iJPublishing

Aerospace Engineering Vol: 1, No 3, 2024, Page: 1-11

Deep Learning-Based Classification of Remote Sensing Images: Challenges, Techniques, and Future Directions in Global Sustainability

Biplov Paneru[1], Bishwash Paneru[2], Sanjog Chhetri Sapkota[3]

[1)] Department of Electrical and Communication, Nepal Engineering College, Affiliated to Pokhara University, Bhaktapur, Nepal

[2)] Department of Applied Science Engineering, Institute of Engineering Pulchowk Campus, Affiliated to Tribhuvan University, Lalitpur, Nepal

[3)] Nepal Research and Collaboration Center, Nepal

DOI: https://doi.org/10.47134/aero.v1i3.2772
*Correspondence: Biplov Paneru
Email: thebplvstar001@gmail.com

Received: 01-05-2024
Accepted: 15-06-2024
Published: 31-07-2024

Abstract: With its high accuracy and efficiency, deep learning has greatly improved the classification of remote sensing (RS) photos. In order to categorize RS photos, this research analyzes the effectiveness of three cutting-edge deep learning models: ResNet-50, EfficientNetB2, and MobileNetV2. The models' accuracy on training and validation data were noted after they were trained and assessed on a dataset containing a variety of situations. Our findings illustrate each model's advantages and disadvantages and shed light on how well suited each is for various RS image categorization applications. The ResNet-50 model performed well in our study, achieving 74.41% training accuracy and 75.00% validation accuracy. With a training accuracy of 74.66% and a higher validation accuracy of 80.33%, the EfficientNetB2 model performed marginally better, demonstrating its strong generalization capabilities. On the other hand, the MobileNetV2 model had severe overfitting, as evidenced by its validation accuracy of 22.79%, which was much lower than its extraordinary high training accuracy of 99.21%. In order to achieve balanced performance between training and validation datasets in remote sensing image classification tasks, these results emphasize the significance of model architecture and regularization strategies. The proposed model can be utilized for sustainable remote sensing based applications in global water, environment and air health.

Keywords: Remote Sensing, GIS, EfficientNetB2, MobileNetV2, ResNet-50

Introduction

The process of gathering and evaluating information about the Earth's surface from satellite or aerial sensors is known as remote sensing (RS). Applications including agriculture, urban planning, disaster management, and environmental monitoring depend on the interpretation of RS images (Ahmed, 2024; Sadr, 2024). Conventional techniques for classifying RS images frequently depended on labor-intensive manual feature extraction and antiquated machine learning algorithms, which were unable to handle the volume and complexity of RS data (Canicattì, 2024; Nath, 2024). The analysis of remote sensing (RS) images and its numerous applications have advanced significantly in recent years. The need for automatic RS picture interpretation is growing since these images are more widely available than previously. The benchmark datasets are necessary preconditions in this regard for creating and evaluating intelligent interpretation algorithms (Akbar, 2024; Wang, 2024).

With the ability to automatically extract features and understand intricate patterns from massive datasets, recent advances in deep learning have completely changed the field of RS image categorization (Grund, 2024; Zekrifa, 2024). In particular, Convolutional Neural Nets (CNNs) have shown remarkable performance in image categorization tasks. In order to categorize RS photos, this research compares the effectiveness of three well-known deep learning models: ResNet-50, EfficientNetB2, and MobileNetV2.

Figure 1. Remote sensing in sustainability

We give an in-depth examination of the findings, present the outcomes, and talk about the approaches that were employed (Almashnowi, 2024; Manchanda, 2024). The implemented model can be used to support decision-making processes and advance sustainable development in a variety of sustainable activities, including environmental monitoring, land cover classification, water resource management, and climate change analysis. Overall, the created model's useful uses and promising future directions highlight how important it is as a potent instrument for tackling environmental issues and achieving sustainability objectives. The model can play a major role in building a more resilient and sustainable future for our planet with further study, innovation, and cross-disciplinary collaboration (Avtar et al., 2020).

Literature Review

This research uses Sentinel-2 satellite pictures and deep learning algorithms to identify and map rural villages in the Souss-Massa region. The authors first evaluated the UNet convolutional neural network architecture's outcome. The authors then adjusted the number of filters in the convolution layer to assess the effect of filter number on UNet performance. Thirdly, the authors put the deep Residual UNet (ResUNet) into practice. Specific metrics including accuracy, precision, recall, F1-score, and the ROC curve are used to assess the quality of tested models. With more filters, the UNet algorithm achieves an accuracy of 87% and an F1-score of 54%, outperforming the other two algorithms, ResUNet and UNet, which have precisions of 81.2% and 86.2%, respectively (Wahbi et al., 2023).

Recent years have seen the emergence of new deep learning (DL) techniques, leading to major advancements in approaches for RS image classification with DL. These developments present exciting new avenues for RS image classification research and development. This paper begins with a quick summary of common DL models (Hamilton, 2024; Patil, 2024; Rasheed, 2024). A thorough analysis of pixel-by-pixel and scene-by-scene RS image classification techniques based on the use of DL comes next. Additionally, a comparison of the effectiveness of common DL-based RS techniques is included. Lastly, the difficulties and possible paths for more study are spoken about (Li et al., 2018).

The three basic CNN models—ResNet50, ResNet101, and GoogleNet—are already trained. They have more sequence layers added to them in relation to CNN, and LAB channel operations are used for pre-processing data. Using GoogleNet over the pre-processed dataset, the best accuracy of 99.68%, precision of 99.42%, recall of 99.51%, and F-Score of 99.45% are attained. When the suggested work is compared to the most advanced techniques, it is found that for medium-sized datasets, adding additional layers to CNN does not always result in superior results. The 50 layers CNN (ResNet50) and the 101 layers CNN (ResNet101) are outperformed by the 22-layer CNN GoogleNet (Yadav et al., 2024).

Among the CNN-based models investigated are InceptionV3, DenseNet, ResNet, VGG, and EfficientNet. Three publicly accessible datasets with categories of images—the EuroSAT, UCMerced-LandUse, and NWPU-RESISC45—were used to assess the models. In terms of accuracy, recall, precision, and F1-score, the models show promise. The ability of Deep Learning techniques to understand the intricate and non-uniform characteristics of high-resolution remote sensing photos is demonstrated by this performance (Adegun et al., 2023).

In order to categorize items and facilities from the IARPA Functional Map of the World (fMoW) dataset into 63 distinct classes, the authors present a deep learning method. The system is made up of additional neural networks and a group of convolutional neural networks that combine image attributes and satellite metadata. The deep learning frameworks Keras and TensorFlow are utilized in its Python implementation, which operates on an NVIDIA Titan X graphics card-equipped Linux server. The system is now ranked #2 in the fMoW TopCoder competition. Its F 1 score is 0.797, its overall accuracy is 83%, and it correctly classifies 15 of the classes with accuracy of 95% or above (Pritt & Chern, 2017).

The Fast Fourier Transform (FFT) algorithm is used to convert the data into a spectral signature after preprocessing the dataset images. Next, each image's data is reduced by picking the top 20 features and applying the Vector Quantization (VQ) algorithm to convert them from a two-dimensional matrix to a one-dimensional vector matrix. There are two types of data: testing and training. Next, information is input into deep neural networks (DNN) with 23 layers for the purpose of classifying satellite photos. There are 2,145,020 parameters in the result, and accuracy, loopback, and F1 were all evaluated as 100% of the performance metrics (Satellite Image Classification using Spectral Signature and Deep Learning, 2023).

For satellite image categorization, a convolutional neural network design with a scaling technique is suggested. Using a compound coefficient, the scaling approach can scale depth, width, and resolution all equally. It can be applied as a first step in monitoring, satellite surveillance, urban planning, etc. Systems for monitoring ships and geo-information can also benefit from it. The suggested approach is predicated on a scalable, end-to-end interpretation of satellite images. It divides these into four groups using the spatial data from satellite photos. The suggested approach yields positive and promising outcomes on a difficult dataset with a high degree of intra- and inter-class heterogeneity. The RSI-CB256 dataset demonstrates 99.64% accuracy for the suggested technique (Tehsin et al., 2023).

With its fully connected layers (FC-1024), batch normalization (BN), L2 regularization, dropout layers, dense layer, and data augmentation, this research presents a novel snapshot-based residual network (SnapResNet). While data augmentation addresses the issue of imbalanced classes, architectural modifications address the problem of inter-class similarity. The network's performance is further enhanced by avoiding over-fitting with the use of the snapshot ensemble technique. The most difficult Large-Scale Cloud photos Dataset for Meteorology Research (LSCIDMR) is used by the suggested SnapResNet152 model to classify thousands of high-resolution photos into ten classifications. With an overall accuracy of 97.25%, the proposed model beats the current deep learning-based methods (e.g., AlexNet, VGG-19, ResNet101, and EfficientNet) (Yousaf et al., 2023).

We propose 4 different models for the purpose of classifying satellite Remote Sensing Images. The images are trained and analysed using the deep learning transfer learning algorithms for the purpose of classification. The dataset obtained from Shoshany et al. (2013) is utilized for the purpose of classisfying remote sensing images. The images captured from various areas can be a good aspect for promoting sustainability using remote sensing and deep learning based technique so that in future the obtained model can be used for AI based treatment on environmental hazards. We also describe some concepts of real life model implementation for sustainable practices.

Methodology

Dataset Description

The photos in the dataset utilized for this study are divided into numerous classes, each of which represents a different scenario. The preprocessed photos were divided into training and validation sets after being uniformly sized at 256 by 256 pixels.

Table 1. Distribution of Files by Class

Class	Number of Files
Cloudy	1500
Desert	1131
Green Area	1500
Water	1500

The distribution of files inside the dataset by class is shown in this table, along with the quantity of files that are available for each class.

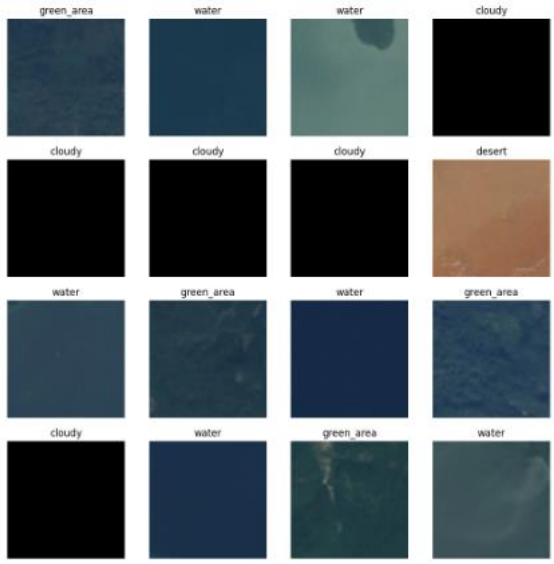

Figure 2. Sample dataset of satellite images

Model Development

ResNet-50: A 50-layer deep neural network, renowned for its residual connections that aid in mitigating the issue of disappearing gradients. In image classification challenges, the deep convolutional neural network architecture ResNet-50 has shown exceptional performance. Its 50-layer deep structure and creative use of residual connections, which aid in resolving the vanishing gradient issue, define it. ResNet-50 is especially well-suited for difficult tasks like remote sensing image classification because of its architecture, which allows it to learn rich and complicated features from images. ResNet-50's usage of residual connections, which enable the training of deep networks without the degradation issue, allows it to remain somewhat efficient despite its depth.

EfficientNetB2: EfficientNetB2 belongs to the EfficientNet model family, which is renowned for its exceptional balance between efficiency and accuracy. The "Efficient" in EfficientNet comes from the model's special scaling technique, which uses a compound coefficient to scale the network's depth, width, and resolution equally. When compared to other models, EfficientNet models can attain state-of-the-art performance with a substantially less number of parameters because to this scaling strategy. In particular, EfficientNetB2 is built for efficiency as well as accuracy, which makes it a strong option for environments with limited resources or applications that need high-performance models without incurring excessive computational expenses.

MobileNetV2: A lightweight neural network architecture called MobileNetV2 was created with embedded and mobile vision applications in mind. Its use of depthwise separable convolutions, which drastically lower the number of calculations and parameters as compared to conventional convolutions, is what makes it unique. Because of its architectural design, MobileNetV2 excels in situations when computing power is scarce, like on edge or mobile devices. Even while MobileNetV2 is lightweight, it can nevertheless perform well on tasks like picture classification, which makes it a popular option for real-time applications where speed is critical.

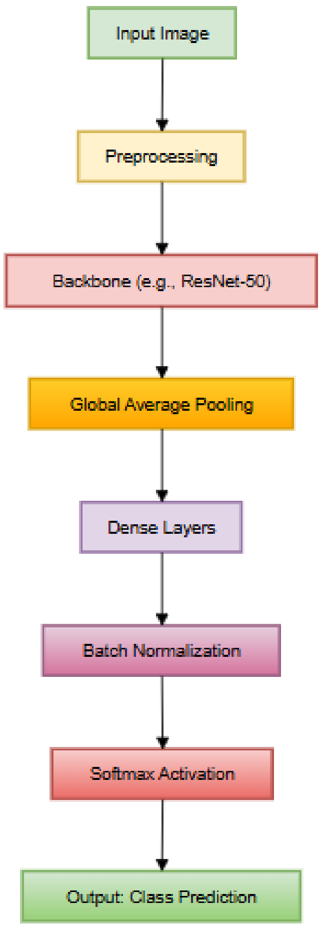

Figure 3. Proposed architecture

Common Architecture Components of Models

Table 2. Common architecture components of models

Component	Description
Input Layer	All models take input of shape (256, 256, 3).
GlobalAveragePooling2D	Applied after the backbone to reduce dimensions.
Dense Layers	Two dense layers with ReLU activation, followed by a final dense layer with softmax activation for classification into 4 classes.
BatchNormalization	Applied after the first dense layer to normalize activations.

Training and Evaluation Metrics used

- Loss Function: Categorical Crossentropy
- Optimizer: Adam with a learning rate of 0.01
- Metrics: Accuracy and Top-K Accuracy

Results and Discussions

ResNet-50

ResNet-50 obtained 75.00% validation accuracy and 74.41% training accuracy. Although it performed well on the training set, its validation accuracy indicates that it did not generalize to previously unseen data as well. This suggests that overfitting may occur, which might be avoided by using methods like dropout or data augmentation.

EffectiveNetB2

Training accuracy of 74.66% and validation accuracy of 80.33% were attained by EfficientNetB2. This model showed strong generalization capabilities by balancing training and validation performance well. The higher performance of EfficientNetB2 on the validation set can probably be attributed to its architecture, which strikes a balance between network depth, width, and resolution.

MobileNetV2

MobileNetV2 only managed a 22.79% validation accuracy compared to an astonishing 99.21% training accuracy. Severe overfitting, or the model learning the training data too well but failing to generalize to the validation data, is indicated by this large mismatch. It's possible that MobileNetV2's lightweight design was insufficient to fully capture the intricacy of the RS pictures in this dataset.

Table 3. Final model accuracies

Model	Training Accuracy	Validation Accuracy
ResNet-50	0.7441	0.7500
EfficientNetB2	0.7466	0.8033
MobileNetV2	0.9921	0.2279

Practical Applications and future Prospects

The created model can be extremely helpful in implementing sustainable practices because it offers insightful data and insights from satellite photography. It can be used, for instance, to follow changes in cloud cover patterns that may have an impact on weather and climate patterns, monitor deforestation in ecosystems that are vulnerable, evaluate the effects of urbanization on green spaces, and identify places that are vulnerable to pollution or water scarcity. The model helps to conserve and preserve the environment for future generations by utilizing deep learning and satellite imagery to enable more effective and efficient management of natural resources.

Such a model can also help with response and management during disasters. Its ability to analyze satellite imagery in afflicted areas allows it to immediately determine the magnitude of natural disasters like hurricanes, wildfires, and floods. By planning evacuation routes, allocating resources effectively, and prioritizing locations for aid and recovery activities, emergency responders can lessen the impact of disasters on infrastructure and human lives (Rao, 1991).

The model can offer useful insights for sustainable farming methods in the field of agriculture. It can evaluate crop health, track soil moisture content, and forecast yield potential by evaluating satellite data. With this information, farmers can reduce their environmental impact while increasing crop yield. They can also use it to plan irrigation schedules more effectively, use herbicides and fertilizers more effectively, and make well-informed decisions.

There are a number of intriguing opportunities for the model's continued application and improvement in the future. New developments in satellite images, such as increased spectral and spatial resolution, will improve the model's capacity to extract more precise information and provide more precise forecasts (Rao, 1991). Integrating the model with additional data sources, such meteorological information, measurements of soil quality, and socioeconomic variables, can enhance its analysis and facilitate more thorough decision-making (Avtar et al., 2020).

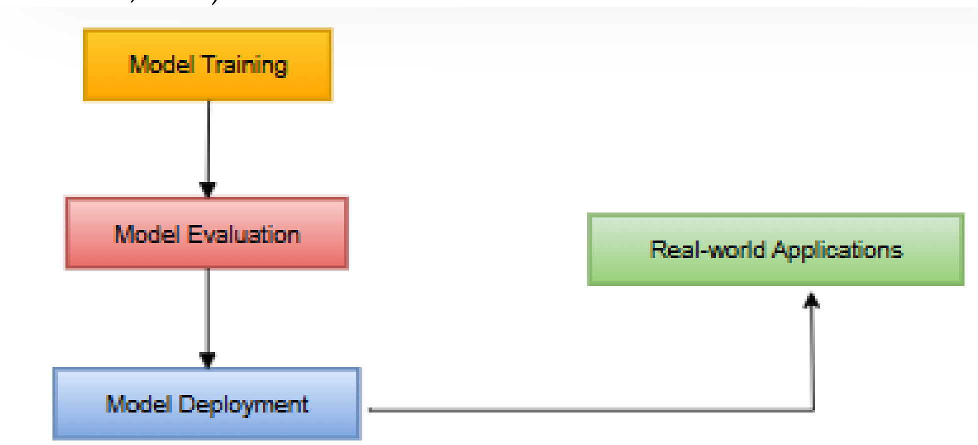

Figure 4. Real world application with AI and GIS integration

Conclusions

The pros and cons of each model for RS image categorization are shown by the comparison of ResNet-50, EfficientNetB2, and MobileNetV2. Both ResNet-50 and EfficientNetB2 performed admirably, with EfficientNetB2 having the best capacity for generalization. Even with its excellent training accuracy, MobileNetV2 suffered from overfitting, which made it less effective for this task in the absence of additional regularization strategies.

Subsequent research endeavours may encompass investigating broader data augmentation techniques, optimising hyperparameters, and integrating transfer learning from models that have been pre-trained on a wider range of RS image datasets. Further testing of these models on various RS image datasets with differing class distributions and levels of complexity would offer more thorough insights into their functionality. Additionally, such AI model with sensors and IoT techniques integrating can support international efforts aimed at mitigating and adapting to climate change. It can assist in tracking progress toward climate targets and provide information for policy interventions aimed at reducing greenhouse gas emissions and promoting sustainable development by keeping an eye on environmental indicators like rates of deforestation, carbon emissions, and changes in land use.

References

Adegun, A. A., Viriri, S., & Tapamo, J. R. (2023). Review of deep learning methods for remote sensing satellite images classification: Experimental survey and comparative analysis. *Journal of Big Data, 10*, 93. https://doi.org/10.1186/s40537-023-00772-x

Ahmed, A. A. (2024). Applications of machine learning to water resources management: A review of present status and future opportunities. *Journal of Cleaner Production*, *441*. https://doi.org/10.1016/j.jclepro.2024.140715

Akbar, J. U. M. (2024). A Comprehensive Review on Deep Learning Assisted Computer Vision Techniques for Smart Greenhouse Agriculture. *IEEE Access*, *12*, 4485–4522. https://doi.org/10.1109/ACCESS.2024.3349418

Almashnowi, M. Y. A. (2024). A Hybrid Deep Learning-Based Power Management Strategy for PV-Assisted Desalination Plant. *Particle and Particle Systems Characterization*. https://doi.org/10.1002/ppsc.202400053

Avtar, R., Komolafe, A. A., Kouser, A., Singh, D., Yunus, A. P., Dou, J., Kumar, P., Gupta, R. D., Johnson, B. A., Minh, H. V. T., Aggarwal, A. K., & Kurniawan, T. A. (2020). Assessing sustainable development prospects through remote sensing: A review. *Remote Sensing Applications, 20*, 100402. https://doi.org/10.1016/j.rsase.2020.100402

Canicattì, M. (2024). Drones in vegetable crops: A systematic literature review. *Smart Agricultural Technology*, *7*. https://doi.org/10.1016/j.atech.2024.100396

Dataset obtained from: Mahmoudreda55. "Satellite Image Classification Dataset." Kaggle, 2022, https://www.kaggle.com/datasets/mahmoudreda55/satellite-image-classification

Grund, J. (2024). Emotions and transformative learning for sustainability: a systematic review. *Sustainability Science*, *19*(1), 307–324. https://doi.org/10.1007/s11625-023-01439-5

Hamilton, J. R. (2024). Adding External Artificial Intelligence (AI) into Internal Firm-Wide Smart Dynamic Warehousing Solutions. *Sustainability (Switzerland)*, *16*(10). https://doi.org/10.3390/su16103908

Li, Y., Zhang, H., Xue, X., Jiang, Y., & Shen, Q. (2018). Deep learning for remote sensing image classification: A survey. *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery, 8*(6). https://doi.org/10.1002/widm.1264

Manchanda, M. (2024). Optimizing Farm Insect Image Classification: A Transfer Learning Approach with Model Selection. *7th International Conference on Inventive Computation Technologies, ICICT 2024*, 657–662. https://doi.org/10.1109/ICICT60155.2024.10544571

Nath, P. C. (2024). Recent advances in artificial intelligence towards the sustainable future of agri-food industry. *Food Chemistry*, *447*. https://doi.org/10.1016/j.foodchem.2024.138945

Patil, R. (2024). Role of Data-Centric Artificial Intelligence in Agriculture. *Data-Centric Artificial Intelligence for Multidisciplinary Applications*, 159–176. https://doi.org/10.1201/9781003461500-13

Pritt, M., & Chern, G. (2017). Satellite image classification with deep learning. In *2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, Washington, DC, USA, pp. 1-7. https://doi.org/10.1109/AIPR.2017.8457969

Rao, U. R. (1991). Remote sensing for sustainable development. *Journal of Indian Society of Remote Sensing, 19*, 217–235. https://doi.org/10.1007/BF03023969

Rasheed, K. (2024). Industry 4.0 and Construction. *AI in Material Science: Revolutionizing Construction in the Age of Industry 4.0*, 65–87. https://doi.org/10.1201/9781003438489-4

Sadr, A. V. (2024). Operational greenhouse-gas emissions of deep learning in digital pathology: a modelling study. *The Lancet Digital Health, 6*(1). https://doi.org/10.1016/S2589-7500(23)00219-4

Satellite Image Classification using Spectral Signature and Deep Learning. (2023). *Iraqi Journal of Science, 64*(6), 3153-3163. https://doi.org/10.24996/ijs.2023.64.6.42

Shoshany, M., Long, D., & Bonfil, D. (2013). Remote sensing for sustainable agriculture. *International Journal of Remote Sensing, 34*(17), 6021–6023. https://doi.org/10.1080/01431161.2013.795004

Tehsin, S., Kausar, S., Jameel, A., Humayun, M., & Almofarreh, D. K. (2023). Satellite image categorization using scalable deep learning. *Applied Sciences, 13*, 5108. https://doi.org/10.3390/app13085108

Wahbi, M., Bakali, I. E., Ez-Zahouani, B., Azmi, R., Moujahid, A., Zouiten, M., Alaoui, O. Y., Boulaassal, H., Maatouk, M., & Kharki, O. E. (2023). A deep learning classification approach using high spatial satellite images for detection of built-up areas in rural zones: Case study of Souss-Massa region - Morocco. *Remote Sensing Applications, 29*, 100898. https://doi.org/10.1016/j.rsase.2022.100898

Wang, R. (2024). Enhancing biomass conversion to bioenergy with machine learning: Gains and problems. *Science of the Total Environment*, *927*. https://doi.org/10.1016/j.scitotenv.2024.172310

Yadav, D., Kapoor, K., Yadav, A. K., et al. (2024). Satellite image classification using deep learning approach. *Earth Science Informatics*, *17*, 2495–2508. https://doi.org/10.1007/s12145-024-01301-x

Yousaf, R., Rehman, H. Z. U., Khan, K., Khan, Z. H., Fazil, A., Mahmood, Z., Qaisar, S. M., & Siddiqui, A. J. (2023). Satellite imagery-based cloud classification using deep learning. *Remote Sensing*, *15*, 5597. https://doi.org/10.3390/rs15235597

Zekrifa, D. M. S. (2024). Advanced deep learning approach for enhancing crop disease detection in agriculture using hyperspectral imaging. *Journal of Intelligent and Fuzzy Systems*, *46*(2), 3281–3294. https://doi.org/10.3233/JIFS-235582