Mapping Irrigated Croplands in Africa Using Combined Sentinel-1 and Sentinel-2 Data

Nasser Ahmed\textsuperscript{1} and Shuanggen Jin\textsuperscript{1}

\textsuperscript{1}Affiliation not available

April 05, 2024
Mapping Irrigated Croplands in Africa Using Combined Sentinel-1 and Sentinel-2 Data

Nasser A. Mohamed, Shuanggen Jin

Abstract—In order to effectively manage agriculture and promote sustainable land use in Africa, accurate mapping of irrigated croplands is imperative. This paper pioneers a novel Hybrid Vegetation Index (HVI) and integrated an approach harnessing the synergies between Radar imagery from Sentinel-1 and optical data from Sentinel-2, key vegetation indices (HVI, MSAVI), and Principal Component Analysis (PCA) to significantly enhance mapping accuracy across diverse African landscapes. Employing Convolutional Neural Network (CNN), Random Forest (RF), and Support Vector Machine (SVM) classifiers, the combined potential of these satellite datasets is thoroughly Assessed. The integrated utilization of both Sentinel-1 and Sentinel-2 demonstrates a substantial enhancement in overall accuracy, elevating classification results by approximately 5-6% when compared to individual sensor applications. Specifically, in Kenya, Egypt, and Nigeria, the amalgamation of Sentinel-1 and Sentinel-2 enables CNN to achieve remarkable accuracies of 97.5%, 98.01%, and 98%, respectively. These findings underscore the superior performance of the fused Sentinel data and highlighted the pivotal role of their integration in advancing precise agricultural mapping. This research emphasizes the promising impact of HVI and Sentinel satellite missions in empowering informed decision-making and promoting sustainable land management practices across vital agricultural regions in Africa.

Keywords—croplands; CNN; crop mapping; Sentinel-2; Sentinel-1; vegetation index

I. INTRODUCTION

Precise and prompt data regarding the spatial allocation of irrigated agricultural areas is fundamental for studies related to agricultural practices, hydrological resources, land utilization patterns, and climate variations in the domain of remote sensing applications[1]. Agriculture serves as the cornerstone for human sustenance, offering vital resources such as fibers, food, raw materials, and fuel essential for human existence and well-being[2]. The comprehensive mapping and continuous monitoring of agricultural land use hold significant importance. It aids in predicting crop yields, evaluating factors impacting crop stress, and estimating the impact of natural hazards on agricultural productivity. Additionally, it plays a crucial role in determining irrigation water requirements at the farm level, contributing to improved management of water resources[3].

Vegetation Indices (VIs) in remote sensing serve as straightforward, efficient, and commonly employed techniques for both quantitative and qualitative assessments of vegetation cover, vitality, and growth dynamics[4]. VIs offer valuable insights for diverse precision agriculture applications, furnishing quantitative data on the growth and health of crops[5]. Greater variations in vegetation reflectance are observed in the Red-Green-Blue (RGB) and Near-Infrared (NIR) bands. Due to this characteristic, many Vegetation Indices (VIs) are formulated using combinations of RGB bands along with infrared, including red-edge or near-infrared[6]. Commonly employed Vegetation Indices (VIs) that exhibit sensitivity to biomass and vegetation density include the Normalized Difference Vegetation Index NDVI (Caruso et al., 2017), Soil-Adjusted Vegetation Index SAVI[6], Enhanced Vegetation Index EVI[7], Modified Soil-Adjusted Vegetation Index MSAVI[8]. To assess water content in vine leaves accurately, the optimal combination involves utilizing Near-Infrared (NIR) and Shortwave Infrared (SWIR), as it proves to be more effective in predicting and showcasing the water stress levels within the vineyard.

The advent of the big data of remote sensing has resulted in an abundance of comprehensive global observational datasets, providing a unique avenue for exploring the spatial arrangement of irrigated croplands[9]. Over the Previous decades, numerous global land-cover datasets have been generated across diverse spatial resolutions, spanning from 1 km to 300 m[10]. Several products for mapping land cover utilizing satellite data have adeptly integrated irrigation as a distinct category within their classifications. These published
products encompass datasets like the global map of irrigation areas (GMIA)\cite{11}, \cite{12}. The majority of the satellite-derived products maintain a spatial resolution of approximately 250 meters or coarser\cite{13}. In areas like Africa, where cropping patterns are fragmented, employing remote sensing data with enhanced spatial resolution is anticipated to significantly enhance the precision of identifying irrigated croplands\cite{14}.

Recently, because of the availability of freely accessible remote sensing data with high resolution like Sentinel-1 (S1) and Sentinel-2 (S2), along with the rapid expansion of data storage and computational capacities, there has been successful development of land-cover products on a global scale with highly detailed spatial resolutions (10 meters) \cite{15}. The increased spatial resolution enables finer discrimination and outlining of individual fields or smaller agricultural plots, thereby enhancing the accuracy in mapping irrigated areas\cite{16}. The Sentinel-1 mission operates via a dual-satellite system, consisting of Sentinel-1A and Sentinel-1B, both hosting a dual-polarized C-band synthetic aperture radar (SAR) with a spatial resolution of 20 meters. These satellites revisit the same area every six days and offer various operational modes for data acquisition. This capability allows users access to an extensive and openly available archive of data over the long term, facilitating applications that demand extended time series analyses \cite{17}, \cite{18}. The radar system on Sentinel-1, which has the capability to penetrate through clouds, offers uninterrupted observations across various weather conditions, day or night, making it a valuable asset to observe and monitor changes occurring on the Earth's surface. Particularly beneficial in regions with persistent cloud cover, this capability ensures continuous data availability. Its applications span diverse domains such as observation and tracking of disturbances within forest environments \cite{19}, mapping of rice croplands \cite{20}, tracking ground deformation \cite{21}, and monitoring inland water areas \cite{22}.

The SAR (Synthetic Aperture Radar) method has been shown to be highly effective for seasonal agricultural monitoring purposes \cite{23}. The SAR system demonstrates sensitivity towards various vegetation biophysical variables, the dynamic traits of plant targets, and foundational soil parameters, including the moisture held within plant tissues, geometric properties, deflection, Irregularity, roughness of the soil surface, and moisture levels (Soria-Ruiz, Fernandez-Ordonez, and Bugden-Storie 2007). The Sentinel-2 spacecraft, which form the second series of satellites among the array of ESA's Sentinel missions, are utilizing scanners that capture information across multiple spectral bands \cite{25}. The core purpose of the Sentinel-2 satellites is to furnish satellite data characterized by high levels of detail and resolution crucial for land cover and land use monitoring, Observation of climate change, monitoring of disasters, and to complement existing satellite missions like Landsat.\cite{26}.

The process of Creating maps that specifically delineate or depict irrigated areas using remote sensing images entails the classification of land cover, wherein each pixel in the images is assigned a predefined land cover class. Numerous studies have relied on shallow classifiers like maximum likelihood classification (MLC) for this purpose \cite{27}\cite{28}, random forest (RF)\cite{29}, \cite{30}, and multi-layer perceptron (MLP) \cite{31}, support vector machine (SVM) \cite{32}, for land cover classification. \cite{33} conducted a comprehensive mapping study that focused on differentiating irrigated and non-irrigated corn fields in Nebraska. They further employed images sourced from Landsat data to enhance their mapping methodology. \cite{34} performed the mapping of irrigated areas utilizing a combination of Landsat and Sentinel-2 imagery alongside the platform Google Earth Engine. \cite{35} introduced an approach to map irrigation based on soil water content, utilizing the optical trapezoid method derived from a range of Landsat datasets. \cite{36} demonstrated the utilization of high-resolution optical imagery and SAR time series, amalgamating Landsat 8 and Sentinel 1 data to enhance the early detection of crop types. They recommended integrating Sentinel 2 data for the preliminary identification of crops. In their evaluation involving the fusion of Both Optical data, alongside multipolar radar data were employed for the purpose of mapping land in Brazil, (De Oliveira Pereira et al. 2013) observed that radar data significantly enhance user accuracy. Specifically, they found that the HH polarization (horizontal transmission and reception) provides greater differentiation among various land use classes compared to HV polarization (horizontal transmission and vertical reception). However, the merging or combination of radar and optical data yielded the most favorable statistical outcomes for land mapping. (Zhou et al. 2017b) explored the potential of winter wheat mapping by employing SAR data, optical images, and the fusion of both data sets. They conducted classification mapping by combining Sentinel-1 data with optical images using the random forest method.

Optical data from Sentinel-2 is sensitive to changes in vegetation color and structure, while SAR data from Sentinel-1 can penetrate clouds and provide information about vegetation water content and structure \cite{37}. Creating hybrid vegetation indices by integrating bands from both sensors can enhance the ability to monitor vegetation under various environmental conditions. This proposed hybrid vegetation index, which includes bands from both Sentinel-2 and Sentinel-1, reflects a comprehensive approach to vegetation monitoring, taking advantage of the synergies between optical and radar data. This approach is particularly beneficial in regions where cloud cover is frequent or in areas with dense vegetation where optical data alone may be limited \cite{38}.

This study aims to explore and establish a novel Hybrid Vegetation Index (HVI) and a comprehensive method for mapping irrigated croplands in African environments using the
time series data from Sentinel-1 and Sentinel-2 satellites and comparing the results with the traditional classification methods. Section 2 shows the materials and methods, results and discussion are outlined in Section 3, and ultimately, the conclusions are provided in Section 4.

II. MATERIALS AND METHODS

A. Study Area

This study encompassed three distinct regions in Africa as shown in fig.1, each characterized by unique agricultural landscapes and diverse environmental features. The first study area, Kenya, spans from approximately 5°N to 4°S latitude and 34°E to 42°E longitude, encompassing various agro-ecological zones. It includes arid and semi-arid lands in the north and fertile highlands in the central and western parts (e.g., regions near 0.5°S, 36°E). The second study area, Egypt, predominantly covers the Nile River Valley and Delta, extending approximately from 31°N to 24°N latitude and 25°E to 35°E longitude. These coordinates encapsulate the regions where intensive agricultural practices supported by Nile River irrigation occur, notably around 30°N, 31°E, and 30°N, 32°E. Lastly, Nigeria, spanning from about 4°N to 14°N latitude and 3°E to 15°E longitude. Nigeria's varied landscape offers a rich tapestry for examining irrigated croplands in distinct environmental contexts. These study areas were meticulously chosen due to their agricultural significance, diverse land cover, and the prevalence of irrigated croplands, providing ideal settings for evaluating the efficacy of satellite data in mapping these critical agricultural regions.

B. Data

Sentinel-2 level 2A (L2A) optical images and Sentinel-1 Ground Range Detected (GRD) scenes were acquired through the website Copernicus Data Space Ecosystem | Europe's eyes on Earth (accessed on 25 Oct, 2023). The images from the ground range detected used within this framework are obtained through interferometric wide swath (IW) mode. The Sentinel-1 GRD images maintain a 10-meter pixel spacing, ensuring complete information detail, as stated by [18] regarding their spatial resolution, a time series of Sentinel-1 images (6 images) from 11 Apr, 2022 to 20 Oct, 2022 were used in this study (Table I). The satellites of Sentinel-2 are outfitted with Multispectral Imaging Instruments (MSI) (Table II) that possess the ability to capture 13 bands across wide swaths as shown in fig.2[25].

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>Product Type</th>
<th>Acquisition Orbit</th>
<th>Mode</th>
<th>Polarization</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.04.2022</td>
<td>15:56:43</td>
<td>IW GRDH JS</td>
<td>ASCENDING</td>
<td>IW</td>
<td>VV + VH</td>
</tr>
<tr>
<td>13.05.2022</td>
<td>15:57:09</td>
<td>IW GRDH JS</td>
<td>DESCENDING</td>
<td>IW</td>
<td>VV + VH</td>
</tr>
<tr>
<td>16.07.2022</td>
<td>15:57:14</td>
<td>IW GRDH JS</td>
<td>ASCENDING</td>
<td>IW</td>
<td>VV + VH</td>
</tr>
<tr>
<td>21.08.2022</td>
<td>15:57:16</td>
<td>IW GRDH JS</td>
<td>DESCENDING</td>
<td>IW</td>
<td>VV + VH</td>
</tr>
<tr>
<td>14.09.2022</td>
<td>15:57:17</td>
<td>IW GRDH JS</td>
<td>ASCENDING</td>
<td>IW</td>
<td>VV + VH</td>
</tr>
<tr>
<td>20.10.2022</td>
<td>15:57:17</td>
<td>IW GRDH JS</td>
<td>ASCENDING</td>
<td>IW</td>
<td>VV + VH</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location</th>
<th>Date</th>
<th>Time</th>
<th>Product Type</th>
<th>Acquisition Orbit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenya</td>
<td>13.05.2022</td>
<td>07:54:44</td>
<td>MSI L2A</td>
<td>ASCENDING</td>
</tr>
<tr>
<td>Egypt</td>
<td>20.10.2022</td>
<td>08:32:55</td>
<td>MSI L2A</td>
<td>ASCENDING</td>
</tr>
<tr>
<td>Nigeria</td>
<td>14.09.2022</td>
<td>07:06:46</td>
<td>MSI L2A</td>
<td>ASCENDING</td>
</tr>
</tbody>
</table>

Fig 2: Sentinel-2 satellite images (a) false color of Kenya, (b) false color of Egypt and (c) false color of Nigeria

C. Methods

The methodology for this research as shown in fig.3 entailed the systematic collection and processing of Sentinel-1 and Sentinel-2 time series imagery. Initially, Sentinel-1 data underwent preprocessing utilizing the SNAP, also known as the Sentinel Application Platform, open software [39].
The preprocessing involved specific steps to extract sigma-naught (σ°) values from both Vertical Transmit/Vertical Receive (VV) and Vertical Transmit/Horizontal Receive (VH) polarized images through comprehensive backscatter analysis techniques [40]. Subsequently, the obtained sigma-naught values were utilized to generate irrigated maps employing two distinct classification approaches: Random Forest (RF) [30], [34] and Convolutional Neural Network (CNN) models. The RF classification technique was applied independently, while the CNN model, tailored for this research, employed sophisticated deep learning methodologies to classify irrigated croplands based on Sentinel-1 data.

A novel aspect of this study involved the novel hybrid vegetation index and the integration of Sentinel-1 and Sentinel-2 imagery. The combined dataset was employed to develop a novel HVI and an innovative CNN model specifically designed to map irrigated areas by harnessing the synergistic information from both sensors. This novel CNN model was rigorously tested and compared against the results obtained from RF and SVM classifications, ensuring a comprehensive evaluation of the mapping accuracy across the studied areas in Kenya, Egypt, and Nigeria. The methodology aimed to leverage the complementary nature of Sentinel-1 and Sentinel-2 data, culminating in the development and assessment of a novel CNN model tailored for accurate mapping of irrigated croplands. The comparisons with traditional classification methods (RF and SVM) provided critical insights into the efficacy and superiority of the proposed approach in delineating irrigated areas across diverse landscapes within the studied African regions.

1) Remote sensing data processing
a) Sentinel-1 time series processing

Applications of SAR data measurement require radiometric calibration in order to precisely determine the relationship between apparent pixel values of SAR image data and the corresponding geophysical information on the earth's surface [41]. Since speckle naturally affects all SAR images, speckle filtering is a crucial step in the preprocessing of SAR images [40]. The local measurement in radar image processing, the Lee filter is among the most widely used and well-known despeckling methods. [42].

Essential beginning products for various applications ranging from ecological to hazard monitoring are Radiometric Terrain Correction (RTC) data. To generate RTC products from Ground Range Detected (GRD) level SAR data, a number of open-source software packages are available, such as the European Space Agency's (SNAP) Sentinel-1 Toolbox. RTC includes normalizing the measured backscatter with respect to terrain slope and eliminating radiometric distortions that depend on geometry [39].

Radar backscatter measurements are expressed as raw digital data on the linear scale. Effective data visualization or analysis may be challenging due to the wide dynamic range of these values, particularly if there are notable variations in backscatter levels throughout the image. The dynamic range of data is compressed by converting the linear scale to a logarithmic scale expressed in decibels (dB). This conversion improves the contrast between various image elements, making it simpler to see and understand the data. The dB scale makes fluctuations in backscatter intensity easier to see, which helps identify minute variations in surface properties like surface roughness or different types of land cover [40].

b) Backscatter Analysis

The part of radar radiation that is reflected back toward the radar detector following interaction with targets or the Earth's surface is known as backscatter. It is dependent upon a number of surface characteristics, including material composition, geometric structure, amount of water, and roughness. Rougher surfaces or items with intricate features typically produce higher backscatter, whereas smooth surfaces typically result in reduced backscatter [41]. The quantity of radar signal that the sensor receives in relation to the signal that is transmitted is measured by the backscatter coefficient. It measures the target's radar cross-section and is affected by polarization, incidence angle, and surface roughness [42]. The normalized radar cross-section is expressed by sigma-naught (σ°), which is obtained from the backscatter coefficient. The backscatter results are normalized by this standardized measurement, which reduces their sensitivity to changes in imaging geometry such incidence angle [39], [42].

2) Novel Hybrid Vegetation Index (HVI)

By combining data from Sentinel-2 and Sentinel-1, we can leverage the strengths of both sensors to overcome limitations associated with using a single data source. Optical data from Sentinel-2 is sensitive to changes in vegetation color and structure, while SAR data from Sentinel-1 can penetrate clouds...
and provide information about vegetation water content and structure [37]. The signal penetrates and interacts with the vegetation canopy components in a magnitude that depends on the SAR wavelength [43]. Creating vegetation indices by integrating bands from both sensors can enhance the ability to monitor vegetation under various environmental conditions. This proposed vegetation index, which includes bands from both Sentinel-2 and Sentinel-1, reflects a comprehensive approach to vegetation monitoring, taking advantage of the synergies between optical and radar data. This approach is particularly beneficial in regions where cloud cover is frequent or in areas with dense vegetation where optical data alone may be limited [38].

By using the optical bands from sentinel 2 images and the dual-pol (VV-VH) Sentinel-1 SAR data, a Novel Hybrid Vegetation Index HVI was introduced by the following equation.

$$HVI = \frac{(\text{NIR} + 0.5 \times \text{VV}) - (\text{Red} + 0.5 \times \text{VH}) - (\text{Blue} + 0.5 \times \text{SWIR})}{(\text{NIR} + 2 \times \text{VV}) + (\text{Red} + 0.5 \times \text{VH}) - (\text{Red} + 0.5 \times \text{VH})}$$

NIR (Near-Infrared) is sensitive to the presence of chlorophyll in vegetation. High NIR values indicate healthy and vigorous vegetation. Red, Blue, SWIR (Shortwave Infrared) Sentinel-2 bands capture information across the visible and shortwave infrared spectrum. Red is sensitive to chlorophyll absorption, Blue represents overall reflectance, and SWIR provides insights into vegetation moisture content and structure [44]. VV (Vertical transmit, Vertical receive) and VH (Vertical transmit, Horizontal receive) represent the radar backscatter signal in the vertical polarization and horizontal polarization respectively. It is particularly sensitive to vegetation structure and water content [45]. The subtraction of (Red+0.5×VH) in the final term serves to suppress non-vegetation signals and enhance the sensitivity of the index to vegetation-specific features [46].

Through rigorous comparative analyses against NDVI, this novel index consistently demonstrates its prowess in delineating subtle variations in vegetation conditions, emphasizing its applicability across diverse ecosystems. The amalgamation of Sentinel-1 and Sentinel-2 data in this index not only broadens the scope of information but also positions it as a powerful tool for remote sensing applications, showcasing a paradigm shift in vegetation monitoring capabilities.

3) Sentinel-2 feature extraction

Specialized vegetation indices like HVI and MSAVI are frequently used in feature extraction processes from multispectral remote sensing data. In the realm of vegetation indices, HVI stands as a groundbreaking advancement in capturing the intricacies of vegetation dynamics, amalgamates the strengths of multispectral and radar data, showcasing a meticulous fusion of near-infrared (NIR), red, blue, and shortwave infrared (SWIR) bands from Sentinel-2, along with the vertical (VV) and horizontal (VH) polarization from Sentinel-1 [43]. By encompassing these diverse spectral and radar components, the index goes beyond the limitations of traditional vegetation indices like NDVI.

Furthermore, the Modified Soil-Adjusted Vegetation Index (MSAVI) modifies the NDVI formula to account for the effects of the soil background, providing enhanced sensitivity to changes in vegetation in regions with different soil conditions [47]. By including soil brightness into its calculation, MSAVI improves its applicability in a variety of settings and allows for more precise evaluations of the stress and vigor of the vegetation. Comprehensive assessments of the dynamics and conditions of the vegetation are made possible by these indices, which are useful instruments in ecosystem research, precision agriculture, and environmental monitoring[48].

4) Principal component analysis (PCA)

PCA represents a linear transformation, while an autoencoder is a non-linear neural network method designed for data dimension reduction through its decoder and encoder components. In the context of Sentinel-2 images, PCA has been utilized to enhance information derived from the optical sensor bands. Its output, distinct from the primary bands, aids in the differentiation between cropland and non-croplands [49].

5) Support vector machines (SVM)

The SVM, a widely applied supervised learning method in various remote sensing applications, works by identifying the optimal minimization, known as the decision boundary, within a problem space with ambiguous classifier outputs. The hyperplane, denoted as the decision boundary, effectively separates the classification problem into a predetermined set of classes consistent with the training data.[32], [50].

6) Random Forest (RF) classification

The RF classifier, functioning as an ensemble classifier, generates multiple decision trees by employing randomly selected subsets of training data and variables. Its widespread adoption within the remote sensing community is owed to the accuracy it achieves in classifications [51]. Notably, the RF classifier has been effectively applied in mapping various land cover classes, particularly in the identification of irrigated croplands [34] Supervised machine learning algorithms for image classification encompass methodologies like Random Forest. The management of the ever-expanding volumes of remotely sensed data and related products poses a challenge when traditional processing techniques are employed (Abdi 2020).
7) Convolutional Neural Network (CNN)

The CNN is an adaptable architecture with multiple layers composed of multiple feature-extraction stages. Each stage consists of three layers: 1) a convolutional layer, 2) a nonlinearity layer, and 3) a pooling layer [51], [52]. In this study, we present a comprehensive workflow of CNN designed to analyze Sentinel-1 and Sentinel-2 satellite imagery alongside Principal Component Analysis (PCA) and vegetation indices of Sentinel-2 data as shown in fig. 4. Employing the TensorFlow Keras API, a CNN architecture is constructed. The approach involves segmenting the images into small patches in size of one pixel to enable a more granular analysis.

The dataset was meticulously divided into three subsets for robust model evaluation. 70% of the data was allocated for training the CNN model, allowing it to learn patterns and features within the images. The subsequent 20% was set aside as testing data, enabling rigorous assessment of the model’s generalization capabilities on unseen examples. Finally, a 10% variation dataset was designated to gauge the model’s performance stability and consistency across different data instances. This architecture integrates convolutional layers for feature extraction, max-pooling layers for dimensionality reduction, flattening layers for data transition, and densely connected layers for classification. The inclusion of the softmax activation function enables multiclass classification tasks.

8) Evaluation

The model undergoes comprehensive training using meticulously prepared datasets and subsequent validation using a separate dataset earmarked for this purpose. Evaluation metrics encompassing accuracy, confusion matrix, precision, recall, and F1-score are rigorously computed to evaluate the model's effectiveness in accurately classifying the test dataset.

To ascertain the classification process's accuracy, the quantity of reference data samples is typically determined by a guideline suggesting a minimum of 50 samples per class, as a rule of thumb for constructing the error matrix [53]. The accuracy assessment involves the application of methods such as overall accuracy and kappa coefficient to evaluate the precision and reliability of classification outcomes.

---

III. RESULTS AND DISCUSSION

A. Hybrid Vegetation Index (HVI)

The correlation coefficient of 0.87, 0.75 and 0.81 between HVI and NDVI at Kenya, Egypt, and Nigeria respectively stands as a testament to the robust relationship and shared information content between the two indices. The magnitude of 0.87 suggests a substantial and reliable connection, affirming that the two indices effectively capture similar trends and variations in vegetation conditions. This compelling correlation underscores the efficacy of the novel index HVI, reinforcing its capacity to mirror the patterns identified by the widely used NDVI. The strength of this positive correlation reinforces the potential of HVI to serve as a valuable alternative or enhancement to NDVI.

Furthermore, through visualization analysis, the superiority of HVI over NDVI becomes evident. The visual examination of vegetation patterns using HVI reveals a more detailed and nuanced representation of vegetation health and structure compared to NDVI as shown in figure 5. Notably, the values of HVI consistently surpass those of NDVI, signifying that HVI provides higher quantitative measures of vegetation properties as shown in figure 6. Particularly in regions like Kenya, Egypt, and Nigeria, where HVI not only demonstrates its prowess in providing a more comprehensive understanding of vegetation dynamics but also yields higher values, it stands out as a reliable and advanced tool for vegetation monitoring.

Fig 5: HVI and NDVI images of (a) Kenya, (b) Egypt (c) Nigeria

B. Sentinel-1 Results

Sentinel-1 backscatter analysis represents a fundamental aspect of this research, entailing the examination and interpretation of radar signals emitted and received by the Sentinel-1 satellite sensor. The analysis primarily focuses on
the backscattered microwave signals, providing essential information about surface characteristics and changes within the observed areas. Backscatter analysis involves the extraction of $\sigma_0$ (sigma naught) values from Sentinel-1 imagery, obtained through the analysis of radar backscattered signals in VV (vertical transmit and vertical receive) and VH (vertical transmit and horizontal receive) polarization modes. This examination of backscatter data facilitates insights into surface roughness, and changes in vegetation cover, thereby aiding in the identification and characterization of land features, including different land covers and land use patterns. The following figure (7) shown the sigma naught values of the time series of sentinel-1 images.

The classification of Sentinel-1 data involves two approaches: Random Forest (RF) and Convolutional Neural Network (CNN). RF employs decision trees on $\sigma_0$ values from VV and VH polarization modes for land cover mapping. In contrast, CNN, a deep learning architecture, directly learns complex patterns from Sentinel-1 backscatter data. These methods are compared to map irrigated croplands in Kenya. Evaluating RF and CNN aids in understanding their effectiveness for Sentinel-1 data classification in agricultural mapping. Figure (8) shows the irrigated croplands map for CNN and RF respectively. The comparison of classification results reveals high accuracy for both the Convolutional Neural Network (CNN) and Random Forest (RF) models. The CNN achieves a superior overall accuracy of 92.00%, 93.00% and 92.00% for Kenya, Egypt and Nigeria respectively. The Kappa coefficient were 0.878, 0.892 and 0.878 indicating its robust performance in mapping irrigated croplands across the study areas. Meanwhile, the RF model achieves a slightly lower accuracy of 90.01%, 91.00% and 90.20% with a Kappa coefficient of 0.865, 0.876 and 0.868 for Kenya, Egypt and Nigeria respectively. Despite the marginal difference, both models demonstrate strong performance in accurately delineating land cover types using Sentinel-1 data. The CNN's slightly better accuracy highlights its capability in capturing intricate features within radar data, suggesting its potential for more detailed and precise agricultural mapping compared to RF.
C. Irrigated croplands from combined methods

In this study, we implemented a comprehensive strategy by merging Sentinel-1 and Sentinel-2 imagery with critical vegetation indices (HVI, MSAVI) and PCA techniques, to map irrigated croplands. Utilizing CNN, RF, and SVM classifiers, this approach aimed to capitalize on diverse datasets and indices' strengths in characterizing agricultural landscapes across Kenya, Egypt, and Nigeria. By integrating spectral and temporal information from Sentinel-1 and Sentinel-2 data alongside vegetation indices and PCA, the study sought to enhance the accuracy and effectiveness of irrigated cropland mapping. The following figure (9) shown the irrigated croplands map for CNN, RF and SVM respectively.

The CNN model showcased exceptional performance, achieving 99% accuracy in land cover classification with minimal misclassifications across various land cover categories. This high accuracy, as confirmed by the Confusion Matrix, reflects minimal misclassification across various land cover categories. With precision, recall, and F1 Score all at 99%. These results highlight the CNN's strong potential as a dependable and precise tool for land cover classification, particularly in agricultural mapping tasks.

The amalgamation of Sentinel-1 and Sentinel-2 datasets has notably enhanced the accuracy of classification methodologies for mapping irrigated croplands across Kenya, Egypt, and Nigeria. Leveraging both Sentinel-1 radar and Sentinel-2 optical imagery alongside key vegetation indices and PCA, the combination of these diverse data sources has significantly improved the overall accuracy of Convolutional Neural Network (CNN) classifier by approximately 5-6%. Notably, in Kenya, the utilization of Sentinel-1 and Sentinel-2 together elevates the CNN's overall accuracy to 97.5% with a Kappa coefficient of 0.9621, outperforming RF (94% accuracy, Kappa 0.9094) and SVM (93% accuracy, Kappa 0.8946). Similarly, in Egypt, the integrated use of these satellite datasets enhances CNN's accuracy to 98.01% with a Kappa of 0.9712, surpassing RF (93% accuracy, Kappa 0.8935) and SVM (92% accuracy, Kappa 0.8791). Additionally, in Nigeria, the combined Sentinel-1 and Sentinel-2 approach boosts CNN's accuracy to 98% with a Kappa coefficient of 0.9711, exceeding RF (93% accuracy, Kappa 0.8934) and SVM (94% accuracy, Kappa 0.9083).

These outcomes underscore the invaluable impact of integrating multiple satellite datasets in improving the precision and robustness of classifiers, particularly CNN, in accurately mapping irrigated croplands across diverse landscapes. The fused information from Sentinel-1, Sentinel-2, alongside vegetation indices, and PCA seems pivotal in achieving superior classification results, reinforcing the effectiveness of this comprehensive approach in agricultural land mapping. Table III show the summary of all results.

### Table III

<table>
<thead>
<tr>
<th>Study area</th>
<th>accuracy</th>
<th>CNN</th>
<th>RF</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenya</td>
<td>Overall accuracy</td>
<td>97.5%</td>
<td>94%</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td>Kappa coefficient</td>
<td>0.962</td>
<td>0.909</td>
<td>0.895</td>
</tr>
<tr>
<td>Egypt</td>
<td>Overall accuracy</td>
<td>98.01%</td>
<td>93%</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>Kappa coefficient</td>
<td>0.971</td>
<td>0.894</td>
<td>0.879</td>
</tr>
<tr>
<td>Nigeria</td>
<td>Overall accuracy</td>
<td>98%</td>
<td>93%</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td>Kappa coefficient</td>
<td>0.971</td>
<td>0.893</td>
<td>0.908</td>
</tr>
</tbody>
</table>

D. Accuracy improvement

This study pioneers a novel integrated approach that harnesses the synergies between Sentinel-1’s radar imagery, Sentinel-2's optical data, and key vegetation indices (HVI, MSAVI), complemented by Principal Component Analysis (PCA), thereby significantly enhancing the accuracy of mapping irrigated croplands across diverse African landscapes. The innovative fusion of Sentinel-1 and Sentinel-2 datasets alongside essential vegetation indices facilitated a substantial leap in classification accuracy, marking a noteworthy advancement in remote sensing-based agricultural mapping.
By leveraging Convolutional Neural Network (CNN), Random Forest (RF), and Support Vector Machine (SVM) classifiers, our approach showcased an approximate 5-6% enhancement in overall accuracy when compared to the utilization of individual sensor applications. Notably, this integration achieved remarkable accuracies of 97.5%, 98.01%, and 98% in Kenya, Egypt, and Nigeria, respectively, underscoring the superior performance of the combined Sentinel data.

This innovative methodology highlights the pivotal role of integrated satellite datasets in revolutionizing precise agricultural mapping practices and advancing informed decision-making in vital African agricultural regions. Table. IV shows the accuracy improvement due to the innovate method.

### TABLE IV

<table>
<thead>
<tr>
<th>Study Area</th>
<th>Classification Method</th>
<th>Sentinel-1 Only (Accuracy %)</th>
<th>Sentinel-1 &amp; Sentinel-2 Integrated (Accuracy %)</th>
<th>Improvement of Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenya</td>
<td>CNN</td>
<td>92.00</td>
<td>97.50</td>
<td>+5.50</td>
</tr>
<tr>
<td>Egypt</td>
<td>CNN</td>
<td>93.00</td>
<td>98.01</td>
<td>+5.01</td>
</tr>
<tr>
<td>Nigeria</td>
<td>CNN</td>
<td>92.00</td>
<td>98.00</td>
<td>+6.00</td>
</tr>
</tbody>
</table>

### E. Discussion

The innovative method employed in this study revolves around the integrated use of Sentinel-1 and Sentinel-2 datasets, coupling key vegetation indices (HVI, MSAV1), and employing PCA techniques. This comprehensive approach aimed at mapping irrigated croplands across diverse African landscapes leveraged the strengths of multiple datasets and analytical techniques. The integration of Sentinel-1 radars imagery and Sentinel-2 optical data, alongside vegetation indices and PCA, marked a pioneering strategy in remote sensing-based agricultural mapping. The fusion of these diverse datasets served as a cornerstone, enabling a holistic understanding of agricultural landscapes by capturing spectral, temporal, and spatial information in a unified framework.

Adding to the discussion, the inclusion of the novel Hybrid Vegetation Index HVI significantly impacted the study's outcomes. The HVI, derived from the synergistic combination of Sentinel-1 and Sentinel-2 bands, played a crucial role in enhancing the discriminatory power of the classification models.

Previous studies have discussed that issue, such as Y. Chen et al. 2018 whom got overall accuracy of 90% and Bazzi,H et al. 2019 whom got overall accuracy of 94%. This study has got 98% overall accuracy and 0.971 Kappa coefficient. However, the study is not without limitations. The absence of ground truth data, which restricted the validation of the classification results is a limitation point. Furthermore, despite achieving high accuracy in delineating irrigated croplands with moderate -resolution Sentinel imagery, the study's methodology might encounter limitations in capturing finer details or smaller-scale features due to inherent constraints in the resolution of the satellite data.

Looking ahead, the future of agricultural mapping holds promising avenues for improvement and advancement. Incorporating very high-resolution satellite imagery from platforms like WorldView and GeoEye presents an exciting prospect. The utilization of these cutting-edge satellites, with their exceptional spatial resolution capabilities, holds the potential to overcome limitations associated with moderate-resolution data. The increased spatial detail offered by such very high-resolution imagery promises enhanced accuracy and precision in mapping agricultural landscapes. Integrating these advanced datasets into the existing framework could refine classification models, enabling more detailed assessments and validations. This advancement is poised to revolutionize precision agriculture, facilitating more informed decision-making in agricultural planning, resource allocation, and sustainable land management initiatives across diverse landscapes.

### IV. Conclusion

This research signifies a notable advancement in accurately delineating irrigated croplands across diverse landscapes in Africa, leveraging the combined potential of Sentinel-1 and Sentinel-2 satellite datasets. The critical importance of precisely delineating agricultural areas cannot be overstated, particularly in influencing agricultural policies, optimizing resource allocation, and bolstering food security initiatives. Through a meticulous integration of cutting-edge satellite data with advanced machine learning models like Convolutional Neural Networks (CNN), Support Vector Machines (SVM), and Random Forest (RF), this study showcases a significant advancement in accuracy compared to traditional methods. Importantly, the integration of the novel hybrid vegetation index (HVI) into the classification models has further elevated the precision of irrigated cropland mapping. The HVI, derived from a synergistic combination of Sentinel-1 and Sentinel-2 bands, has played a pivotal role in enhancing the discriminatory power of the classifiers. The incorporation of HVI resulted in an additional improvement in classification accuracy. Significantly, the fusion of Sentinel-1 and Sentinel-2 datasets demonstrated a remarkable improvement—approximately 5-6%—in classification accuracy when compared to using either dataset in isolation. This integration enabled the CNN model to achieve unparalleled accuracy peaks of 97.5%, 98.01%, and 98% in Kenya, Egypt, and Nigeria, respectively, surpassing RF and SVM by considerable margins.
These findings underscore the indispensable role of synergizing diverse satellite datasets, especially the powerful combination of Sentinel-1 and Sentinel-2, in significantly enhancing classifier precision for meticulous and accurate mapping of irrigated croplands. The collaborative use of Sentinel-1’s radar capabilities and Sentinel-2’s optical data represents a transformative approach in remote sensing-based agricultural mapping. This integrated methodology holds immense promise in optimizing agricultural management practices, bolstering food security initiatives, and facilitating informed decision-making in land-use planning across Africa. The results underscore the undeniable potential of Sentinel data as a pivotal tool in mapping irrigated areas, offering a robust framework that stands at the forefront of revolutionizing precision agriculture and resource allocation strategies in the region.

ACKNOWLEDGMENTS

We extend our heartfelt gratitude for the invaluable support received from the Jiangsu Marine Science and Technology Innovation Project, specifically through Grant No. JSZRHYKJ2020202, which played a crucial role in making this study possible.

V. REFERENCES


