), 2–16 (2010).
17. M. C. Hansen et al., "High-resolution global maps of 21st-century forest cover change," *Science* **342**(6160), 850–853 (2013).
18. D.-X. Song et al., "Use of landsat and corona data for mapping forest cover change from the mid-1960s to 2000s: case studies from the eastern United States and Central Brazil," *ISPRS J. Photogramm. Remote Sens.* **103**, 81–92 (2015).
19. X. Zhang et al., "Rapid generation of global forest cover map using landsat based on the forest ecological zones," *J. Appl. Remote Sens.* **14**, 022211 (2020).
20. J. Im and J. R. Jensen, "A change detection model based on neighborhood correlation image analysis and decision tree classification," *Remote Sens. Environ.* **99**(3), 326–340 (2005).
21. Q. Zhao et al., "Comparison of machine learning algorithms for forest parameter estimations and application for forest quality assessments," *For. Ecol. Manage.* **434**, 224–234 (2019).
22. F. Qamer et al., "Mapping deforestation and forest degradation patterns in Western Himalaya, Pakistan," *Remote Sens.* **8**(5), 385 (2016).
23. C. Voight et al., "Predictive modeling of future forest cover change patterns in Southern Belize," *Remote Sens.* **11**(7), 823 (2019).
24. P. Helber et al., "Introducing EuroSAT: a novel dataset and deep learning benchmark for land use and land cover classification," in *IGARSS 2018-2018 IEEE Int. Geosci. and Remote Sens. Symp.*, IEEE, pp. 204–207 (2018).
25. H. Song, Y. Kim, and Y. Kim, "A patch-based light convolutional neural network for land-cover mapping using landsat-8 images," *Remote Sens.* **11**(2), 114 (2019).
26. K. He et al., "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 770–778 (2016).
27. C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 1–9 (2015).
28. P. P. de Bem et al., "Change detection of deforestation in the Brazilian Amazon using landsat data and convolutional neural networks," *Remote Sens.* **12**(6), 901 (2020).
29. The Bonn Challenge, "The challenge. A global effort," <http://www.bonnchallenge.org/content/challenge> (accessed December 2020).
30. NASA, "Landsat timeline," <https://landsat.gsfc.nasa.gov/a-landsat-timeline/> (accessed December 2020).
31. A. Kamal, M. Yingjie, and A. Ali, "Significance of billion tree tsunami afforestation project and legal developments in forest sector of Pakistan," **1**, 157–165 (2019).
32. Google, "Google Earth engine," <https://earthengine.google.com/> (accessed December 2020).
33. J. Xue and B. Su, "Significant remote sensing vegetation indices: a review of developments and applications," *J. Sens.* **2017** (2017).

34. Sentinel-Hub, “NDVI (normalized difference vegetation index),” <https://www.sentinel-hub.com/eoproducts/ndvi-normalized-difference-vegetation-index> (accessed December 2020).
35. Sentinel-Hub, “EVI (enhanced vegetation index),” <https://www.sentinel-hub.com/eoproducts/evi-enhanced-vegetation-index-0> (accessed December 2020).
36. The Landscape Toolbox, “Soil-adjusted vegetation index,” https://wiki.landscapetoolbox.org/doku.php/remote_sensing_methods:soil-adjusted_vegetation_index (accessed December 2020).
37. USGS, “Landsat surface reflectance-derived spectral indices—Landsat modified soil adjusted vegetation index,” <https://www.usgs.gov/land-resources/nli/landsat/landsat-modified-soil-adjusted-vegetation-index> (accessed December 2020).
38. Z. Jiang et al., “Interpretation of the modified soil-adjusted vegetation index isolines in red-NIR reflectance space,” *J. Appl. Remote Sens.* **1**(1), 013503 (2007).
39. S. Barati et al., “Comparison the accuracies of different spectral indices for estimation of vegetation cover fraction in sparse vegetated areas,” *Egypt. J. Remote Sens. Space Sci.* **14**(1), 49–56 (2011).
40. E. Olmos-Trujillo et al., “Spatio-temporal response of vegetation indices to rainfall and temperature in a semiarid region,” *Sustainability* **12**(5), 1939 (2020).
41. USGS, “Landsat surface reflectance-derived spectral indices—normalized difference moisture index,” <https://www.usgs.gov/land-resources/nli/landsat/normalized-difference-moisture-index> (accessed December 2020).
42. USGS, “Landsat surface reflectance-derived spectral indices—landsat normalized burn ratio,” <https://www.usgs.gov/land-resources/nli/landsat/landsat-normalized-burn-ratio> (accessed December 2020).
43. USGS, “Landsat surface reflectance-derived spectral indices—landsat normalized burn ratio 2,” <https://www.usgs.gov/land-resources/nli/landsat/landsat-normalized-burn-ratio-2> (accessed December 2020).
44. O. Ronneberger, P. Fischer, and T. Brox, “U-net: convolutional networks for biomedical image segmentation,” *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
45. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *3rd Int. Conf. Learn. Represent., ICLR 2015*, San Diego, California, Y. Bengio and Y. LeCun, Eds., (2015).
46. T.-Y. Lin et al., “Focal loss for dense object detection,” in *Proc. IEEE Int. Conf. Comput. Vision*, pp. 2980–2988 (2017).
47. C. H. Sudre et al., “Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations,” *Lect. Notes Comput. Sci.* **10553**, 240–248 (2017).
48. N. Nazir et al., “A system dynamics model for billion trees Tsunami afforestation project of Khyber Pakhtunkhwa in Pakistan: model application to afforestation activities,” *J. Mount. Sci.* **16**, 2640–2653 (2019).
49. N. Khan et al., “Socioeconomic impacts of the billion trees afforestation program in Khyber Pakhtunkhwa province (KPK), Pakistan,” *Forests* **10**(8), 703 (2019).
50. M. I. Khan et al., “Third party monitoring of billion trees afforestation project in Khyber Pakhtunkhwa phase-II,” WorldWide Fund for Nature Pakistan (WWF-Pakistan) (2017).

Annus Zulfiqar: Biography is not available.

Muhammad M. Ghaffar received his BSc degree in electrical engineering from the Federal Urdu University, Islamabad, Pakistan, and his MSc degree in electrical and computer Engineering from the TU Kaiserslautern, Kaiserslautern, Germany, in 2009 and 2013, respectively. He is currently pursuing his PhD in the Microelectronic System Design Research Group, Department of Electrical and Computer Engineering, University of Kaiserslautern, Kaiserslautern, Germany. His research interests include machine learning and deep learning with a special emphasis on energy-efficient implementations.

Muhammad Shahzad received his PhD in radar remote sensing and image analysis from the Department of Signal Processing in Earth Observation, Technische Universität München, Munich, Germany, in 2016. His PhD topic was the automatic 3-D reconstruction of objects from point clouds retrieved from spaceborne synthetic-aperture-radar image stacks. Since October

2016, he has been working as an assistant professor at the School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, Pakistan. His research interests include processing both unstructured/structured 3D point clouds, optical RGBD images, and very high-resolution radar data.

Christian Weis received his PhD in electrical engineering from the TU Kaiserslautern, Kaiserslautern, Germany, in 2014. From 1998 to 2009, he was with Siemens Semiconductor, Infineon Technologies AG, and Qimonda AG, Munich, Germany, in DRAM design. During this time frame, he was involved in DRAM design for graphics and commodity DRAM products. Since 2009, he has been with the Microelectronic System Design Research Group, TU Kaiserslautern, Kaiserslautern, Germany. He holds more than 20 patents related to DRAMs and DRAM design and published more than 60 papers. His current research interests include DRAM controller design, near- and in-memory processing, 3D-integrated DRAMs, heterogeneous memory architectures, physical design, and MPSoCs.

Muhammad I. Malik received his masters' degree in computer science and his PhD in artificial intelligence in 2011 and 2015, respectively, from Technische Universität Kaiserslautern, Kaiserslautern, Germany. Currently, he is an assistant professor at the NUST, Islamabad, Pakistan. In addition, he worked as a visiting researcher for 3 months at Kyushu University, Fukuoka, Japan, and University of Western Australia, Perth, Australia, in 2013 and 2014, respectively. His research interests include machine learning and computer vision with a special emphasis on applications in document image analysis and recognition.

Faisal Shafait received his PhD in computer engineering from the Technical University Kaiserslautern, Kaiserslautern, Germany, in 2008. He is currently working as a professor at the School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, Pakistan. In addition, he is an adjunct professor at the School of Computer Science and Software Engineering, The University of Western Australia, Perth, Australia. His research interests include computer vision and machine learning, in which he has published over 150 peer-reviewed international conference and journal publications. He is the recipient of IAPR/ICDAR Young Investigator Award 2019 by the International Association of Pattern Recognition which is awarded every 2 years to the most outstanding young scientist (below 40 years of age) worldwide in the field of pattern recognition and document analysis.

Norbert Wehn received his diploma and PhD degrees from the TU Darmstadt, Darmstadt, Germany. He currently holds the chair for microelectronic systems design in the Department of Electrical and Computer Engineering, TU Kaiserslautern, Kaiserslautern, Germany. He has more than 350 publications in various fields of microelectronic system design and holds 20 patents. His special research interests are VLSI architecture for mobile communication, forward error correction techniques, low-power techniques, advanced SoC and memory architectures, 3-D integration, reliability issues in SoC, the IoT, and hardware accelerators for big data applications. He is an associate editor of various journals and a member of several scientific industrial advisory boards.