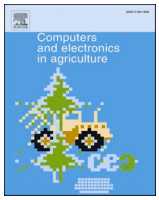
Computers and Electronics in Agriculture 221 (2024) 108964

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag



Empirical curvelet transform based deep DenseNet model to predict NDVI using RGB drone imagery data

Mohammed Diykh [a,f,g], Mumtaz Ali [a,b,*], Mehdi Jamei [b,d], Shahab Abdulla [a], Md Palash Uddin [e], Aitazaz Ahsan Farooque [b,c,*], Abdulhaleem H. Labban [h], Hussein Alabdally [a]

[a] UniSQ College, University of Southern Queensland, QLD 4305, Australia
[b] Canadian Centre for Climate Change and Adaptation, University of Prince Edward Island, St Peters Bay, PE, Canada
[c] Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE C1A4P3, Canada
[d] Faculty of Civil Engineering and Architecture, Shahid Chamran University of Ahvaz, Ahvaz, Iran
[e] School of Information Technology, Deakin University, Geelong, VIC 3220, Australia
[f] University of Thi-Qar, College of Education for Pure Science, Iraq
[g] New Era and Development in Civil Engineering Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar, Nasiriyah 64001, Iraq
[h] Department of Meteorology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Keywords:
NDVI
RGB
DenseNet
Curvelet coefficients
Drone image
Prediction

ABSTRACT

Predicting accurately the Normalized Difference Vegetation Index (NDVI) trends from RGB images are essential to monitor crops and identify issues related to plant diseases, and water shortages. The current NDVI prediction models are primarily based on traditional machine learning models which lack reliability due to the problem related to atmospheric conditions. To predict NDVI in Prince Edward Island using RGB drone imagery data, this paper proposed a novel framework integrating empirical curvelet transform and DenseNet models. Each channel of RGB drone imagery data was passed through empirical curvelet transform method where the curvelet coefficients were analysed which result in creating a new formula to design NDVI. The output of the new formula was sent to the deep DenseNet to predict the final NDVI. The proposed model was evaluated using quantitative metrics including, Q-Q plot, regression, correlation coefficients, structural similarity (SSIM), peak signal to noise ratio (PSNR) and mean square error (MSE) as well as accuracy (ACC), sensitivity (SEN), f1-score, specificity. The obtained results showed that the proposed model outperformed the previous models by scoring the highest values of SSIM = 0.98, and lowest MSE = 120. It is believed that the proposed model is helpful to support farmers in monitoring the growth and plant health as well as to identify crops problems.

1. Introduction

NDVI prediction from RGB drone imagery data is crucial for identifying crop problem, growth, and plant health. Mainly, Near-infrared (NIR) wavelength is required to generate a NDVI map. The NIR is designed based on multispectral cameras and sensors which are time consuming and expensive compared to the RGB cameras. However, designing a low cost NDVI map using standard RGB cameras is crucial to provide decision makers with the important crop health information.

Several techniques have been suggested to monitor crops health for example, Remote sensing (RS) which is an old tool, and from time to time, it has come in handy for a lot of research work. Li et al. (2020) showed how remote sensing data could help monitor ecological elements like vegetation index, soil moisture, and evapotranspiration with an accuracy of 78%-94%, 17.7%–32.8%, and <12%, respectively. Vinukollu et al. (2011) used multi-sensor remote sensing data for global estimates of evapotranspiration for climate studies. Eguchi et al. (2008) used remote sensing for disaster management and offered to build consensus-based damage assessment criteria based only on remote-sensed data for standardizing the results of multi-event assessments. Devi et al. (2015) showed another application of remote sensing in satellite oceanography. Their paper showed how IR and microwave radiometers could be used to measure temperature at different depths in the ocean. In the past years, RS has also become a potential tool to identify potential fishing zones. Canada Centre for Remote Sensing (CCRS) scientists are also working on developing a sustainable way to monitor land cover, water, and wetland transformation (Development, 2023).

Generative Adversarial Network (GAN) is a deep learning model

* Corresponding authors.
E-mail addresses: mumtaz.ali@unisq.edu.au (M. Ali), afarooque@upei.ca (A.A. Farooque).

https://doi.org/10.1016/j.compag.2024.108964
Received 18 December 2023; Received in revised form 13 March 2024; Accepted 16 April 2024

Nomenclature

PEI	Prince Edward Island
CCRS	Canada Centre for Remote Sensing
NDVI	Normalized Difference Vegetation Index
DGPS	Differential Global Positioning System
EWT	Empirical wavelets transform
SSIM	Structural similarity
ACC	Accuracy
SPE	Specificity
PSNR	Peak signal to noise ratio
MSE	Mean square error
SEN	Sensitivity
RS	Remote sensing
UAVs	Unmanned Aerial Vehicle
NIR	Near-infrared
GANs	Generative Adversarial Networks
DN	Digital number
FB	Filter bank

introduced by Ian Goodfellow and his colleagues (Giles, 2018). In recent years, GANs have been used as a critical tool to perform image transformation from one domain into another while preserving its essential characteristics. GAN consists of two major components: Generator and Discriminator. The former generates images that are indistinguishable from the actual photos in the target domain, and the latter is a binary classifier that evaluates whether an image is real or fake. The generator produces realistic photos to fool the discriminator while the discriminator tries to become better at telling real from counterfeit images (Aggarwal et al., 2021). GANs have been successfully applied to a wide range of image translation tasks like SR-GAN for crop health detection (Maqsood et al., 2021), weed detection (Sapkota et al., 2022), SA-GAN for animal farming (Li and Tang, 2020), Cycle GAN for fruit detection (Barth et al., 2020) and aquaculture (Zhang et al., 2021). GANs have not only performed well on the ground level but have also done well as Underwater Multiscene Generative Adversarial Network (UMGAN) to enhance underwater images. Although the mentioned studies predicted NDVI, establishing an accurate and standard NDVI based on deep learning remains a challenging task. It is important to monitor plant growth accurately in Prince Edward Island. To achieve that We found that empirical wavelet transform (EWT) can reveal the important features of RGB images. As a results, in this paper, we utilise the empirical curvelet transform to extract curvelet coefficients to design a new NDVI formula.

Many recent studies have been designed to analyse time series data to predict NDVI for example Beyer et al., (2023) designed a NDVI prediction model based on graph based deep learning model. In that study, Graph WaveNet model was proposed to predict NDVI from time series data. The data collected form the period of June 2002 to May 2019, including 763 different dates. Omar and Kawamukai, (2021) suggested an architecture based on Holt-Winters to design NDVI index to monitor crops. Li et al., (2022) generated NDVI time series using Transformer Temporal-spatial Model based deep learning model. Their model was tested using time series data collected from South of the Songnen region. Guo et al., (2024) designed a deep learning model for NDVI prediction in YRB of Chaina. In that study, multilayer multivariate LSTM neural network was employed to predict the actual values of NDVI. Emami et al. (2020) introduced the spatial GAN model (SPA-GAN), which lets the generator focus more on the most discriminative regions between the source and the target domains. Yi et al. (2017) performed a study and introduced the concept of DualGAN, which simultaneously learns two image translators from one domain to another and, as a result, performs a wide variety of image-to-image translation tasks; comparing the results showed that DualGAN could outperform supervised, trained methods on labelled data. DenseNet model has been used for melanoma cancer detection (Girdhar et al., 2023), image classification (Abai and Rajmalwar, 2019; Zhang et al., 2019), recognition and mapping of landslide (Gao et al., 2021, Liu et al., 2021), and remote sensing scene classification (Zhang et al. 2019).

Most designed NDVI prediction models are primarily based on analyzing time series data using traditional neural network techniques. Those models showed several limitations regarding the designing NDVI formula, causing the NDVI based time series prediction models to lack reliability. The study conducted in this paper demonstrates an excellent approach due to the researchers' advantage in utilizing advanced technology and abundant resources, distinguishing it from other studies. To predict NDVI from drone imagery data, this study constructed an artificial model based on curvelet decomposition and deep learning model to eliminate the issues associated with those models.

Although the previous studies have suggested several NDVI models that used all image channels information. In this paper, we made an attempt to utilise the texture feature of RGB image to predict NDVI. The main contribution of this paper is to design an accurate NDVI model utilizing an empirical curvelets transform based model to analyse texture patterns of RGB images. The coefficients of empirical curvelets were tested and used to design a new NDVI formula. The designed NDVI formula was fed to DenseNet model directly instead of using the original images pixels. It is widely acknowledged that wavelet coefficients were more powerful to analyse RGB images.

2. Data collection and dataset generation

The data collection consisted of a fixed-wing Unmanned Aerial Vehicle (UAV) model called the Ebee X, developed by SenseFly Inc. based in Cheseaux-Sur-Lausanne, Switzerland, was employed. This UAV was equipped with a RedEdge-MX sensor, developed by MicaSense, headquartered in Seattle, Washington, USA. The data collection occurred in Prince Edward Island during the 2021 growing season, from May to October. The selected fields were located in various regions of the island: to the west (O'Leary at coordinates 47.0601196° N, 62.4924976° W), to the east (Souris at 46.259217° N, 63.137842° W), and in the central areas (Oyster Cove at 46.471967° N, 63.694683° W; Tryon at 46.2416057° N, 63.5351343° W; and Bedeque at 46.3477236° N, 63.7577183° W).The imagery was captured throughout the potato growing season, covering key growth stages, including:

1. Vegetation stage (15–30 days after planting, DAP).
2. Tuber formation stage (30–45 DAP).
3. Tuber bulking stage (75–110 DAP).
4. Tuber maturation stage (more than 110 DAP).

The RedEdge-MX sensor possesses the capability to capture images in a variety of spectral bands, including blue, red, green, near-infrared (NIR), and red-edge, with respective central wavelengths of 475, 560, 668, 840, and 717 nm. To ensure precise positional accuracy, the captured imagery uses a built-in Differential Global Positioning System (DGPS) with real-time kinematics protocol integrated into the Ebee X UAV. The execution of regular interval image capture from agricultural fields, with a 75% frontal and side image overlap, was facilitated using eMotion flight mission planning software, developed by SenseFly Inc. of Cheseaux-sur-Lausanne, Switzerland.

Before and after each survey, the RedEdge-MX sensor underwent radiometric calibration to correct for brightness and ensure the accuracy of the collected data. Most surveys were conducted on days with ample sunlight and wind speeds not exceeding 35 km per hour, contributing to stable imagery from the sensor.

After data collection, radiometric calibration was employed to convert the image's brightness values into spectral radiance values. This conversion facilitated the determination of nitrogen, chlorophyll, and

biomass levels within the images, as well as the calculation of vegetation indices. Additionally, raw digital number (DN) values were ascertained for each pixel in each image. The standard radiometric calibration procedure for drone imagery involved the utilization of reflectance panels and the built-in functions of Pix4D software. Gains and offsets for the radiometric sensors were calculated by comparing the sensors' DN values to those of the standard reflectance panel. The conversion to radiance accounted for various factors, including reflected and incident radiances, as well as solar scattering and absorption. However, atmospheric variables, such as water vapor and humidity, could still be present, potentially leading to inaccurate data. To mitigate these atmospheric variables, the radiance data were further transformed into reflectance, yielding more precise and reliable data.

The images collected by the Ebee X UAV were prepared in a suitable format for the training of Generative Adversarial Networks (GANs). This study made use of two image datasets: RGB and NDVI indices. An RGB image is a data array where each pixel is defined in terms of red, green, and blue channels. RGB cameras are commonly employed in standard drones, enabling the capture of colour aerial images via the combination of these three channels. The RGB option is favoured for crop image capture due to its relatively low cost and ease of use. On the other hand, NDVI, a combination of NIR and red channels, is widely utilized in vegetation monitoring. However, the inclusion of the NIR band increases the cost of the sensor, making it more expensive for vegetation monitoring purposes. The NDVI index can be calculated as follows:

$$\text{Discriminator Loss} = \frac{\text{NIRChannel} - \text{RedChannel}}{\text{NIRChannel} + \text{RedChannel}}$$

Incorporating the NIR band makes the sensor more expensive compared to those operating in the visible domain. The computed NDVI index image reflects the health of vegetation and symptoms of plant stress. Healthy plants are depicted in green, while barren land, roads, sheds, and soil appear in various shades of red. An index scale ranging from 0 to 1 can be added to quantify plant health in different areas, but, for GAN training purposes in this study, an index scale was not included in the NDVI computed images and may be added in the future. The GAN training dataset consisted of 500 pairwise images, distributed across training (80%, 400 images), validation (10%, 50 images), and testing (10%, 50 images) phases.

3. Methodology and model development

Most previous models for NDVI prediction do not adequately consider analysis of the texture features from RGB images, which were not fully considered in the designing of NDVI formulas. However, the analysis of texture features using decomposition techniques is very important to design standard NDVI formula. In response, this study establishes a more realistic NDVI prediction model for plants diseases detection, monitoring the productivity and health of crops. The proposed approach is novel in terms of empirical curvelet transform method, construction of the NDVI based on the new formula, and DenseNet model to predict NDVI using RGB drone imagery data. Here each channel of drone RGB imagery data was passed through empirical curvelet transform method to decompose the data where the curvelet coefficients were analysed and calculated. The curvelet coefficients were then used to construct the new formula to design the NDVI. Lastly, the output of the new formula was sent to DenseNet model to predict the final NDVI.

In this paper, we proposed a newly designed model for NDVI prediction using empirical curvelet transform coupled with DenseNet model. The methodology consists of 3 steps, obtaining Curvelet coefficients for RGB channels, designing NDVI index, training and testing the model using DenseNet. These are discussed in detail in the coming sections. Fig. 1 depicts the schematic view of proposed modelling framework.

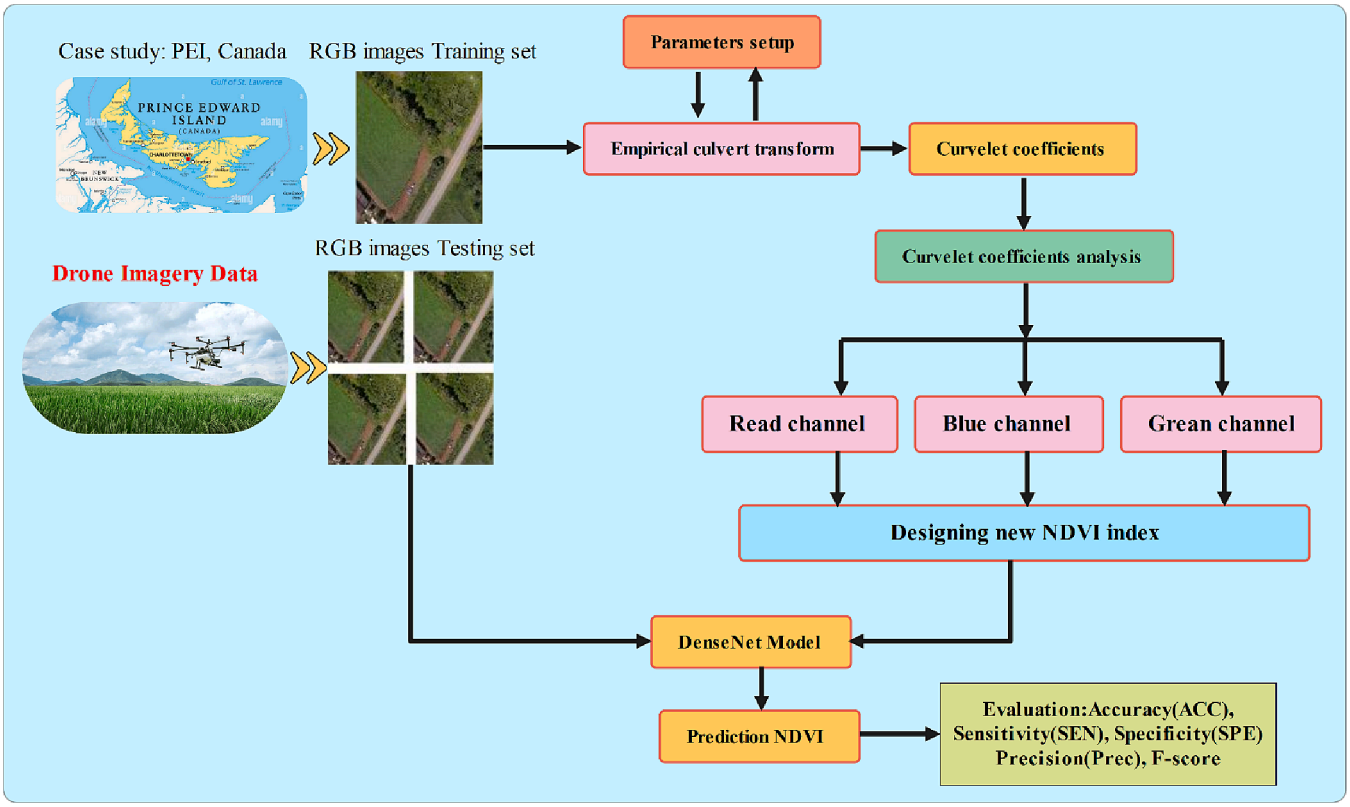


Fig. 1. Flowchart of the proposed model for NDVI prediction.

3.1. Two-dimension empirical wavelet transform (EWT)

NDVI images are comprised of various texture patterns. These texture patterns are complex and can be depicted as a combination of several oscillations. The goal of this section is to give an overview about EWT. Texture analysis is a difficult task as it requires to investigate all types of irregularity within textures.

The EWT was originally designed as a signal decomposition tool. It has been used to separate and detect principal harmonic modes of signals (Zhang et al., 2022). The EWT comprises of two phases. First, it divides the Fourier spectrum into N supports, this step involves building the Fourier domain a filter bank. Second, it filters the input signal using the built filter bank to obtain the different components. The filter bank (FB) contains N wavelet filters in which one low-pass filter represents the approximation component while the $N-1$ bandpass filters represent the details components (Drees et al., 2021).

Suppose the Fourier spectrum is separated into N segments with boundaries $\{\varphi_n\}_{n=0}^{N}$ where $\varphi_n = \pi, \varphi_0 = 0$. Using Meyers wavelet, a filter bank of wavelet is constructed.

$$\{\varphi_1(x), \{\mu(x)\}_{n=1}^{N-1}\} \tag{1}$$

The Fourier transform (f_1) of empirical scaling function is expressed as

$$f_1(\varphi_1)(\omega) = \begin{cases} 1 if |\omega| \le (1-\mu)\varphi_1 \\ \cos[\frac{\pi}{2}\beta(\frac{1}{2\mu\varphi_1}(|\varphi| - (1-\mu)\varphi_1))] if (1-\mu)\varphi_1 \le |\omega| \le (1+\mu)\varphi_1 \\ 0 otherwise \end{cases} \tag{2}$$

where the FT of EWD is given by

$$f_1(\varphi_n)(\omega) = \begin{cases} 1 if (1+\mu)\omega_n \le |\omega_n| \le (1-\mu)\omega + 1 \\ \cos[\frac{\pi}{2}\beta(\frac{1}{2\mu\omega_{n+1}}(|\omega| - (1-\mu)(\omega_n+1)\,] if (1-\mu)\omega_n \le |\omega| \le (1+\mu)\omega_n + 1 \\ \sin[\frac{\pi}{2}\beta(\frac{1}{2\mu\omega_n}(|\omega| - (1-\mu)(\omega_n)\,] if (1-\mu)\omega_n \le |\omega| \le (1+\mu)\omega_n \\ 0 othersie \end{cases} \tag{3}$$

where $n = 1, \cdots., N$, and β is an arbitrary function. The EWT is defined as the same way as the WT. The coefficients are obtained as

$$w_f^{\varepsilon}(n,x) = F_1(F_1((f)(\varphi)\overline{F_1(\varphi_n)(\omega)})(x) \tag{4}$$

$$W_m(\varphi) = \begin{cases} 1 if (1+\mu)\omega_m \le |\omega| \le (1-\mu)\omega_m + 1 \\ \cos[\frac{\pi}{2}\beta(\frac{1}{2\mu\omega_m+1}(|\omega| - (1-\mu)\omega_m))] if (1-\mu)\omega_m \le |\omega| \le (1+\mu)\omega_m + 1; \\ \sin[\frac{\pi}{2}\beta(\frac{1}{2\mu\omega_m+1}(|\omega| - (1-\mu)\omega_{n-1}))] if (1-\mu)\omega_{n-1} \le |\omega| \le (1+\mu)\omega_n + 1; \\ 0 otherwise \end{cases} \tag{7}$$

And the approximated coefficients are given as

$$w_f^{\varepsilon}(o,x) = F_1(F_1((f)(\varphi)\overline{F_1(\varphi_1)(\omega)})(x) \tag{5}$$

3.2. Curvelet coefficients analysis to construct new NDVI index

To construct an empirical curvelet filter, the procedure in (Emami, Aliabadi et al. 2020) was adopted in this paper. For each input image *im*, s set of scale boundaries $\varphi_{\omega}^{i} = \{\omega_i^n\}, n = 0, \cdots.., n = N_s^i$, and a set of angular boundaries $\varphi_{\alpha}^{i} = \{\alpha_i^m\}, m = 0, \cdots.., n = N_{\alpha}^i$ are detected and analysed, Where N_s^i and N_{α}^i are to the number of detected angular sectors and scales for the image *im*. Then, we employed the merging technique illustrated above on $\varphi_{\omega}^{i} = \{\omega_i^n\}, \varphi_{\alpha}^{i} = \{\alpha_i^m\}$ to remove the useless supports. A threshold was employed to explore small supports. Threshold was set to $T_{\omega} = 0.2, T_{\alpha} = 0.06$. As a result, we obtained two final sets of scales and angular boundaries.

Based on above notation, the empirical curvelet filter are designed using a lowpass filter $\varnothing 1$ constructed using $\omega, and \alpha$ that corresponds to the polar coordinates in Fourier domain and the 2D Fourie terraform $f2$, where $\forall \alpha \in [0, 2\pi]$

$$f2(\varnothing 1)(\omega,\alpha) = \begin{cases} 1 if |\omega| \le (1-\mu)\omega_1 \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\mu\omega}(|\omega| - (1-\mu)\omega)\right)\right] if 1 if |\omega| \le (1-\mu)\omega_1 \\ 0 otherwise \end{cases} \tag{6}$$

where β is an arbitrary function belonging to $C^k([0,1])$, and fulfilling the following conditions.

$\beta(x) = 0 if x \le 0$

$\beta(x) = 1 if x \ge 1$

$\beta(x) + \beta(1-x) = 1, \forall x \in [0,1]$

The parameter μ allows us to ensure that only two consecutive filters can be overlapped. More details regarding chosen the two parameters in [1,3]. The bandpass curvelet filter f_{nm} is equivalent to the polar wedges in Fourier domain. It's defined as the product of polar window W_m and radial windowV_n. $.F_2(f_{mn}) = W_n V_m$ where

while V_m is calculated as

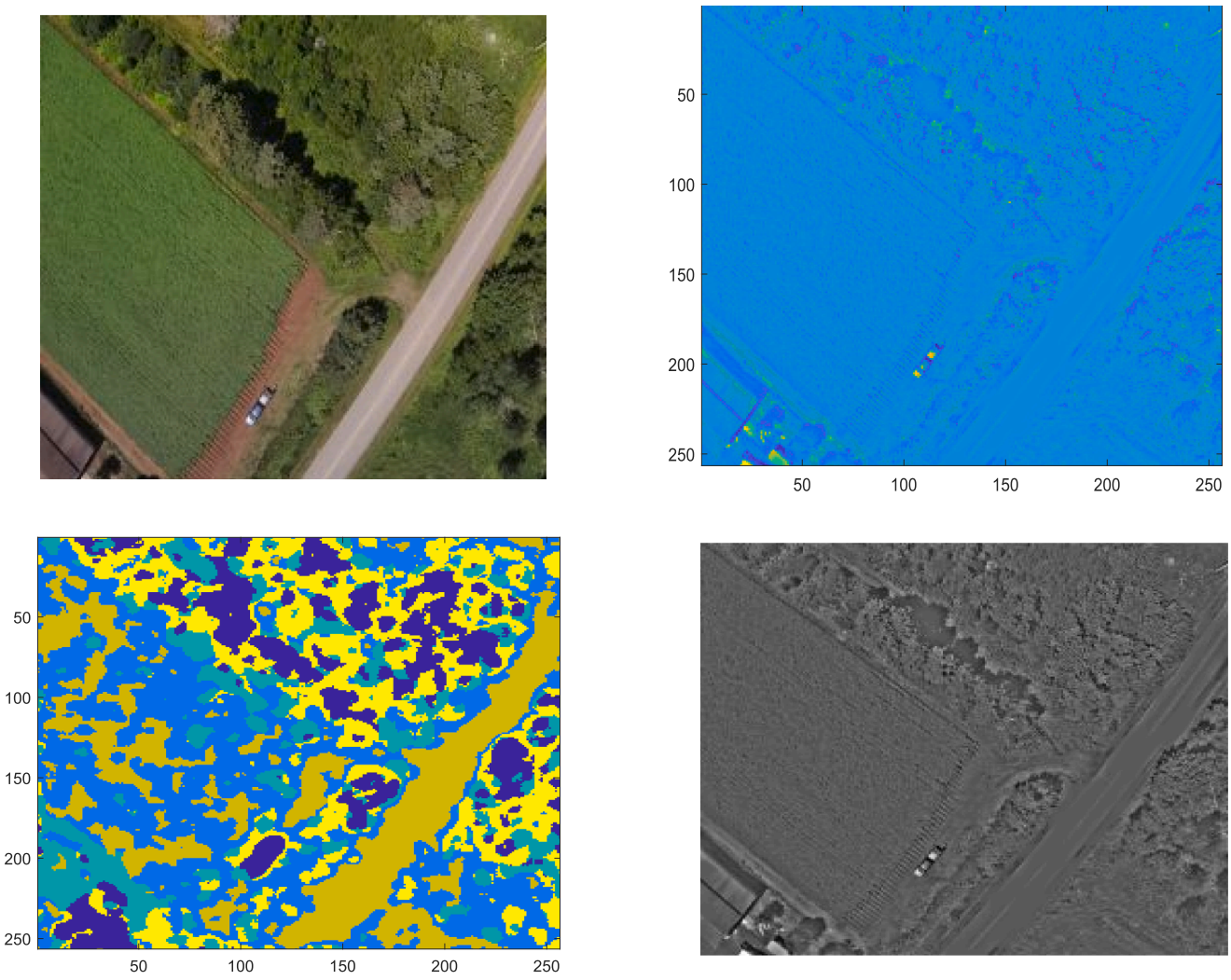


Fig. 2. Texture analysis for (a) RGB image, (b) Texture output, (c) Visual segmentation texture of RGB image and (d) Texture in grey color.

$$V_m = \begin{cases} 1 if \theta_n + \Delta\theta \le \theta \le \theta_{n+1} - \Delta\theta \\ \cos[\frac{\pi}{2}\beta(\frac{1}{2\Delta\theta}(\theta - \theta_{n+1} + \Delta\theta))] if \theta_{n+1} - \Delta\theta \le \theta \le \theta_{n+1} + \Delta\theta \\ \sin[\frac{\pi}{2}\beta(\frac{1}{2\Delta\theta}(\theta - \theta_n + \Delta\theta))] if \theta_n - \Delta\theta \le \theta \le \theta_1 + \Delta\theta \\ 0 otherwise \end{cases} \quad (8)$$

Finally, the corresponding filter bank is obtained by

$$B^{ec} = \{f1(x), \varnothing_{mn}9x)\}, n = 1, \cdots N_s, m = 1, \cdots., N_\theta \quad (9)$$

where x is a pixel in the 2D plane. The empirical curvelet transform is denoted as W^{ec} of image im. This can be expressed as

$$W^{ec}(x) = ((\varnothing^* f)(x), (\varphi_{11}{}^* f)(x), (\varphi_{12}{}^* f)(x), \cdots., (\varphi_{mn}{}^* f)(x), \cdots., \left(\varphi_{m_s n_\theta}{}^* f\right)(x) \quad (10)$$

where * is convolution product. All convolutions are performed as pointwise product in Fourier transform and then, the inverse Fourier transform is performed to obtain curvelet coefficients. Figs. 2, 3, and 4 show examples of texture patterns of RGB, and NDVI image.

3.3. Designing new NDVI index

Our main objective is to use the curvelet coefficients to construct the new NDVI index. Based on equation (9) to design an NDVI index from RGB channels, we utilized the curvelet coefficients of three channels including red, green, and blue. The proposed index is formulated using the following equation:

$$New_{NDVI} = (R_{texture}{}^* weight, G_{texture}{}^* weight, B_{texture}{}^* weight) \quad (11)$$

Where $R_{texture}$, $G_{texture}$, $RB_{texture}$, are the curvelet coefficients of the three channels (R, G, B) respectively. The weights (*weight*) is empirically chosen. We conducted several experiments based on hit and trail technique to determine the optimal weight value. At each experiment, the new formula was sent to the prediction model (i.e., DenseNet), and the MSE was calculated. The procedure was repeated until, the optimal weight value was obtained for each channel. The histogram of the actual NDVI and the proposed NDVI was plotted in Fig. 5 to show how the two models are close to each other.

Fig. 6 shows an example of comparison between the proposed NDVI against the reference NDVI. The proposed NDVI was formulated from the RGB channels using equation (11). The regression between the proposed NDVI and the actual NDVI was calculated for pixel value.

3.4. NDVI prediction using DenseNet model

In deep learning-based approaches, Convolutional neural networks have been proven to be an effective tool to achieve higher recognition rates. The existing deep learning methods are designed to extract features from single scale. Thus, the model cannot fully extract the feature

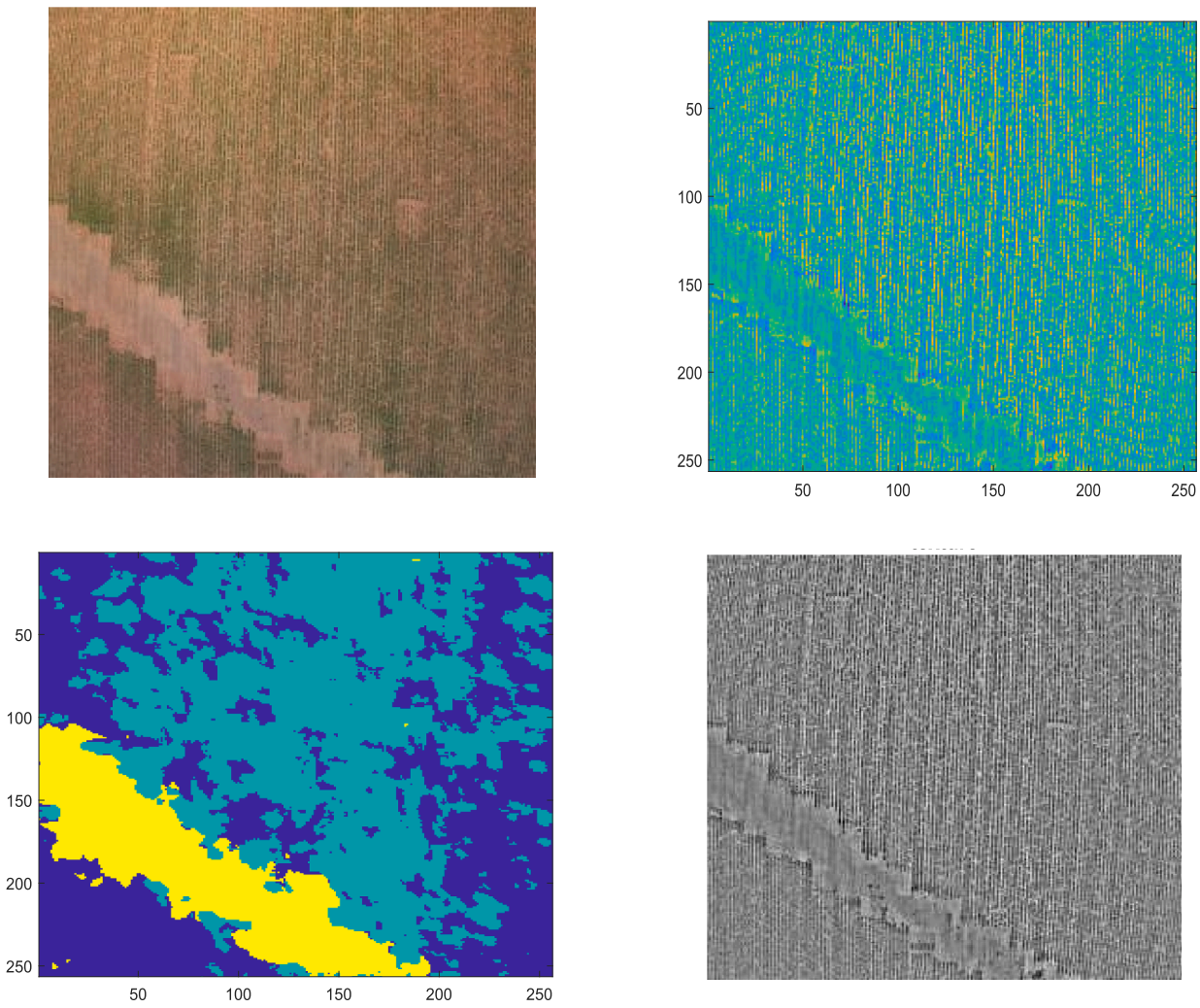


Fig. 3. Example of texture analysis field image for (a) RGB image, (b) Texture output, (c) Visual segmentation texture of RGB image and (d) Texture in grey colour.

information from the data, resulting in poor classification results. The architecture of the proposed DenseNet networks is shown in Fig. 7. The proposed DenseNet is constructed to predict NDVI images correspond to RGB images. The DenseNet model has been applied for various prediction issues due to its ability to accurate predict the target values. For example, Seong et al., (2024) employed DenseNet model for deformation prediction. Albahli et al., (2023) proposed an advanced DenseNet model for prediction of stock market from time series data. The proposed DenseNet model includes an input layer, convolutional layers, a batch normalization layer, max-pooling, and fully connected layers (Yi et al., 2017, Bui et al., 2020). The convolutional layer performs convolution computations on the input images using a convolution kernel. Batch-normalization is responsible for the fostering generalization of DenseNet model. The max pooling layer avoids an overfitting problem that could be appeared by down-sampling.

Finally, the fully connected layer employees the extracted features to deliver the final classification (Gilles, 2013). The DenseNet model consists of three multi-scale blocks with three fully connected layers of sizes 256, 128, and 2.

$$fuulyconnected = \alpha(f^{*}x + c) \tag{12}$$

where f refers to the kernel weight, c denotes bias, α activation function, and x is the input image. A soft max layer was used to determine the probability RGB images using the following equation

$$SW_j = \frac{E^{p_j}}{\sum_{i=1}^{M} E^{p_j}} \tag{13}$$

where SW_j is the probability of the sample belonging to the class p_j, M is the total of classes. In this paper, cross entropy was employed as a loss function. The main formula for the loss function is expressed as

$$Lossfucntion = 1\frac{1}{M} y_i log\overline{y_i} \tag{14}$$

where y_i and $\overline{y_i}$ denotes to the predicted and actual samples. For recurrent updating, and parameters learning, the Mini-Batch Gradient Descent was employed.

$$\alpha = \alpha_r - \theta^{*}\frac{1}{m}\sum_{i=j}^{j+m}(y^{(i)} - h_{\alpha}(x^{i}))x_j^i \tag{15}$$

where m, α, θ, α_r refers to the learning rate, mini-batch size, weight parameters, and feature index. y is the actual label, x refers to the input images, h_{α} denotes the predicted output.

3.5. Hyperparameters setting

Selecting the optimal parameters has a great effect on the performance of the prediction model. In this paper, the parameters of all models used were selected. Table 1 lists the hyperparameters of all

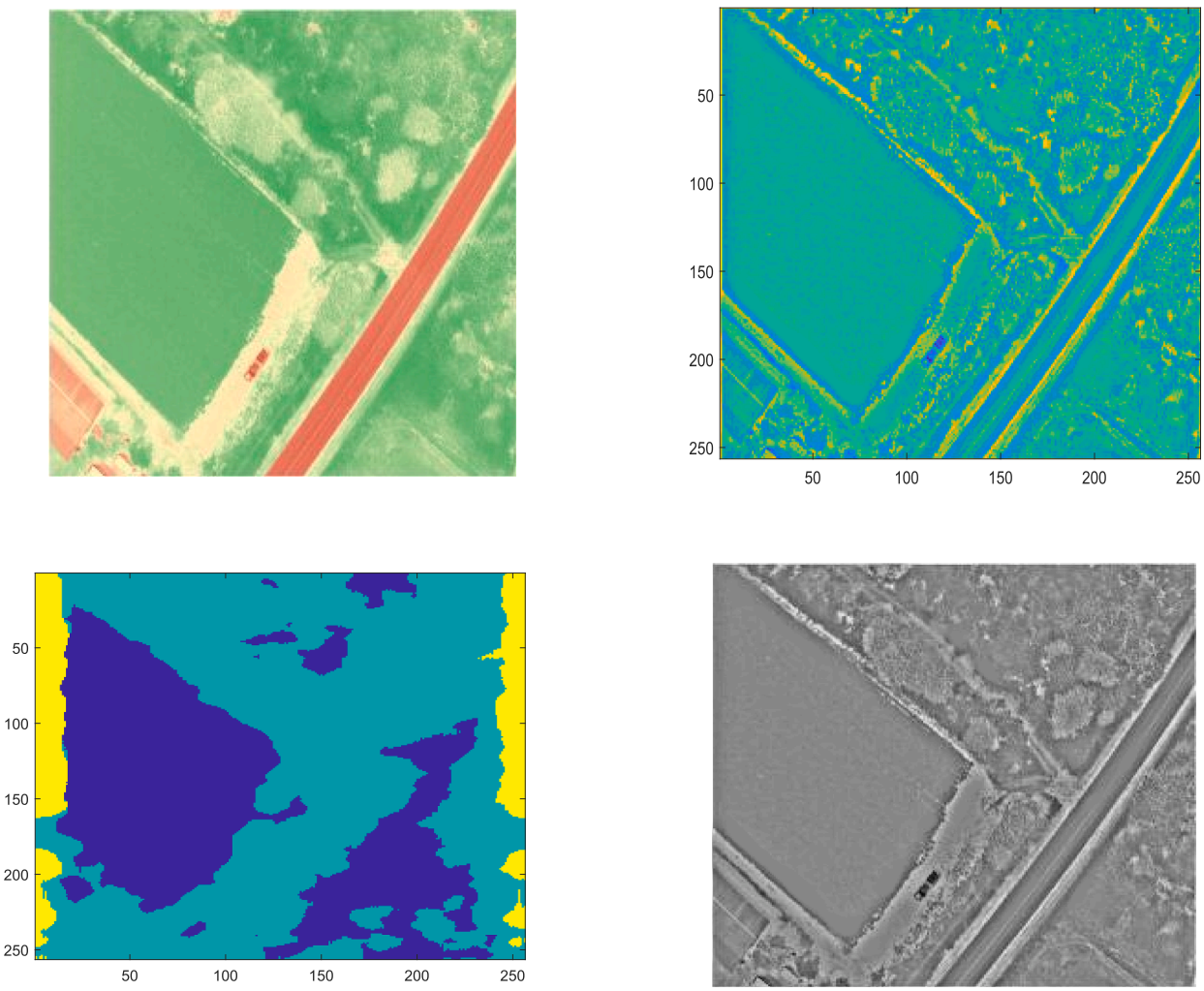


Fig. 4. Texture analysis of (a) NDVI image, (b) Texture output, (c) Visual segmentation texture of NDVI image and (d) Texture in grey colour.

models used in NDVI prediction. Multi-experiments were conducted to select the parameters carefully for optimum accuracy of each model. As mentioned before the loss function was cross entropy, the rate of learning was set to 0.004. The network was trained for 50 iterations with a 34 size of mini batch. The LeakyReLU was used as an activation function. Adam optimizer is used for the network optimizer algorithm.

3.6. Evaluation metrics

To evaluate the proposed model, several metrics were used including accuracy, sensitivity, specificity, precision, recall, and f1-score, and three quantitative metrics named structural similarity (SSIM), peak signal to noise ratio (PSNR) and mean square error (MSE) were employed in this study.

The different metrics used are defined as follows (Winkler and Mohandas, 2008, Huynh-Thu and Ghanbari, 2012, Gilles, 2013, Huang et al., 2019, Zeng et al., 2019, Gan et al., 2021, Jahmunah et al., 2023):

$$\text{Accuracy(ACC)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{16}$$

$$\text{Sensitivity(SEN)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{17}$$

$$\text{Specificity(SPE)} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{18}$$

$$\text{Precision(Prec)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{19}$$

$$\text{F} - \text{score} = 2 \times \frac{\text{Prec} \times \text{SEN}}{\text{Prec} + \text{SEN}} \tag{20}$$

where TP is the total number of correctly predicted NDVI samples, FN is the total number of mistakenly predicted NDVI samples, FP is the total number of incorrectly classified HC samples as MDD cases, and TN is the total number of correctly predicted NDVI. In addition, quantitative accuracy of the model was calculated using structural similarity (SSIM), the following metrics:

$$SSIM(x, y) = \frac{(2av_x 2av_y C_1) + (2covr_{xy} + C_2)}{(av_x^2 av_y^2 yC_1)(ver_x^2 + ver_y^2 + C_2}} \tag{21}$$

where x, and y has the same length, av_x is the average of x, av_y is the average of y, ver_x^2 is the variance of x, ver_y^2 is the variance of y, $covr_{xy}$ is the covariance of x, y, and C_1, and C_2 are the variables to stabilize the week dominator. The value of SSIM rages from 0 to 1, where 0 indicates no image similarity, and 1 indicates identical image.

We also utilized MSE and RMSE. The MSE is defined as the average squared error between predicted, and actual data. While RMSE is the average magnitude of the errors of prediction model. The values of MSE, and RMSE ranges from 0 to ∞. A good predictor model should produce lower values of MSE, and RMSE.

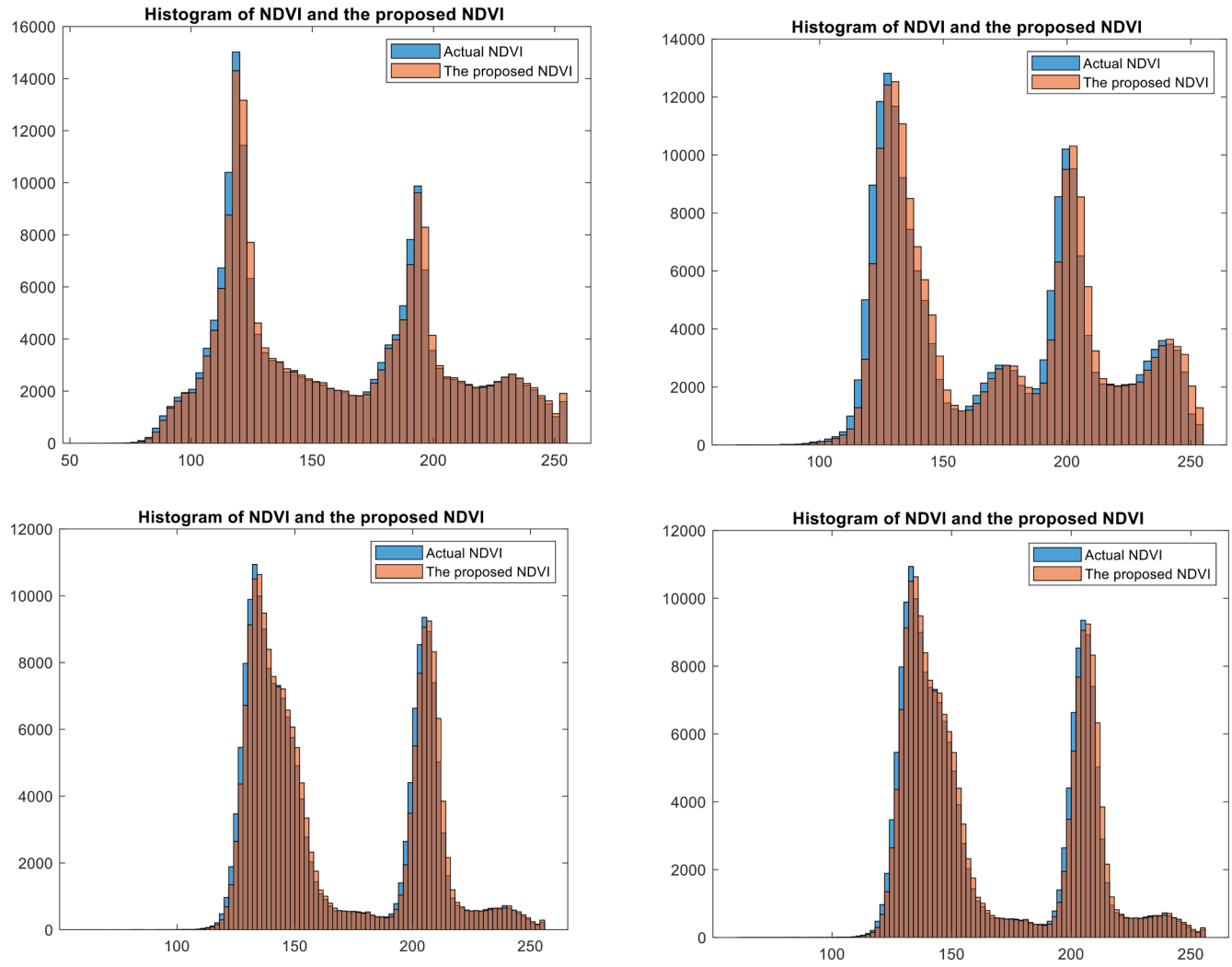


Fig. 5. Histogram comparison between the proposed NDVI and the actual NDVI.

$$PSNR = 20log_{10}(\frac{\max i}{\sqrt{MSE}} \tag{22}$$

$$MSE = \frac{1}{mn}\sum_{0}^{m-1}\sum_{0}^{n-1}||f(i,j) - g(i,j)||^2 \tag{23}$$

where max$i$ is the maximum signal value in the original image, g is the compared image, f is the original image, m, n are the dimension of row and column of images.

4. Experimental results

In this section, we discussed the obtained results in three contexts i. e., evaluation using Quantile-Quantile (Q-Q) plot, comparison with other models, and evaluation based on SSIM, PSNR, and MSE. The proposed NDVI was designed using the RGB channels obtained from the camera as an input to the proposed model based empirical curvelet coupled with deep learning along with comparing models. To evaluate the proposed model, several metrics were used including accuracy, sensitivity, specificity, and three quantitative metrics named structural similarity (SSIM), peak signal to noise ratio (PSNR), mean square error (MSE), Precision, f1-score and Recall were employed in this study.

4.1. Model evaluation using Quantile-Quantile (Q-Q) plot

The Q-Q plot is a graphic tool which has been designed to validate two models according to their distributions, and behaviors. In this experiment, we evaluate the degree of similarity between the proposed NDVI model with the actual NDVI based on color distributions. The quantiles of the prosed NDVI were plotted against the quantiles of actual NDVI. The quantiles indicated to the number of pixels that lie below a given value. Based on Q-Q plot hypothesis, a total of 30% of the data-points should be below the given value, however, the rest of the data-points should remain above that value. A 45-degree reference line was mapped. If two models have the same behaviors and distribution, all points should be fall along the reference line. Fig. 8 shows the Q-Q plot of the proposed NDVI model and the actual NDVI.

4.2. Performance comparison with other models

The proposed DenseNet was compared with convolution neural networks (CNN), ResNet, simple neural networks (NN). In this experiment, the output of the proposed model from Eq. (11) was fed to the proposed DenseNet model as well as to the CNN, simple NN, and ResNet. The comparisons were carried out in terms of accuracy, precision, recall, and F1-score. Table 2 displays the obtained results. The results demonstrated that the proposed DenseNet outperform the other models. It recorded the highest accuracy with 97% compared to CNN, simple NN,

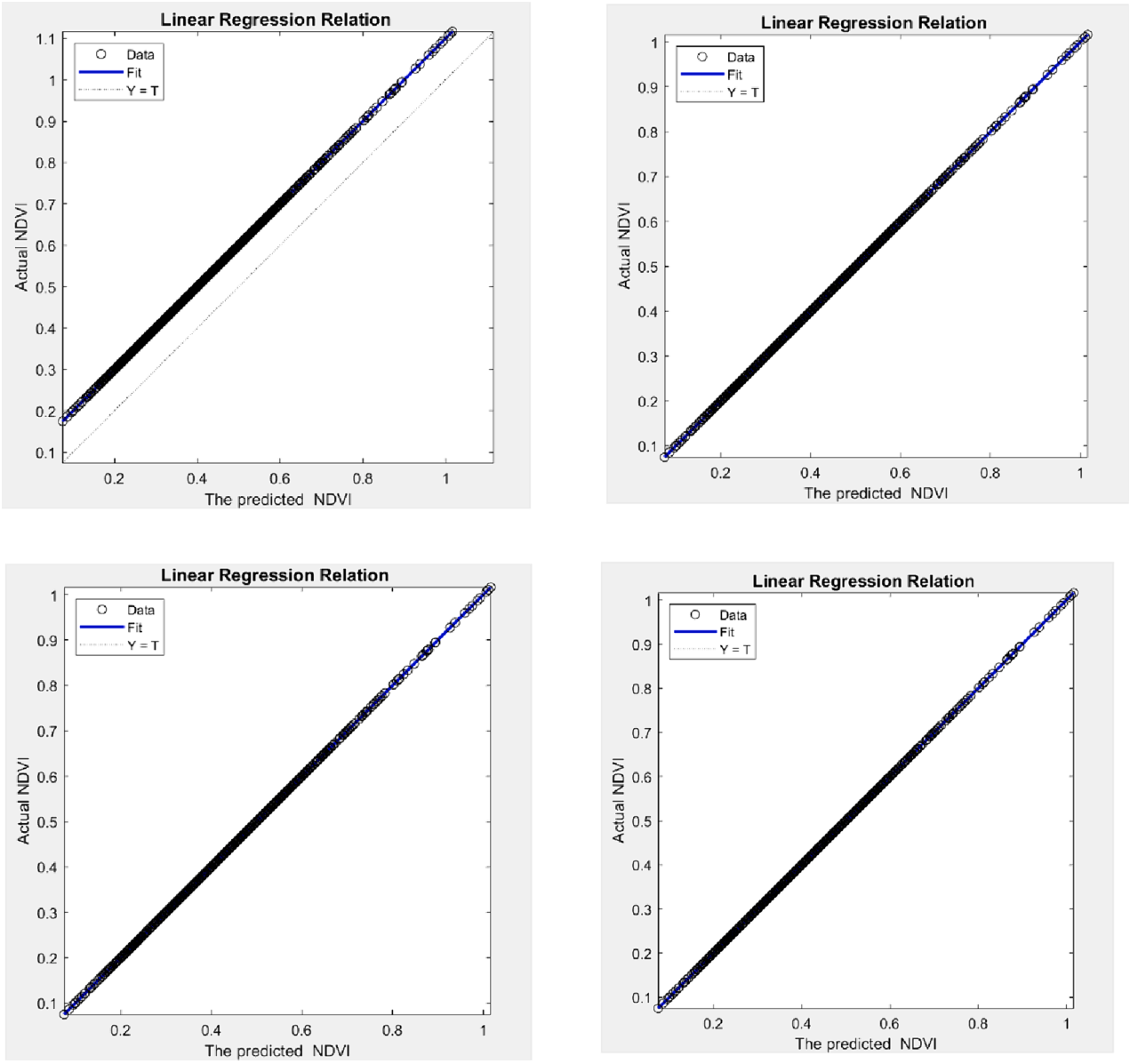


Fig. 6. Regression curve for the proposed NDVI and the actual NDVI.

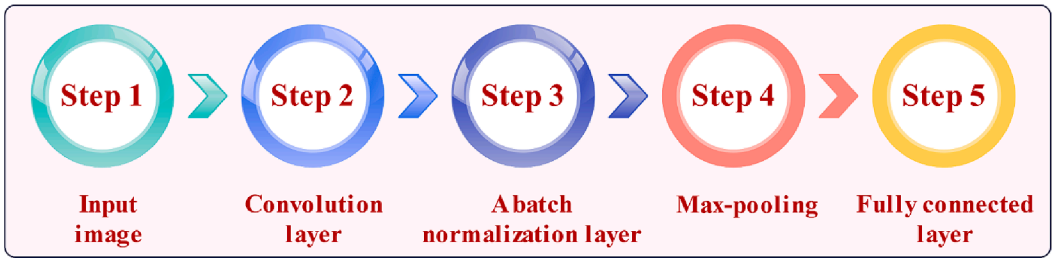


Fig. 7. The architecture of the DenseNet model.

and ResNet. However, the CNN obtained the second highest accuracy. Fig. 9 reports the prediction results.

The error rate for the proposed DenseNet model was also calculated and compared with the CNN, ResNet, and simple NN. Fig. 10 reports the error rate for DenseNet and the other models. The error rate of DenseNet was 0.02, Simple NN is 0.4, CNN is 0.3, ResNet is 0.4.

4.3. Model evaluation based on SSIM, PSNR, and MSE

To evaluate the performance of the proposed model, three quantitative metrics named structural similarity (SSIM), peak signal to noise ratio (PSNR) and mean square error (MSE) were employed in this study. We calculated the SSIM, PSNR, and MSE for each pair of real NDVI and the predicted NDVI. Fig. 11 shows that the average of SSIM varies 0.96

Table 1
List of hyper-parameters used for each model.

Model	Hyperparameters
DenseNet	Learning rate = 0.004 batch size = 5 dropout rate = 0.2 epoch = 1000 activation function = LeakyReLU
ResNet	Learning rate = 0.005 Batch size = 2 activation function = ReLu
CNN	Learning rate = 0.003 Batch size = 3 Solver = adam
Simple NN	Learning rate = 0.005 Batch size = 5 Solver = adam, activation function = ReLu,

and 0.99, which drop down to 0.90, and 0.89 for three images. Our investigation showed that worse performance was obtained when the proposed model was tested on image that it had not a crop. Fig. 12 shows an example of image in which the proposed model showed a poor performance.

The PSNR metric was also calculated to evaluate the performance of the proposed model. Fig. 13 shows the PSNR for all predicted images and it can be observed that the proposed model showed a high correlation with actual NDVI. However, as mentioned before, the performance of the proposed model decreased when it was tested on some types of images as shown in Fig. 12.

To further quantify how well the proposed performed across the NDVI images, another performance evaluation was carried out using MSE metric. Fig. 14 shows the obtained results by MSE proven that the proposed model predicted most NDVI images correctly. To shed more light on the performance of the proposed model, the proposed DenseNet model was compared with the CNN, simple NN, and ResNet. Table 3 reports the prediction results in terms of the SSIM, PSNR, and MSE. From the results obtained, we can notice that the DenseNet outperforms the other models, and it scored the highest values of the SSIM, PSNR, and MSE. However, the simple NN recorded the lowest average of the SSIM,

Table 2
Prediction results of the proposed and comparing models.

Model	Accuracy	Precision	f1-score	Recall
DenseNet	97 %	95 %	95 %	96 %
CNN	88 %	87 %	87 %	86 %
ResNet	86 %	85 %	86 %	85 %
Simple NN	84 %	82 %	83 %	82 %

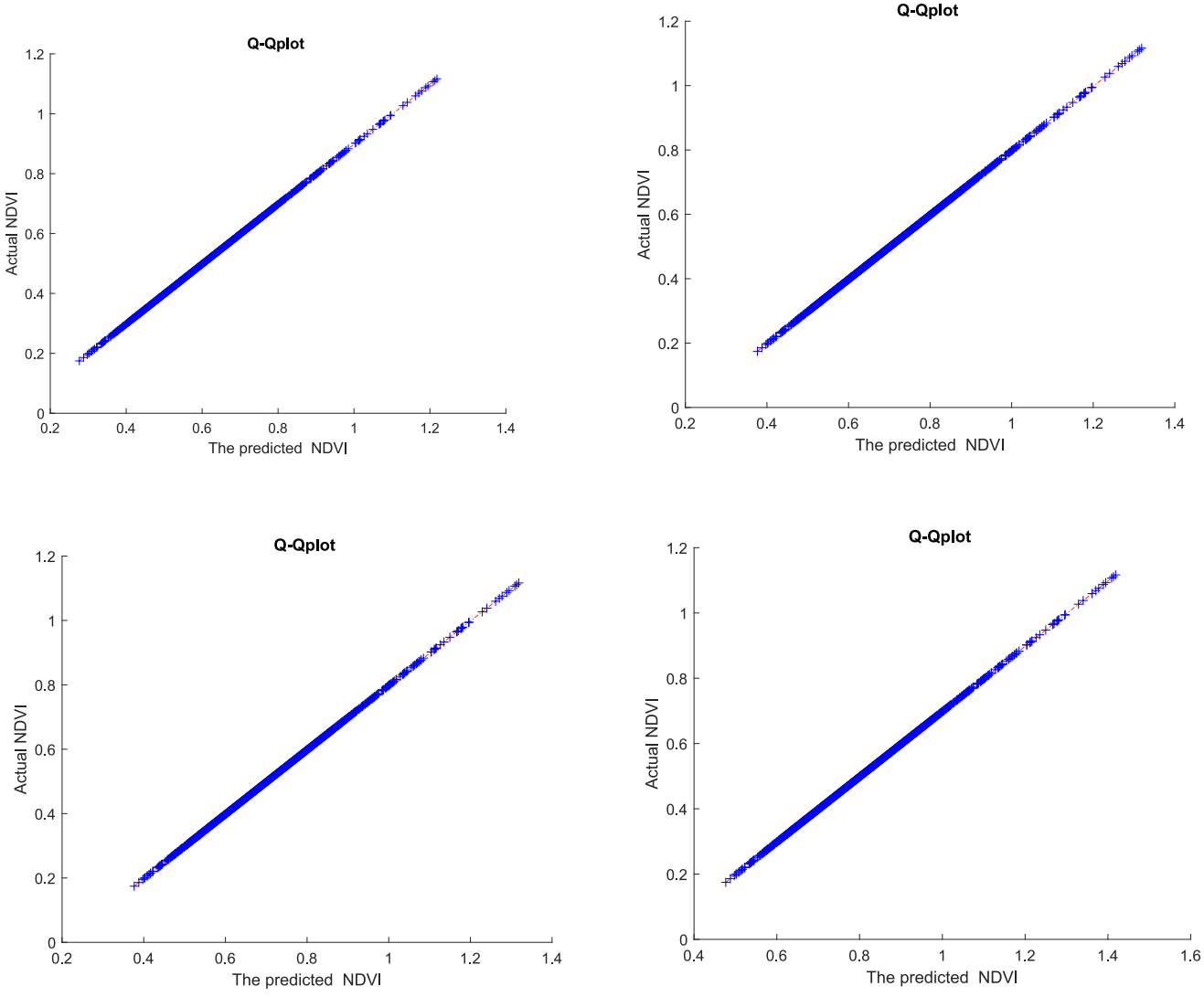


Fig. 8. Q-Q plot curve for the proposed NDVI and the actual NDVI.

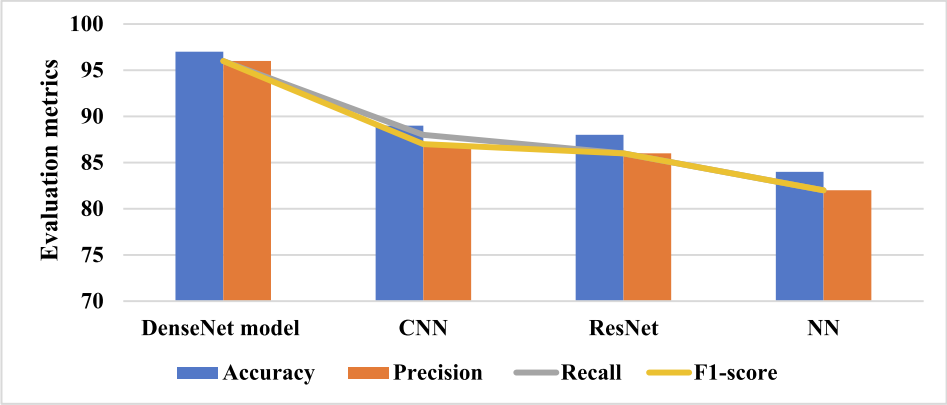


Fig. 9. Comparisons among the DenseNet and other models.

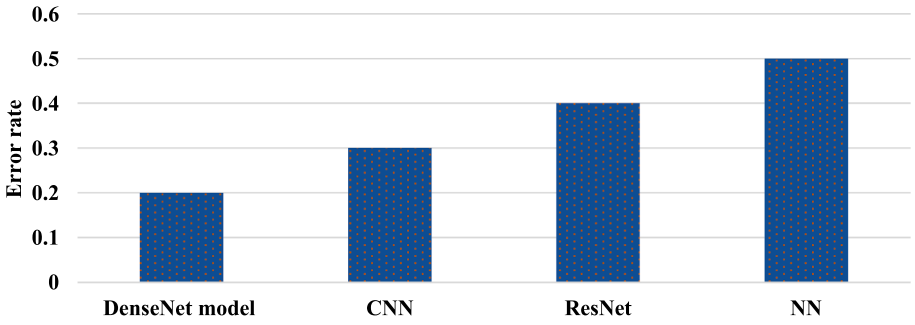


Fig. 10. Error rate of the DenseNet model and other models.

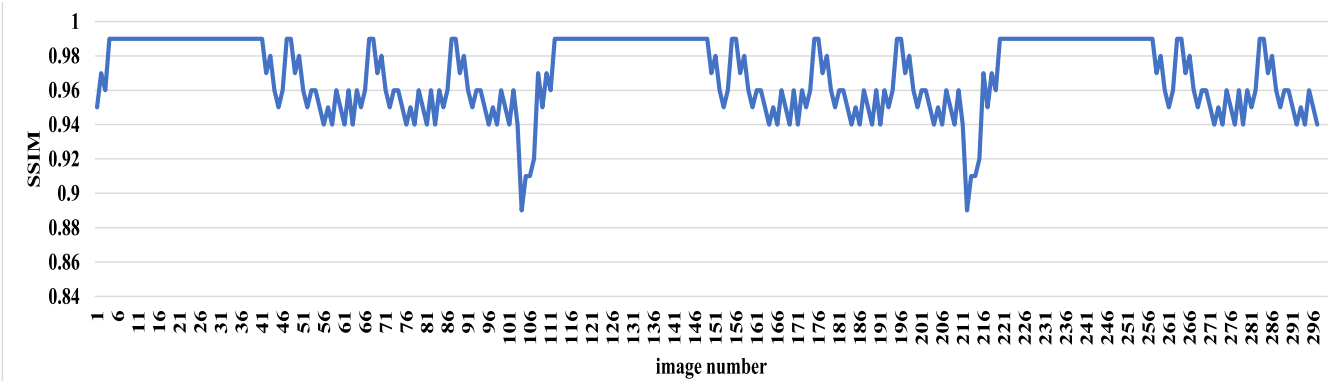


Fig. 11. SSIM results of all RGB inputs.

PSNR, and MSE.

5. Discussions

In this paper, we proposed NDVI prediction model based on empirical curvelets transform method and deep DenseNet model using the RGB drone imagery data form Prince Edward Island, Canada. In this section, the main findings are summarized as follows:

1. Although the model presented in this study provided a new perspective to predict and monitor crops that could be useful to identify issues related to plant diseases to assist forest managers and framers, the proposed model only tested with small dataset. In future research, we aim to develop more efficient solutions to generalize the proposed model, a large dataset is required with different crops to evaluate the effectiveness of the proposed model. In addition, further improving the overall efficiency of the framework will also be made by testing more texture features.
2. In this model, only six deep learning models were coupled with the empirical curvelets transform method to predict NDVI. One of our future works is to test other models such as graph based deep learning, LSTM. In addition, a further investigation will be made to find more accurate models that can predict NDVI with a high prediction rate.
3. To illustrate the efficiency of empirical curvelets transform technique in NDVI prediction. A new experiment was conducted in which the images were forwarded to various models with and without involving curvelet coefficients technique. Fig. 15 reports the obtained results in term of accuracy. It was observed that the prediction rate was declined for all models when the RGB images were forwarded directly to the prediction models. The results approved the ability of the proposed formula to correctly transfer RGB images into

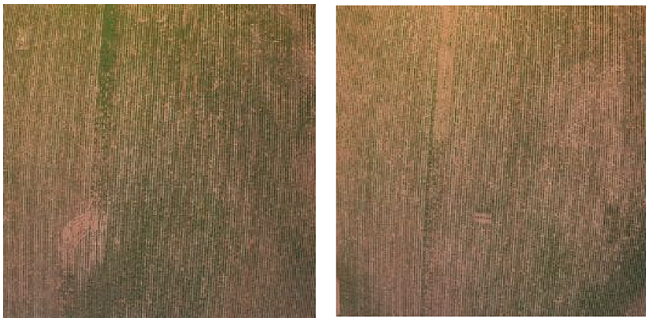


Fig. 12. Example of images with low performance.

NDVI map and the prediction rate was notably increased as show in Fig. 15.

4. In this experiment, the performance of the proposed model for NDVI prediction was evaluated in term of ROC. The specificity was plotted against sensitivity to calculate the ROC curves. Fig. 16 shows the ROC curves of the DenseNet model, CNN, simple NN, and ResNet. The performance of the four models were evaluated was in term of area under the curve (AUCI). The obtained results showed that the DenseNet model outperforms the other model, and it was scored AUCI of 0.96. however, the CNN model recorded AUC1 of 0.90, while the simple NN obtained the lowest AUC1.
5. The correlation coefficient (CC) metric was also utilised to assess the proposed model for NDVI prediction. The correlation coefficient of the actual and predicted NDVI was computed for the DenseNet model, CNN, simple NN, and ResNet. Fig. 17 reports the correlation coefficient over different sample. In this experiment. The average of correlation coefficient of different sample was computed and presented in Fig. 16.
6. The Willmott's Index (d) was also used to evaluate the proposed model. The index calculates the degree of model prediction error which ranges between 0 and 1. The index of agreement indicates the ratio of the mean square error and the potential error. The d value of 1 indicates a perfect match, while 0 indicates no agreement at all. The Willmott's Index detects additive and proportional differences in the observed and simulated means and variances. The results of Willmott's Index in Fig. 18 showed that the proposed model offers substantial benefits in NDVI prediction. The Willmott's Index was calculated for DenseNet model, CNN, simple NN, and ResNet.

Table 3
Performance evaluation in terms of SSIM, PSNR, MSE.

Model	SSIM	PSNR	MSE
DenseNet	**0.98**	**30**	**120**
CNN	0.88	25	370
ResNet	0.79	20	270
Simple NN	0.85	23	400

The NDVI uses specific wavelengths to examine crop health and performance. The proposd model can be implemented to increasie the accessibility of spectral imaging systems through the development of small, low cost, and easy to use platforms for precision agriculture.

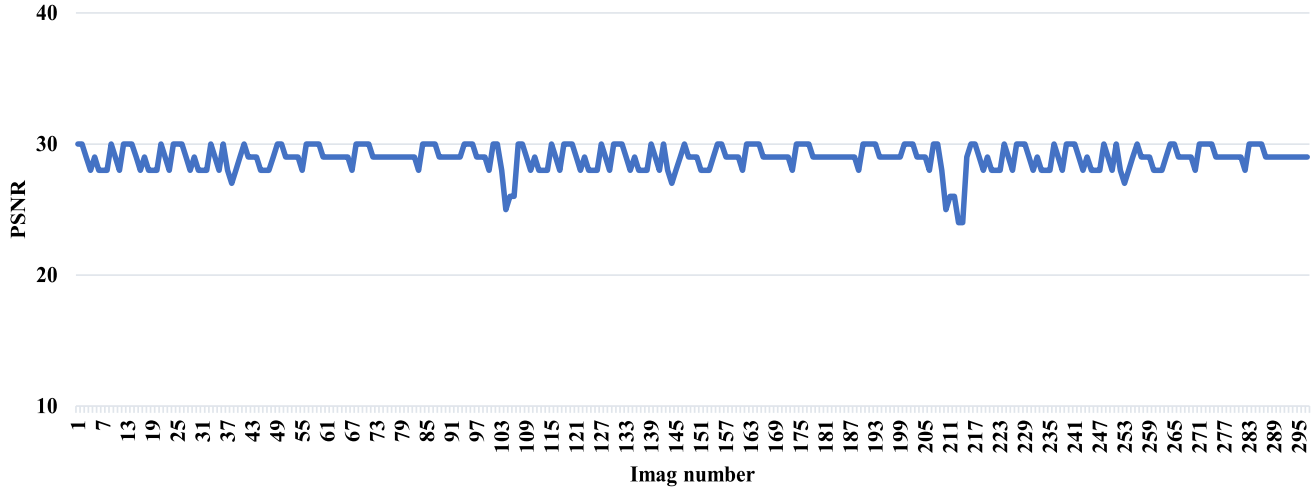


Fig. 13. PSNR results of all RGB inputs.

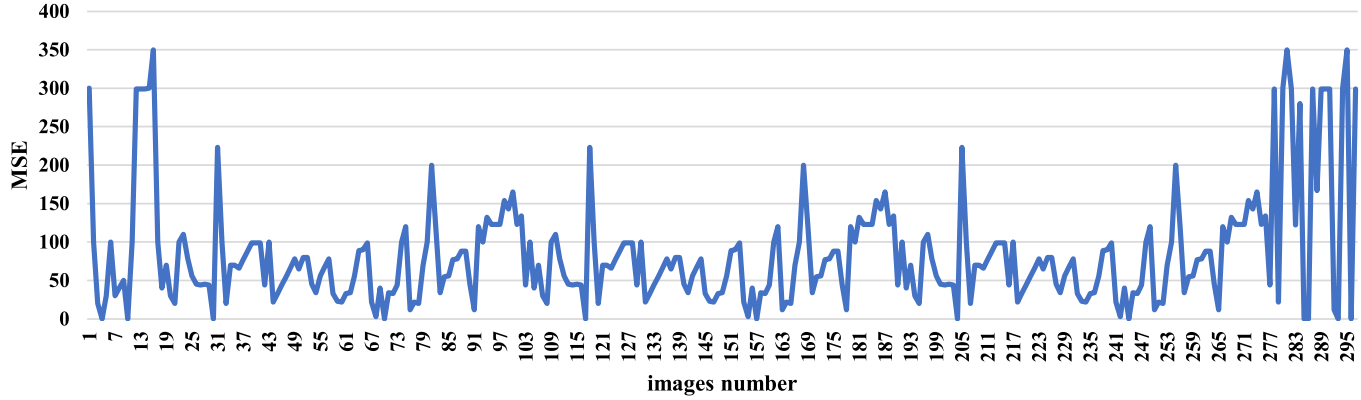


Fig. 14. MSE results of all RGB inputs.

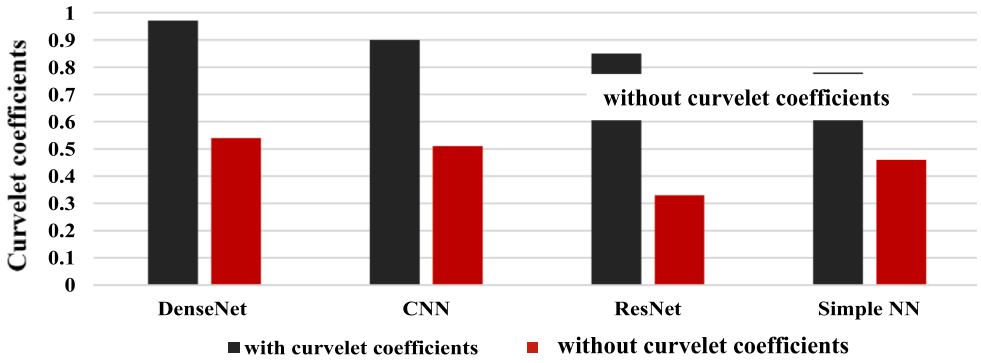


Fig. 15. The influence of curvelet coefficients on prediction rate.

6. Conclusion

This paper proposed a novel empirical curvelet transform based DenseNet model to predict NDVI using RGB drone imagery data. In the proposed model, each channel of drone RGB imagery data was passed through empirical curvelet transform method where the curvelet coefficients were analysed and assessed. As a result, the curvelet coefficients were used to generate a new formula to design NDVI. Finally, the output of the new formula was sent to DenseNet model to predict the final NDVI. The proposed model is compared against CNN, simple NN, and ResNet models using several goodness-of-fit metrics and diagnostic plots to predict NDVI.

The results showed that there was a strong relationship between the actual NDVI and the new formula. According to the goodness-of-fit metrics, the proposed model has potential to generate NDVI from RGB images with high SSIM and PSNR values, and lower MSE. A high QQ plot and correlation were obtained for the DenseNet model as compared to other benchmarking models.

The proposed model can be updated when large region prediction is required. Moreover, the curvelet coefficients can be further examined using graph theory concept to find the relationship among image channels. The scope of this work can be enhanced by employing and validating the proposed model in other areas such as agriculture, climate change, hydrology, environment, and renewable sectors for better decision making. In the future research, other regions can be considered to provide a standard NDVI prediction model. It is believed that the proposed model can provide important references for both farmers and agricultural managers in the region to monitor crops and their health. In addition, the proposed model can reduce the high cost associated with generating an NDVI from multispectral sensors.

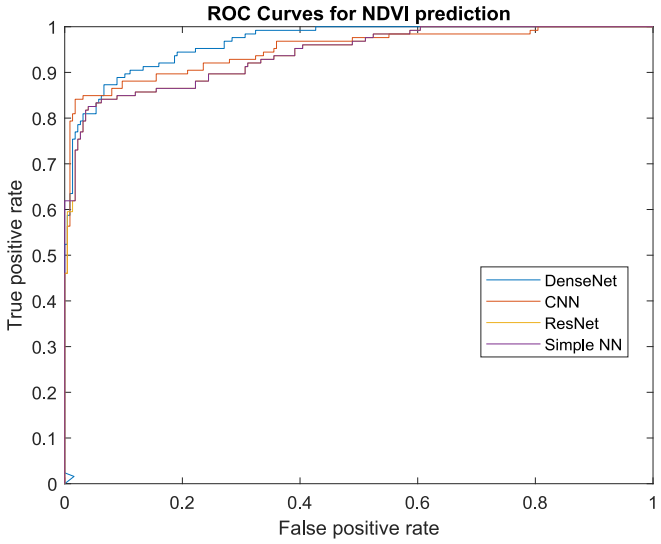


Fig. 16. ROC of the proposed model for NDVI predication.

CRediT authorship contribution statement

Mohammed Diykh: Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Conceptualization. **Mumtaz Ali:** Writing – review & editing, Supervision, Methodology, Investigation,

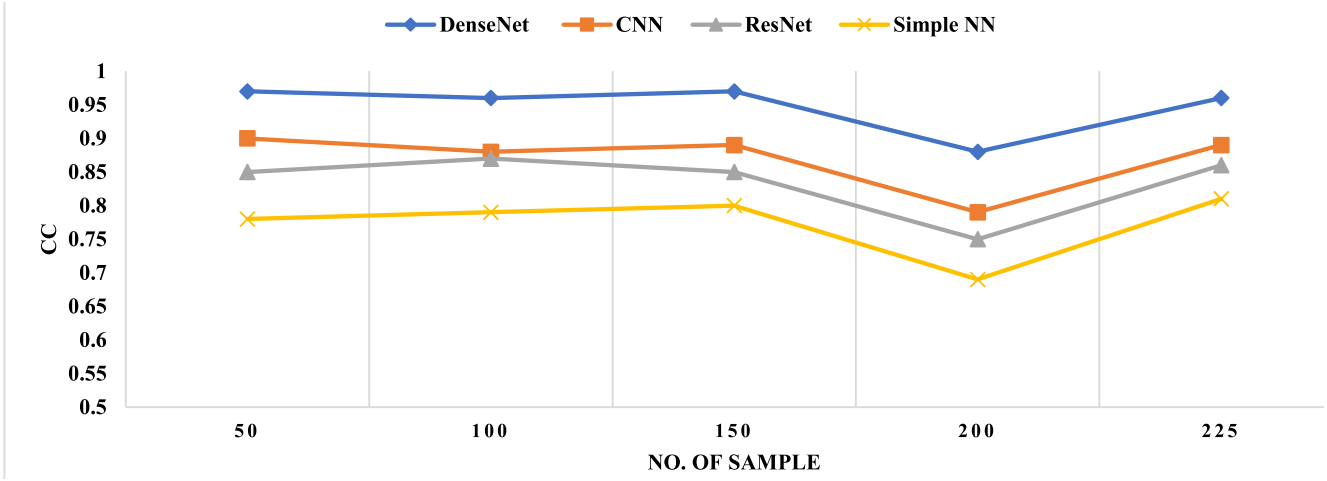


Fig. 17. Correlation coefficient of the proposed model for NDVI predication.

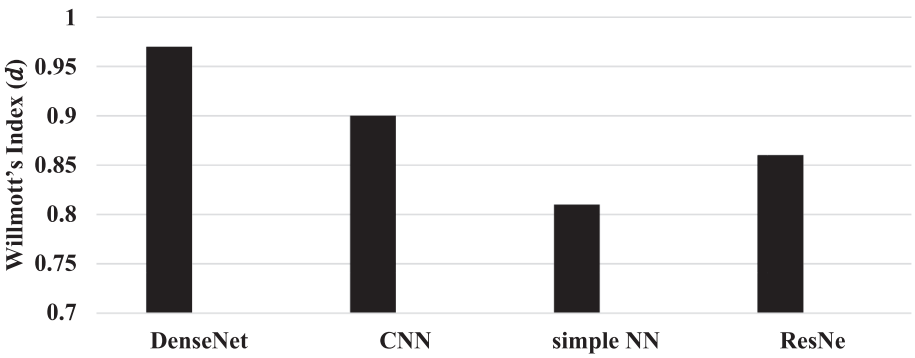


Fig. 18. Willmott's Index (d) agreement between the observed and predicted NDVI.

Conceptualization. **Mehdi Jamei:** Writing – review & editing, Visualization. **Shahab Abdulla:** Supervision, Writing – review & editing. **Md Palash Uddin:** Writing – original draft, Methodology, Data curation. **Aitazaz Ahsan Farooque:** Funding, Data Collection, Conceptulization, Methodology, Supervision, Writing – review & editing. **Abdulhaleem H. Labban:** Writing – review & editing, Supervision. **Hussein Alabdally:** Writing – original draft, Methodology, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors would like to thank Natural Sciences and Engineering Council of Canada and Atlantic Canada Opportunties Agency for financial support to conduct this research. A special thanks to the Precision Agriculture Research Team at the University of Prince Edward Island for their assistance during data collection.

References

Abai, Z. and N. Rajmalwar (2019). "Densenet models for tiny imagenet classification." arXiv preprint arXiv:1904.10429.

Aggarwal, A., Mittal, M., Battineni, G., 2021. Generative adversarial network: an overview of theory and applications. International Journal of Information Management Data Insights 1 (1), 100004.

Albahli, S., Nazir, T., Nawaz, M., Irtaza, A., 2023. An improved DenseNet model for prediction of stock market using stock technical indicators. Expert Syst. Appl., 120903

Barth, R., Hemming, J., Van Henten, E.J., 2020. Optimising realism of synthetic images using cycle generative adversarial networks for improved part segmentation. Comput. Electron. Agric. 173, 105378.

Beyer, M., Ahmad, R., Yang, B., Rodríguez-Bocca, P., 2023. Deep spatial-temporal graph modeling for efficient NDVI forecasting. Smart Agricultural Technology 4, 100172.

Bui, K., Fauman, J., Kes, D., Torres Mandiola, L., Ciomaga, A., Salazar, R., Bertozzi, A.L., Gilles, J., Goronzy, D.P., Guttentag, A.I., 2020. Segmentation of scanning tunneling microscopy images using variational methods and empirical wavelets. Pattern Anal. Appl. 23, 625–651.

Development, R. S. A. I. a. R. (2023). Remote Sensing Applications: Infrastructure and Resource Development.

Devi, G., B. Ganasri G. Dwarakish (2015). Applications of remote sensing in satellite oceanography: A review. Aquatic Procedia, 4: 579–584. International Conference on Water Resources, Coastal and Ocean Engineering (ICWRCOE'15).

Drees, L., Junker-Frohn, L.V., Kierdorf, J., Roscher, R., 2021. Temporal prediction and evaluation of Brassica growth in the field using conditional generative adversarial networks. Comput. Electron. Agric. 190, 106415.

Eguchi, R.T., Huyck, C.K., Ghosh, S., Adams, B.J., 2008. The application of remote sensing technologies for disaster management. The 14th World Conference on Earthquake Engineering.

Emami, H., Aliabadi, M.M., Dong, M., Chinnam, R.B., 2020. Spa-gan: Spatial attention gan for image-to-image translation. IEEE Trans. Multimedia 23, 391–401.

Gan, Y., Shi, J.-C., He, W.-M., Sun, F.-J., 2021. Parallel classification model of arrhythmia based on DenseNet-BiLSTM. Biocybernetics and Biomedical Engineering 41 (4), 1548–1560.

Gao, X., Chen, T., Niu, R., Plaza, A., 2021. Recognition and mapping of landslide using a fully convolutional DenseNet and influencing factors. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 7881–7894.

Giles, M. (2018). The GANfather: The man who's given machines the gift of imagination, TECHNOL REV 1 MAIN ST, 13 FLR, CAMBRIDGE, MA 02142 USA. **121:** 48-53.

Gilles, J., 2013. Empirical wavelet transform. IEEE Trans. Signal Process. 61 (16), 3999–4010.

Girdhar, N., Sinha, A., Gupta, S., 2023. DenseNet-II: An improved deep convolutional neural network for melanoma cancer detection. Soft. Comput. 27 (18), 13285–13304.

Guo, Y., Zhang, L., He, Y., Cao, S., Li, H., Ran, L., Ding, Y., Filonchyk, M., 2024. LSTM time series NDVI prediction method incorporating climate elements: a case study of Yellow River Basin. China. Journal of Hydrology 629, 130518.

Huang, Y., Zhou, F., Gilles, J., 2019. Empirical curvelet based fully convolutional network for supervised texture image segmentation. Neurocomputing 349, 31–43.

Huynh-Thu, Q., Ghanbari, M., 2012. The accuracy of PSNR in predicting video quality for different video scenes and frame rates. Telecommun. Syst. 49, 35–48.

Jahmunah, V., Ng, E., Tan, R.-S., Oh, S.L., Acharya, U.R., 2023. Uncertainty quantification in DenseNet model using myocardial infarction ECG signals. Comput. Methods Programs Biomed. 229, 107308.

Li, J., Pei, Y., Zhao, S., Xiao, R., Sang, X., Zhang, C., 2020. A review of remote sensing for environmental monitoring in China. Remote Sens. (basel) 12 (7), 1130.

Li, J., Li, C., Xu, W., Feng, H., Zhao, F., Long, H., Meng, Y., Chen, W., Yang, H., Yang, G., 2022. Fusion of optical and SAR images based on deep learning to reconstruct vegetation NDVI time series in cloud-prone regions. International Journal of Applied Earth Observation and Geoinformation 112, 102818.

Li, H., Tang, J., 2020. Dairy goat image generation based on improved-self-attention generative adversarial networks. IEEE Access 8, 62448–62457.

Liu, T., Chen, T., Niu, R., Plaza, A., 2021. Landslide detection mapping employing CNN, ResNet, and DenseNet in the three gorges reservoir, China. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 11417–11428.

Maqsood, M.H., Mumtaz, R., Haq, I.U., Shafi, U., Zaidi, S.M.H., Hafeez, M., 2021. Super resolution generative adversarial network (Srgans) for wheat stripe rust classification. Sensors 21 (23), 7903.

Omar, M.S., Kawamukai, H., 2021. Prediction of NDVI using the Holt-Winters model in high and low vegetation regions: a case study of East Africa. Scientific African 14, e01020.

Sapkota, B.B., Popescu, S., Rajan, N., Leon, R.G., Reberg-Horton, C., Mirsky, S., Bagavathiannan, M.V., 2022. Use of synthetic images for training a deep learning model for weed detection and biomass estimation in cotton. Sci. Rep. 12 (1), 19580.

Seong, S., Chang, A., Mo, J., Na, S., Ahn, H., Oh, J., Choi, J., 2024. Crop classification in South Korea for multitemporal PlanetScope imagery using SFC-DenseNet-AM. International Journal of Applied Earth Observation and Geoinformation 126, 103619.

Vinukollu, R.K., Wood, E.F., Ferguson, C.R., Fisher, J.B., 2011. Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: evaluation of three process-based approaches. Remote Sens. Environ. 115 (3), 801–823.

Winkler, S., Mohandas, P., 2008. The evolution of video quality measurement: from PSNR to hybrid metrics. IEEE Trans. Broadcast. 54 (3), 660–668.

Yi, Z., Zhang, H., Tan, P., Gong, M., 2017. Dualgan: Unsupervised dual learning for image-to-image translation. Proceedings of the IEEE International Conference on Computer Vision.

Zeng, X., Feng, G., Zhang, X., 2019. Detection of double JPEG compression using modified DenseNet model. Multimed. Tools Appl. 78, 8183–8196.
Zhang, K., Guo, Y., Wang, X., Yuan, J., Ding, Q., 2019. Multiple feature reweight densenet for image classification. IEEE Access 7, 9872–9880.
Zhang, J., Lu, C., Li, X., Kim, H.-J., Wang, J., 2019. A full convolutional network based on DenseNet for remote sensing scene classification. Math. Biosci. Eng. 16 (5), 3345–3367.
Zhang, Y., Wa, S., Zhang, L., Lv, C., 2022. Automatic plant disease detection based on tranvolution detection network with GAN modules using leaf images. Front. Plant Sci. 13, 875693.
Zhang, J., Yang, G., Sun, L., Zhou, C., Zhou, X., Li, Q., Bi, M., Guo, J., 2021. Shrimp egg counting with fully convolutional regression network and generative adversarial network. Aquacultural Engineering 94, 102175.