Satellite Image Classification using

Classification using Deep Learning

Kinshuk Kushagra

*Department of CSE*

*LPU, Phagwara, Punjab*

kinshukkushagra@gmail.com

Rohan Tyagi

*Department of CSE*

*LPU, Phagwara, Punjab*

tyagirohan798@gmail.com

Himmi Raja

*Department of CSE*

*LPU, Phagwara, Punjab*

irhimmi27@gmail.com

Kamalpreet Kaur

*Department of CSE*

*LPU, Phagwara, Punjab*

Kamalpreet.cse98@gmail.com

***Abstract*—Internet of Things (IoT) has significantly transformed various sectors, including agriculture. However, IoT devices in the rapidly advancing field of smart agriculture often face substantial challenges related to energy consumption and network connectivity, especially in large-scale agricultural operations. This problem is further exacerbated by the increasing global population, which is expected to reach 9.7 billion by 2050, and the unpredictability of natural disasters, necessitating the use of smart agriculture solutions that address the high energy consumption and connectivity challenges faced by IoT devices. The research aims to address this gap by focusing on the optimization of energy-efficient connectivity solutions for IoT-based smart agriculture systems. The objectivity is to enhance the energy efficiency and sustainability of smart agriculture through the use of advance technology such as implementing narrowband IoT for large scale agriculture and integrating NOMA with narrowband IoT for decreasing the interference. The further scope of this research includes enhancing the productivity and energy efficiency of agricultural products securely for a sustainable future.**

***Keywords—Internet of Things (IoT), Smart Agriculture, Energy Efficiency, Connectivity Solutions, Sustainability.***

I. INTRODUCTION

In the realm of remote sensing and satellite imagery analysis, cloud detection methods play a pivotal role in discerning and categorizing cloud cover within captured scenes. These methods are indispensable because clouds can obscure the Earth's surface, impacting the accuracy and reliability of data extracted from these images. Whether for weather forecasting, climate monitoring, agricultural planning, or environmental assessment, understanding cloud dynamics is crucial for interpreting remote sensing data effectively.

The necessity for robust cloud detection methodologies stems from the inherent challenges posed by cloud cover in satellite imagery. Clouds can mask important ground features, alter the spectral characteristics of light reflected from the Earth's surface, and introduce variability that complicates image analysis algorithms. Hence, accurate identification and delineation of clouds are essential for ensuring the integrity and usefulness of subsequent analyses and applications.

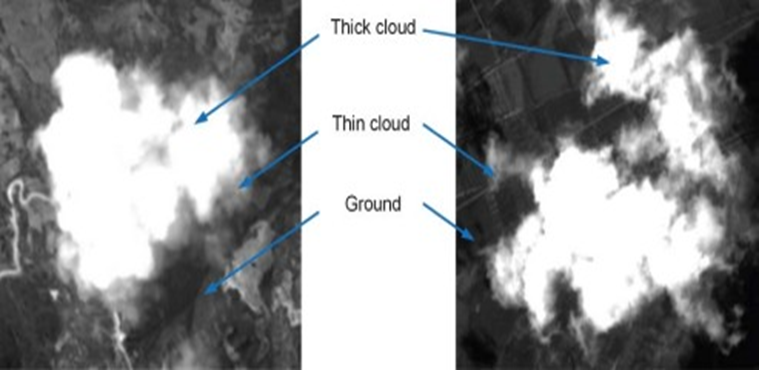


Fig. 1 Cloud detection in a form of thick or thin cloud

In this context, the choice of neural network architecture becomes crucial. Neural networks, especially deep learning models like ResNet-50, VGG-19, Xception, and YOLO, have been employed for various computer vision tasks, including cloud detection. Each architecture brings unique strengths to the table: ResNet-50 with its deep residual connections, VGG-19 with its simplicity and effectiveness in feature extraction, Xception with its efficiency-enhancing depthwise separable convolutions, and YOLO with its real-time object detection capabilities.

Among these architectures, YOLO (You Only Look Once) stands out particularly for cloud detection applications. Unlike traditional architectures that are primarily optimized for image classification or feature extraction, YOLO is designed specifically for real-time object detection. Its ability to swiftly and accurately identify objects across diverse environments makes it well-suited for the dynamic and variable nature of cloud detection in satellite imagery. YOLO processes images in a single forward pass, making it efficient for applications where timely detection and precise delineation of cloud cover are critical.

By leveraging YOLO's strengths in object localization and recognition, cloud detection methods can enhance the speed and accuracy with which clouds are identified and characterized in satellite images. This capability not only improves the reliability of remote sensing data but also enables more informed decision-making in fields reliant on accurate environmental observations.

In essence, cloud detection methods utilizing advanced neural network architectures like YOLO represent a significant advancement in the analysis of satellite imagery. They address the pressing need for rapid and precise identification of clouds, thereby supporting applications that rely on detailed and dependable remote sensing data for scientific, environmental, and operational purposes.

II. RELATED WORK

G. E. Bostanc's study titled "Cloudy/Clear Weather Classification Using Deep Learning Techniques with Cloud Images" (2022) [1] examines the challenges posed by meteorological errors in weather forecasting, despite the increasing accessibility of weather forecasts on smart devices. The study emphasizes the human errors inherent in traditional weather forecasting and proposes the use of deep learning techniques to minimize these errors by classifying large datasets of ground-based cloud images. The research evaluates the performance of four pretrained deep learning models: MobileNet V2, VGG-16, ResNet-152 V2, and DenseNet-201, with VGG-16 achieving the highest accuracy at 91.4%. Similarly, R. Dhir et al. (2022) [2] in their paper "Land Use and Land Cover Classification of Remote Sensing Images Based on Deep Learning Approaches" highlight the significance of deep learning in environmental applications of remote sensing, focusing on land use and land cover (LULC) classification. The study reviews various DL architectures, emphasizing their superiority over traditional methods, and concludes with a precision of 92.75%, recall of 96.08%, and F1-score of 94.39%. S. Ji et al. (2021) [3], in "Simultaneous Cloud Detection and Removal from Bitemporal Remote Sensing Images Using Cascade CNN," introduce a deep learning framework for cloud detection and removal using Landsat 8 datasets. Their method, leveraging VGG16, Deeplab v3+, and U-Net, achieves 96% accuracy by effectively addressing scale-related effects in remote sensing data. Also, N. Zhou et al. (2021) [4] in "Geo-Parcel-Based Change Detection Using Optical and SAR Images in Cloudy Areas" propose a technique for change detection in cloud-affected regions using multisource remote sensing images. Their method, validated on the Gui'an test site, achieves a 94% accuracy rate, demonstrating its efficacy in identifying altered geo-parcels using deep learning models.

Alan Li's February 2020 study[5], "Cloud Detection Algorithms for Multimodal Satellite Imagery Using Convolutional Neural Network," addresses the critical role of cloud detection in remote-sensing applications to optimize data processing by mitigating cloud interference over surfaces of interest. Despite existing cloud detection methodologies, challenges persist, particularly over high albedo surfaces (e.g., snow and sand) and in transferring algorithms between observational platforms. Li presents a CNN algorithm developed within the NASA NeMO-Net project for detecting cloud and cloud shadow fields in multi-channel satellite imagery from World-View-2 (WV-2) and Sentinel-2 (S-2) using RGB and NIR channels. The CNN algorithm, trained on WV-2 data, achieved an 89% accuracy rate for cloud detection on WV-2 and outperformed the original S-2 cloud mask with an 81% prediction rate for clouds. A novel domain adaptation approach further enhances the algorithm's capability to transfer knowledge between satellite platforms, showing promise in increasing the prediction accuracy of both clear and cloudy pixels. Similarly, Seema Mahajan's June 2019 review[6], "Cloud Detection Methodologies: Variants and Development," examines cloud identification techniques from 2004 to 2018, utilizing datasets from various satellites and achieving nearly 90% accuracy through hybrid methods that combine atmospheric parameters and artificial neural networks. Jingyu Yang's August 2019 paper[7], "CNN-Based Cloud Detection for Remote Sensing Imagery With Cloud-Snow Coexistence," introduces CDnetV2, an encoder-decoder neural network with adaptive feature fusing and high-level semantic information guidance, achieving superior accuracy on ZY-3 satellite data. Qingyong Li's May 2012 study[8] "Thin Cloud Detection of All-Sky Images Using Markov Random Fields," addresses the challenge of low contrast and vague boundaries in ground-based sky imaging, achieving an accuracy of 83.4%. In Rachana Gupta and Pradip Panchal's August 2015 paper[9], "Advancement of Cloud Detection Algorithm in Satellite Images With Application to Color Models," explores various color models for cloud detection, achieving a maximum accuracy of 92.4% using the RGB model on VIRR datasets.

Recent advancements in cloud detection methodologies have significantly improved accuracy and robustness across various remote sensing platforms. Li et al. [10] introduced an innovative spectral-spatial classification strategy that combines Threshold Exponential Spectral Angle Map (TESAM), adaptive Markov Random Field (aMRF), and Dynamic Stochastic Resonance (DSR) to enhance onboard cloud detection in hyperspectral images. This method achieved a notable accuracy of 96.28% and outperformed conventional techniques by effectively handling noise and misclassification issues. Arun et al. [11] proposed the Fog Stability Index (FSI) for detecting fog and low clouds using data from AIRS, IASI, and model wind data over the Indo-Gangetic plains. Their technique demonstrated an 84.63% accuracy, proving effective in both day and night conditions and under high clouds. Hayatbini et al. [12] developed a gradient-based cloud-image segmentation algorithm, CDnetV2, which integrates adaptive feature fusion and high-level semantic guidance to improve cloud detection accuracy. This method surpassed existing techniques, offering up to 98% accuracy and significant enhancements over the PERSIANN-CCS system. Changhui et al. [13] proposed a feature extraction-based method using gray scale, frequency, and texture features for cloud detection, achieving a 90% accuracy rate by efficiently distinguishing clouds from ground objects. Frantz et al. [14] refined the Fmask algorithm by introducing the Cloud Displacement Index (CDI), leveraging parallax effects from multi-angular near-infrared bands to distinguish clouds from bright surfaces. Their method improved the overall accuracy to 95%, addressing challenges posed by the absence of a thermal band in Sentinel-2 imagery. Lastly, Song et al. [15] examined cloud detection methods using MODIS multispectral images, achieving up to 95% accuracy by exploiting spectral and spatial correlations. Their work emphasizes the effectiveness of statistical and pattern recognition approaches in leveraging the spectral channels available on MODIS for precise cloud masking.

Sun et al. [16] introduced a refined geometry-based method for detecting cloud shadows in Landsat 8 OLI images by incorporating dynamic cloud height ranges and combining spectral and object-based analyses. This approach achieved an accuracy exceeding 80%, significantly improving upon the 60% accuracy of Fmask and reducing cloud shadow leakage. Ishida et al. [17] proposed a support vector machine (SVM)-based cloud detection method for MODIS, addressing the variability in cloud and surface conditions with adjustable parameters to minimize incorrect results. Their method, which integrates discriminant analysis and feature space adjustments, demonstrated a 90% accuracy, offering robust performance under diverse conditions. Li et al. [18] developed MSCN, a deep learning-based cloud detection method utilizing multi-scale convolutional features. Tested on GaoFen-1 WFV imagery, MSCN achieved an impressive accuracy of 97.85% and excelled in distinguishing clouds from bright surfaces, outperforming traditional methods. Oishi et al. [19] evaluated CLAUDIA3, a new SVM-based cloud discrimination algorithm for GOSAT-2, comparing it with its predecessor CLAUDIA1. The updated algorithm showed a higher accuracy of 89.5% in tropical rainforests, enhancing cloud detection especially in challenging conditions like snow-covered surfaces. Wu and Shi [20] introduced a novel deep learning approach for cloud detection using a multi-level feature extraction technique. Their method, which combines deep convolutional neural network features with a composite image filter, achieved an intersection over union of 85.38% and produced detailed cloud masks, demonstrating significant improvements in handling complex imagery.

Ricciardelli et al. [21] developed the MACSP algorithm for the MSG-SEVIRI data, combining physical, statistical, and temporal approaches. The MACSP algorithm achieved a detection accuracy of 91.8%, surpassing SAFNWC's 89.2%, and demonstrated superior performance in classifying cloudy pixels compared to MODIS cloud masks. Chen et al. [22] addressed the challenges of high-resolution cloud detection by employing multiple convolutional neural networks (MCNNs). Their method, which integrates adaptive segmentation and MCNNs for multiscale feature extraction, achieved high accuracy ranging from 95% to 98%, proving effective in detecting various cloud levels. Deng et al. [23] introduced a novel cloud detection approach based on natural scene statistics and Gabor features, which overcame limitations of existing methods. Their approach, which combines pre-processing, seed extraction, and region-growing steps, demonstrated an overall accuracy of 93%, outperforming traditional CNN and SVM methods. Ghasemian and Akhoondzadeh [24] proposed two Random Forest-based methods, FLFRF and DLFRF, incorporating various spectral and textural features. Their methods achieved high accuracy and kappa values, with FLFRF and DLFRF showing superior performance compared to SVM, KNN, and other machine learning methods. Liu et al. [25] developed a linear combination model for detecting thin clouds, using a tree structure and SVM classifier. Their method, which combines multiple image features and utilizes AdaBoost for feature selection, achieved a high accuracy of 93% and demonstrated significant improvements in computational efficiency.

Savas Ozkan et al. (2018) [26] explore cloud detection from RGB color remote sensing images using Deep Pyramid Networks (DPNs). Their study addresses the challenge of detecting clouds in RGB images, which lack distinct spectral patterns for clouds. The DPN model, enhanced with pre-trained parameters at the encoder layer, was tested on 20 images from RASAT and Gokturk-2 satellites. This method achieved an impressive accuracy of 98%, effectively handling difficult cases such as snowy mountains by accurately segmenting and classifying pixels. M. Reguiegue et al. (2017) [27] propose a cloud detection approach using fuzzy logic and neural networks for MSG SEVIRI images. Their methods utilize spatial and temporal

properties from multiple spectral channels, including solar and thermal infrared bands. The fuzzy logic method achieved an accuracy of 84.41%, while the neural network approach significantly outperformed with an accuracy of 99.69%. This study demonstrates the effectiveness of AI techniques in distinguishing between thick, thin, and less bright clouds, showcasing the neural network's robustness in various conditions.

Wang et al. in November 2018 [28] introduce an object-based Convolutional Neural Network (CNN) method for detecting clouds and snow in high-resolution multispectral images. Given the challenge of differentiating clouds from snow, especially in the absence of shortwave infrared bands, their CNN model learns multiscale semantic features and is combined with a linear iterative clustering algorithm to create superpixels. This approach achieved a detection accuracy of nearly 92%, outperforming traditional methods and improving precision in separating cloud and snow. In “Object-Based Convolutional Neural Networks for Cloud and Snow Detection in High-Resolution Multispectral Imagers”[29], authors focus on developing a cloud detection algorithm for Proba-V satellite imagery using a supervised pixel-based classification method. Their approach, leveraging statistical machine learning techniques, was validated on a large number of Proba-V images. The algorithm demonstrated an accuracy of 93%, highlighting its

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **S. No.** | **Author/Year** | **Title of the Paper** | **Method** | **Performance** | **Advantage** |
| 1. | E. Ricciardelli et al., 2008 | Physical and statistical approaches for cloud identification using MSG-SEVIRI data | Physical, statistical, temporal approaches; K-Nearest Neighbour classifier | 91.8% accuracy (MODIS) | Combines physical and statistical methods; high accuracy in cloud detection |
| 2. | Yang Chen et al., 2018 | Multilevel Cloud Detection for High-Resolution Remote Sensing Imagery Using MCNNs | Multiple Convolutional Neural Networks (MCNNs); A-SCLI for superpixel segmentation | 95%-98% accuracy | High accuracy; effective in detecting multilevel clouds in high-resolution imagery |
| 3. | Chenwei Deng et al., 2019 | Cloud Detection in Satellite Images Based on Natural Scene Statistics and Gabor Features | Natural Scene Statistics model; Gabor features; Support Vector Machine (SVM) | 93% accuracy | Outperforms existing methods; good at detecting small and thin clouds |
| 4. | Nafiseh Ghasemian & Mehdi Akhoondzadeh, 2018 | Introducing two Random Forest based methods for cloud detection in remote sensing images | Feature Level Fusion Random Forest (FLFRF); Decision Level Fusion Random Forest (DLFRF) | 83% accuracy (MODIS) | Higher accuracy for cloud and snow/ice detection; better than other machine learning methods |
| 5. | L. Liu et al., 2018 | Thin Cloud Detection Method by Linear Combination Model of Cloud Image | Linear combination model; AdaBoost Classifier; SVM classifier | 93% accuracy | Fast computation; high accuracy; effectively separates cloud information from surface |
| 6. | Savas Ozkan et al., 2018 | Cloud Detection from RGB Color Remote Sensing Images with Deep Pyramid Networks | Deep Pyramid Network (DPN); pre-trained model; pixel-level segmentation | 98% accuracy | Excellent performance on challenging cases like snowy mountains |
| 7. | Mourad Reguiegue & Fatima Chouireb, 2017 | Automatic day time cloud detection over land and sea from MSG SEVIRI images | Fuzzy logic; Neural network; uses multiple spectral channels | 99.69% accuracy (NN) | High accuracy; detects various cloud types including thin and less bright clouds |
| 8. | Lei Wang et al., 2018 | Object-Based Convolutional Neural Networks for Cloud and Snow Detection | Object-based CNN; extended linear iterative clustering algorithm for superpixels | 92% accuracy | Effective separation of cloud and snow; improves precision in high-resolution images |
| 9. | Luis Gomez-Chova et al., 2017 | Cloud Detection Machine Learning Algorithms for Proba-V | Supervised pixel-based classification; statistical machine learning techniques | 93% accuracy | Accurate cloud detection; reduces errors in land and sea cover parameter retrieval |
| 10. | Zhengsheng Guo et al., 2018 | A Cloud Boundary Detection Scheme Combined with ASLIC and CNN | Convolutional Neural Network (CNN); Adaptive Simple Linear Iterative Clustering (ASLIC) | >94% accuracy | Robust detection of thin and thick clouds; effective in diverse imaging platforms |

capability to accurately detect clouds and reduce errors in sea and land cover biophysical parameter retrieval. Z. Guo et al. [30] in their paper proposed a cloud boundary detection scheme combining Adaptive Simple Linear Iterative Clustering (ASLIC) with Convolutional Neural Networks (CNNs). Tested on GF-1/2 and ZY-3 satellite imagery, their method increases average detection accuracy by over 5% compared to traditional methods, achieving more than 94% overall accuracy. This approach efficiently detects both thin and thick clouds and delineates cloud boundaries, although it faces challenges in detecting thin clouds at the edges of thicker clouds.

III. PROPOSED METHODOLOGY

The workflow of a deep learning algorithm starts with data collection and preprocessing, which includes cleaning, resizing, normalization, and augmentation to prepare the data for training. The processed data is then split into training, validation, and test sets. The model architecture is designed and trained using the training set, during which hyperparameters are tuned. The model's performance is evaluated on the validation set to adjust parameters and avoid overfitting. Finally, the model is tested on the unseen test set to assess its generalization ability. After testing, the model is refined if necessary and deployed for practical use.

*A. Dataset*

The dataset used is the RSI-CB256 Satellite Image Classification Dataset [16], which combines photos from Google Maps and sensors to create 4 different categories. The collection includes 5631 photographs in total, 1500 of which are cloudy images of cloud areas, 1131 of which are desert images of desert places, 1500 of which are green area images of green areas, and 1500 of which are water images of water locations. The number of cloudy and clear images that display the cloudy and clear areas in satellite images are available online. The data on the desert, green area, and water area satellite photos are therefore readily available, whereas the authors have published a relatively large collection of hazy satellite images in this work. The authors built a Kaggle database so that this data is freely accessible for researchers to use, in order for this work to be useful to others.

*B. Image pre-processing*

**1. Resizing:**

* **Uniform Size:** Images are typically resized to a uniform size, often 299x299 pixels, which is the default input size for the Xception model and for VGG-19 Images are resized to a fixed size of 224x224 pixels, which is the standard input size for VGG-19 as well as ResNet 50 models.But for yolo-v8 Images are typically resized to a fixed size (e.g., 640x640 pixels), ensuring consistency across all input images. This helps in maintaining uniformity for the model input.

**2. Scaling:**

* **Normalization:** Pixel values are scaled to a range of [0, 1] by dividing by 255. This normalization helps in speeding up the convergence of the training process. Alternatively, pixel values can be standardized to have zero mean and unit variance. For VGG-19 Subtract the mean RGB values of the ImageNet dataset: [103.939, 116.779, 123.68]. This centers the data around zero. Convert the image from RGB to BGR format, as VGG-19 was trained using BGR images.For ResNet 50 Normalize pixel values by subtracting the mean RGB values of the ImageNet dataset: [0.485, 0.456, 0.406] and then Divide by the standard deviation of the ImageNet dataset: [0.229, 0.224, 0.225].In yolo-v8 Pixel values are scaled to the range [0, 1] by dividing by 255.

**3. Centering:**

* **Mean Subtraction:** For xception The mean of the pixel values in the dataset can be subtracted from each image to center the data. This step can be done per channel (R, G, B).

**4. Data Augmentation:**

* **Random Cropping:** Randomly cropping the image to 299x299 pixels to introduce variability in the training data.
* **Horizontal Flipping:** Random horizontal flipping of images to augment the dataset.
* **Rotation:** Random rotation of images by a few degrees.
* **Zooming:** Random zooming in and out of images.
* **Shifting:** Random shifting of the image in height and width.
* **Shearing:** Random shearing of the image.
* Data augmentation helps in making the model more robust to variations and prevents overfitting.

**5. Colour Jittering:**

* Adjusting the brightness, contrast, saturation, and hue of images to improve the model's robustness to lighting variations.

**6. Normalization:**

* Ensuring that the data is normalized to the same distribution as the data the model was originally trained on, typically using the ImageNet dataset mean and standard deviation.

*C*. Classification

**ReNet 50 :** ResNet-50 is a prominent variant of the ResNet (Residual Network) architecture, introduced by Microsoft Research in 2015. It features 50 layers and is specifically designed to address the challenges of training very deep neural networks. The key innovation of ResNet-50 is its use of residual connections, which allow the network to bypass one or more convolutional layers, facilitating the training process by mitigating the vanishing gradient problem. These residual blocks consist of convolutional layers with skip connections that add the input to the output, enabling the model to learn residual mappings and improve performance. ResNet-50 achieves a balance between depth and computational efficiency, making it a popular choice for various computer vision tasks, including image classification, object detection, and feature extraction. Its design enables high accuracy and robust performance on benchmarks such as ImageNet, and it has become a foundational model for many advanced applications in the field of deep learning.

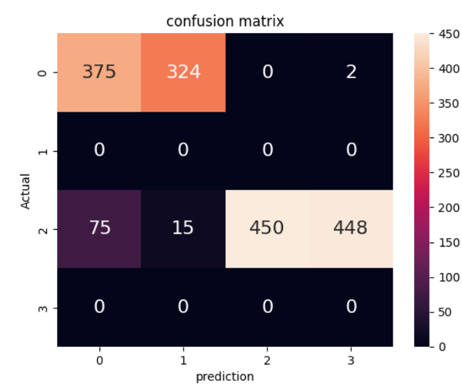


Fig. A. Confusion matrix

A graph with red lines

Description automatically generated

Fig. B. Training and Validation Loss

A graph of a graph

Description automatically generated

Fig. C. Training and Validation Accuracy

A screenshot of a computer program

Description automatically generated

Fig. D. Precision, recall, f1-score

**VGG-19:** VGG-19, a prominent convolutional neural network architecture, was introduced by the Visual Geometry Group at the University of Oxford in 2014. It is a deep network known for its simplicity and effectiveness, built with 19 layers that include 16 convolutional layers and 3 fully connected layers. VGG-19 is characterized by its use of small 3x3 convolutional filters and 2x2 max-pooling layers, which contribute to its ability to capture fine-grained details and hierarchical features in images. The network's architecture follows a straightforward and consistent design: it stacks convolutional layers with increasing depth, followed by max-pooling layers that reduce spatial dimensions while preserving important features. VGG-19 is renowned for its high performance on image classification tasks, such as those in the ImageNet competition, where it demonstrated its capability to achieve state-of-the-art results. Its architecture has become a foundational model for many applications in computer vision, including feature extraction and transfer learning, due to its deep and rich feature representations.

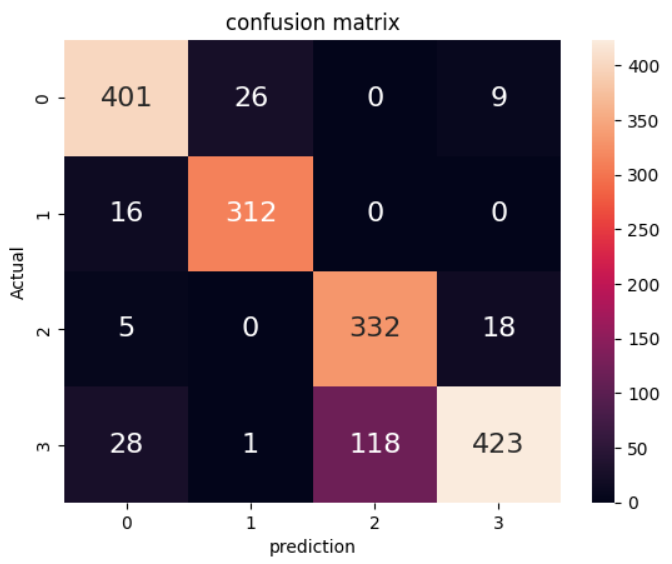


Fig. A. Confusion Matrix

*A graph of a graph

Description automatically generated with medium confidence*

Fig. B. Training and Validation Loss

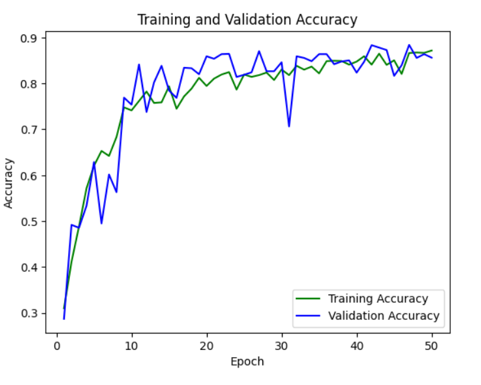
**

Fig. C. Training and Validation Accuracy

A screenshot of a computer screen

Description automatically generated

Fig. D. Precision, recall, f1-score

**XCEPTION:** Xception, or Extreme Inception, is an advanced convolutional neural network architecture introduced by François Chollet in 2017. It builds upon the principles of the Inception modules but replaces standard convolutions with depthwise separable convolutions. This approach divides convolution operations into two stages: depthwise convolutions, which apply a single filter per input channel, and pointwise convolutions, which use 1x1 convolutions to combine the output channels. This method significantly reduces the computational complexity and number of parameters, making Xception both efficient and powerful. The architecture is organized into three main components: the Entry Flow, which reduces the spatial dimensions and increases the depth; the Middle Flow, which performs feature extraction with several depthwise separable convolution blocks; and the Exit Flow, which prepares the final feature maps for classification. Xception utilizes residual connections to facilitate the training of deep networks and has achieved impressive performance on image classification tasks, such as those in the ImageNet dataset. Its efficiency and effectiveness make it suitable for various applications, including object detection and semantic segmentation.

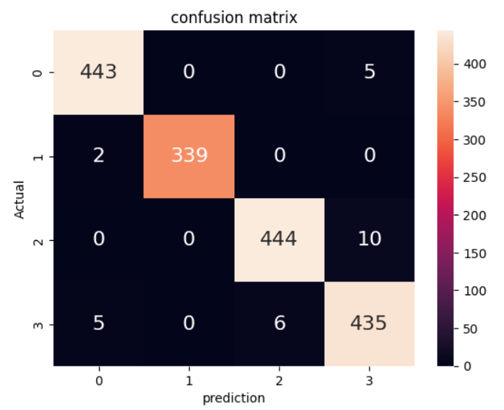


Fig. A. Confusion Matrix

A graph with red lines

Description automatically generated

Fig. B. Training and Validation Loss

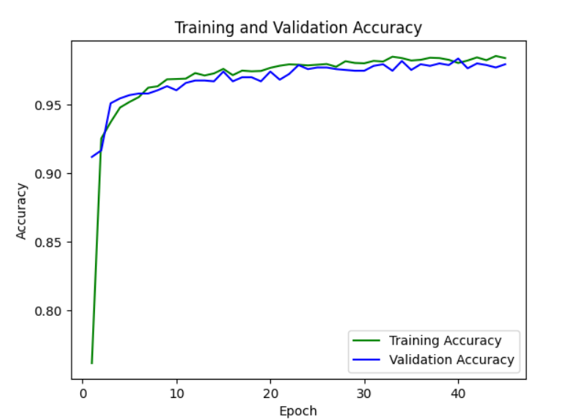


Fig. C. Training and Validation Accuracy

A screenshot of a computer

Description automatically generated

Fig. D. Precision, recall, f1-score

**YOLOv8 :** YOLOv8, the latest iteration of the You Only Look Once (YOLO) series, represents a significant advancement in real-time object detection. Building upon the successes of previous YOLO models, YOLOv8 incorporates several key improvements to enhance accuracy and efficiency. This version features an upgraded backbone network, which improves feature extraction capabilities, and introduces an advanced detection head that refines object localization and classification. YOLOv8 utilizes a combination of convolutional and transformer-based layers, optimizing both speed and precision for detecting objects in images and videos. The model is designed for high performance in various applications, from autonomous driving to surveillance, and maintains the YOLO series' hallmark of real-time processing with high accuracy. Its architecture supports comprehensive preprocessing techniques, including resizing, normalization, and augmentation, ensuring robust and reliable performance across diverse scenarios. YOLOv8 continues to push the boundaries of object detection technology, setting new standards for speed and accuracy in the field.

A blue squares with white text

Description automatically generated

A blue squares with white text

Description automatically generated Fig. A. Confusion Matrix

Fig. B. Training and Validation Loss

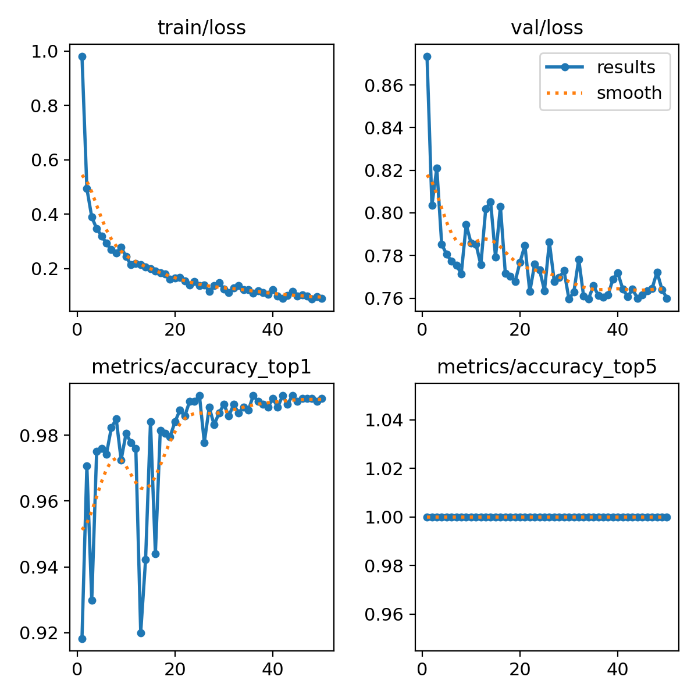


Fig. C. Results

IV. RESULTS

The Satellite image classification done through the various algorithms that are used in our project provides different accuracies. The Resnet50 algorithm provides with an accuracy of 48%, VGG19 algorithm provides with an accuracy of 85%, Xception algorithm on the other hand is 98% accuracy. But the YOLOV8 model used in our project provides the best results with the accuracy of 99%. The accuracy is achieved within the 20-50 epochs.

A graph of red bars

Description automatically generated with medium confidence

Fig. 2. Comparative analysis

V. CONCLUSION

In our research on cloud detection using satellite images, we employed ResNet-50, VGG-19, Xception, and YOLO v8. Among these, YOLO v8 demonstrated the highest accuracy. This trend is supported by literature, highlighting YOLO's growing popularity for detecting clouds in satellite images. Studies indicate that YOLO and other deep learning techniques excel in locating and removing clouds from satellite photos. Enhancing accuracy by incorporating atmospheric models, varied locations, and time frames has proven effective.

Future work in cloud identification can focus on incorporating sophisticated atmospheric models, expanding datasets to cover various geographic locations and time frames, and optimizing algorithms for real-time applications. Utilizing multispectral and hyperspectral imagery can provide additional data for more accurate cloud identification. Combining data from multiple sensors, such as radar and lidar, with optical imagery can enhance detection capabilities. Additionally, applying transfer learning techniques can help adapt models to new environments, improving their generalizability and applicability across different regions. These advancements can significantly improve atmospheric monitoring and understanding.

VI. REFERENCES

[1] Kalkan, M., Bostanci, G. E., Guzel, M., Kalkan, B., Özsarı, Ş., Soysal, Ö., & Köse, G. (2022). Cloudy/clear weather classification using deep learning techniques with cloud images. *Computers and Electrical Engineering* vol. 102, p. 108271, doi: 10.1016/j.compeleceng.2022.108271.

[2] Digra, M., Dhir, R., & Sharma, N. (2022). Land use land cover classification of remote sensing images based on the deep learning approaches: a statistical analysis and review. *Arabian Journal of Geosciences 2022 15:10*, *15*(10), 1–24. https://doi.org/10.1007/S12517-022-10246-8

[3] Ji, S., Dai, P., Lu, M., & Zhang, Y. (2020). Simultaneous Cloud Detection and Removal From Bitemporal Remote Sensing Images Using Cascade Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, pp. 1-17, doi: 10.1109/TGRS.2020.2994349.

[4] Zhou, N., Li, X., Shen, Z., Wu, T., & Luo, J. (2020). Geo-Parcel-Based Change Detection Using Optical and SAR Images in Cloudy and Rainy Areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. PP, pp. 1-1, doi: 10.1109/JSTARS.2020.3038169.

[5] M. Segal-Rozenhaimer, A. Li, K. Das, and V. Chirayath, "Cloud detection algorithm for multi-modal satellite imagery using convolutional neural-networks (CNN)," Remote Sensing of Environment, vol. 237, p. 111446, 2020, doi: 10.1016/j.rse.2019.111446.

[6] S. Mahajan and B. Fataniya, "Cloud detection methodologies: variants and development—a review," *Complex & Intelligent Systems*, vol. 6, pp. 1-11, 2019, doi: 10.1007/s40747-019-00128-0.

[7] J. Yang, J. Guo, H. Yue, Z. Liu, H. Hu, and K. Li, "CDnet: CNN-Based Cloud Detection for Remote Sensing Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, pp. 1-17, 2019, doi: 10.1109/TGRS.2019.2904868.

[8] Q. Li, W. Lyu, J. Yang, and J. Wang, "Thin Cloud Detection of All-Sky Images Using Markov Random Fields," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, pp. 417-421, 2012, doi: 10.1109/LGRS.2011.2170953.

[9] P. Panchal and R. Gupta, "Advancement of cloud detection algorithm in satellite images with application to color models," 2015, doi: 10.13140/RG.2.1.2449.3285.

[10] H. Li, H. Zheng, C. Han, H. Wang, and M. Miao, "Onboard Spectral and Spatial Cloud Detection for Hyperspectral Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 1, pp. 38-50, Jan. 2018, doi: 10.1109/TGRS.2017.2756031.

[11] S. H. Arun, S. Chaurasia, A. Misra, and R. Kumar, "Fog Stability Index: A novel technique for fog/low clouds detection using multi-satellites data over the Indo-Gangetic plains during winter season," *Journal of Atmospheric and Oceanic Technology*, vol. 35, no. 1, pp. 1-15, Jan. 2018, doi: 10.1175/JTECH-D-17-0074.1.

[12] N. Hayatbini, K.-l. Hsu, S. Sorooshian, Y. Zhang, and F. Zhang, "Effective Cloud Detection and Segmentation using a Gradient-Based Algorithm for Satellite Imagery; Application to improve PERSIANN-CCS," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2938-2952, May 2019, doi: 10.1109/TGRS.2018.2875120.

[13] Y. Changhui, Y. Yuan, M. Minjing, and Z. Menglu, "Cloud Detection Method Based on Feature Extraction in Remote Sensing Images," *International Journal of Remote Sensing*, vol. 34, no. 7, pp. 2387-2404, May 2013, doi: 10.1080/01431161.2012.741430.

[14] D. Frantz, E. Ha, A. U., J. Stoffels, and J. Hill, "Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects," *Remote Sensing*, vol. 10, no. 1, p. 38, Sep. 2018, doi: 10.3390/rs10010038.

[15] X. Song, Z. Liu, and Y. Zhao, "Cloud Detection and Analysis of MODIS Multispectral Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 9, pp. 1965-1975, Sep. 2004, doi: 10.1109/TGRS.2004.835424.

[16] L. Sun, X. Liu, Y. Yang, T. Chen, Q. Wang, and X. Zhou, "A cloud shadow detection method combined with cloud height iteration and spectral analysis for Landsat 8 OLI data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 141, pp. 135-148, Apr. 2018, doi: 10.1016/j.isprsjprs.2018.04.004.

[17] H. Ishida, Y. Oishi, K. Morita, K. Moriwaki, and T. Y. Nakajima, "Development of a support vector machine-based cloud detection method for MODIS with the adjustability to various conditions," *Remote Sensing of Environment*, vol. 195, pp. 162-174, Jan. 2017, doi: 10.1016/j.rse.2017.04.015.

[18] Z. Li, H. Shen, Y. Wei, Q. Cheng, and Q. Yuan, "Cloud Detection by Fusing Multi-Scale Convolutional Features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4523-4537, Aug. 2018, doi: 10.1109/TGRS.2018.2794531.

[19] Y. Oishi, H. Ishida, T. Y. Nakajima, R. Nakamura, and T. Matsunaga, "Preliminary verification for application of a support vector machine-based cloud detection method to GOSAT-2 CAI-2," *Journal of Atmospheric and Oceanic Technology*, vol. 35, no. 5, pp. 1063-1077, May 2018, doi: 10.1175/JTECH-D-17-0138.1.

[20] X. Wu and Z. Shi, "Utilizing Multilevel Features for Cloud Detection on Satellite Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8497-8512, Nov. 2018, doi: 10.1109/TGRS.2018.2839937.

[21] [21] E. Ricciardelli, F. Romano, and V. Cuomo, "Physical and statistical approaches for cloud identification using Meteosat Second Generation-Spinning Enhanced Visible and Infrared Imager Data," *Remote Sensing of Environment*, vol. 112, no. 4, pp. 1590-1604, Apr. 2008, doi: 10.1016/j.rse.2007.08.020.

[22] Y. Chen, R. Fan, M. Bilal, X. Yang, J. Wang, and W. Li, "Multilevel Cloud Detection for High-Resolution Remote Sensing Imagery Using Multiple Convolutional Neural Networks," *Remote Sensing*, vol. 10, no. 5, pp. 764-779, May 2018, doi: 10.3390/rs10050764.

[23] C. Deng, Z. Li, W. Wang, S. Wang, L. Tang, and A. C. Bovik, "Cloud Detection in Satellite Images Based on Natural Scene Statistics and Gabor Features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 803-815, Apr. 2019, doi: 10.1109/TGRS.2018.2863155.

[24] N. Ghasemian and M. Akhoondzadeh, "Introducing two Random Forest based methods for cloud detection in remote sensing images," *Remote Sensing*, vol. 10, no. 6, pp. 929-944, Jun. 2018, doi: 10.3390/rs10060929.